Stability Based on Single-Agent Deviations in Additively Separable Hedonic Games*

Felix Brandt[a], Martin Bullinger[b,*], Leo Tappe[a]

[a] *School of Computation, Information and Technology, Technische Universität München, Boltzmannstraße 3, 85748 Munich, Germany*
[b] *Department of Computer Science, University of Oxford, 7 Parks Road, Oxford OX1 3QD, United Kingdom*

Abstract

Coalition formation is a central concern in multiagent systems. A common desideratum for coalition structures is stability, defined by the absence of beneficial deviations of single agents. Such deviations require an agent to improve her utility by joining another coalition. On top of that, the feasibility of deviations may also be restricted by demanding consent of agents in the welcoming and/or the abandoned coalition. While most of the literature focuses on deviations constrained by unanimous consent, we also study consent decided by majority vote and introduce two new stability notions that can be seen as local variants of another solution concept called popularity. We investigate stability in additively separable hedonic games by pinpointing boundaries to computational complexity depending on the type of consent and friend-oriented utility restrictions. 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 a new combinatorial observation that we call the *Deviation Lemma* and that we leverage to prove the convergence of simple and natural single-agent dynamics under fairly general conditions. Our negative results, in particular, resolve the complexity of contractual Nash stability in additively separable hedonic games.

Keywords: Computational Social Choice, Algorithmic Game Theory, Hedonic Games, Single-Agent Stability, Dynamics, Computational Complexity

*This paper unifies and expands results that appeared in the Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI) (Brandt et al., 2022) and the Proceedings of the 47th International Symposium on Mathematical Foundations of Computer Science (MFCS) (Bullinger, 2022). Most of this research was done while the corresponding author was a PhD student at Technische Universität München.

*Corresponding author

1. Introduction

Coalition formation is a central topic in multi-agent systems. It is concerned with 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 applied to model problems from operations research and the mathematical social sciences, such as research team formation (Alcalde and Revilla, 2004), task allocation (Saad et al., 2011), or community detection (Aziz et al., 2019). Moreover, these games have been used in the context of clustering, an important task in machine learning (Feldman et al., 2015; Ahmadi et al., 2022). Identifying desirable coalition structures is often based on the prospect of coalitions staying 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 forming a new one.

In this paper, we focus on deviations by single agents who leave their current coalition to join another coalition or to form a new coalition on their own. Whenever such a deviation is not possible, we speak of *single-agent stability*. The simplest example of a single-agent deviation is a Nash deviation, where some agent unilaterally decides to leave her current coalition in order to join another coalition. Nash stability then captures partitions that do not admit Nash deviations. Consider a scenario with two agents, say, Alice and Bob, where Alice prefers to be in a coalition of her own, but Bob wants to be together with Alice. Then, in the partition where Alice and Bob both form singleton coalitions, Bob has an incentive to perform a Nash deviation to join the coalition of Alice. However, in the partition where Alice and Bob are in a joint coalition, Alice has an incentive to perform a Nash deviation to be in a coalition of her own. This describes a simple *run-and-chase* situation which can occur in most reasonable classes of hedonic games (and in all classes of hedonic games considered in this paper), but it already shows a defect of Nash stability: Nash-stable outcomes need not exist. Moreover, while a Nash deviation clearly captures the incentive of single agents to perform deviations, it completely ignores other agents' opinions about the deviation.

To overcome the shortcomings of Nash stability, 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. This is certainly reasonable in high-stakes coalition formation scenarios like joining the partners of a firm (Drèze and Greenberg, 1980; Meade, 1972). Moreover, unanimous consent is used in the formation process of international bodies like the EU or the NATO (see, e.g., Brandt et al., 2023). Still, it might be impractical and even undesirable in small- or medium-scale coalition formation problems, such as joining a music group or sports club. 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,[1] and we will also define new majority-based stability notions. Since a majority consent is equal to unanimous consent when joining a coalition of only one agent, this is already sufficient to circumvent the nonexistence of stable outcomes in the run-and-chase example described above: majority consent suggests that a reasonable outcome is that both players will end up in singleton coalitions.

Since the number of coalitions an agent can be part of is not polynomially bounded, a lot of effort has been put into identifying reasonable and succinct classes of hedonic games (see, e.g., Aziz et al., 2019; Ballester, 2004; Bogomolnaia and Jackson, 2002; Elkind and Wooldridge, 2009). In many such classes, agents extract cardinal preferences from a weighted graph by some aggregation method. Perhaps the most natural and thoroughly studied way to aggregate preferences is by taking the sum of the weights of edges towards agents in one's own coalition. This leads to the class of additively separable hedonic games (ASHGs) (Bogomolnaia and Jackson, 2002). ASHGs allow 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 goals can have higher priority than the other one (Dimitrov et al., 2006). Our work provides a computational analysis of single-agent stability, focusing on friend-oriented utility restrictions.

1.1. Contribution

A recent line of research on stability notions focuses on the dynamical aspects leading to the formation of stable outcomes (see, e.g., Bilò et al., 2018; Hoefer et al., 2018; Carosi et al., 2019; Brandt et al., 2023). This yields a critical distributed perspective on the coalition formation process. The value of some positive computational results in the context of hedonic games is diminished because 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 dynamics towards stable states (see, e.g., Bogomolnaia and Jackson, 2002; Suksompong, 2015; Brandt et al., 2023; Bullinger and Suksompong, 2024).

We extend the applicability of this approach by considering *nonmonotonic* 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. We demonstrate the power of this lemma via three applications, which in particular yield polynomial running time of dynamics in friend-oriented games for various stability concepts. The Deviation Lemma is not restricted to additively separable utilities or the specific type of single-agent deviations. For instance,

[1]The paper by Gairing and Savani (2019) is a notable exception.

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 an exponential worst-case running time of the dynamics. By contrast, our results hold for restricted sets of *nonsymmetric* 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.

In contrast to the positive results obtained by means of the Deviation Lemma, we also find strong computational boundaries. In the conclusion, Table 2 summarizes our and related complexity results. We obtain NP-hardness of the existence problem for Nash stability in severely restricted ASHGs as well as the existence problem of contractually Nash-stable coalition structures in general ASHGs. Despite knowing that additively separable hedonic games that do not admit a contractually Nash-stable coalition structure exist (Sung and Dimitrov, 2007), previous investigations of single-agent stability have left the complexity of the associated existence problem open (Sung and Dimitrov, 2010). Hence, we complete the picture of the complexity of unanimity-based single-agent stability concepts in ASHGs.

In addition, we also find computational boundaries for majority-based stability concepts. This complements the results obtained by the Deviation Lemma. Our results thus completely pinpoint the complexity of majority-based stability notions in appreciation-of-friends games, aversion-to-enemies games, and friends-and-enemies games. Notably, a major step towards these hardness results is the construction of No-instances, which can then be leveraged in hardness reductions.

Our results are in line with a repeatedly observed theme in hedonic games: the existence of counterexamples is the key to computational intractability (see, e.g., Dimitrov et al., 2006; Sung and Dimitrov, 2010; Aziz et al., 2013; Brandt et al., 2023).[2] On the other hand, we demonstrate that the observed intractabilities

[2] A notable exception is that partitions in the core of aversion-to-enemies games always exist but are hard to compute (Dimitrov et al., 2006). Bullinger and Kober (2021) also identify

lie at the computational boundary by carving out further weak restrictions that lead to the existence and efficient computability of stable states.

1.2. Related Work

The study of hedonic games was initiated by Drèze and Greenberg (1980) but was only popularized two decades later by Banerjee et al. (2001), Cechlárová and Romero-Medina (2001), and Bogomolnaia and Jackson (2002). Aziz and Savani (2016) provide an overview of many important concepts. Two important research questions concern the design of reasonable computationally manageable subclasses of hedonic games and the detailed investigation of their computational properties. The former has led to a broad landscape of game representations. Some of these representations are ordinal and fully expressive, i.e., they can, in principle, express every preference relation over coalitions (Ballester, 2004; Elkind and Wooldridge, 2009). Still, representing certain preference relations requires exponential space. These representations are contrasted by cardinal representations based on weighted graphs (Aziz et al., 2019; Bogomolnaia and Jackson, 2002; Olsen, 2012), which are not fully expressive but only require polynomial space (except when weights are disproportionately large). Apart from the already discussed additively separable hedonic games, important aggregation methods consider the average of weights leading to the classes of fractional hedonic games (Aziz et al., 2019) and modified fractional hedonic games (Olsen, 2012).

Computational properties of hedonic games have been studied extensively and we focus on literature related to additively separable hedonic games. Various versions of stability have been investigated (Dimitrov et al., 2006; Sung and Dimitrov, 2010; Aziz and Brandl, 2012; Aziz et al., 2013; Gairing and Savani, 2019). In particular, Sung and Dimitrov (2010) perform a detailed computational study of single-agent stability and Gairing and Savani (2019) settle the complexity of single-agent stability for symmetric input graphs. Apart from stability, other desirable axioms concern efficiency and fairness. Aziz et al. (2013) cover a wide range of axioms, whereas Elkind et al. (2020) and Bullinger (2020) focus on Pareto optimality, and Brandt and Bullinger (2022) investigate popularity, an axiom combining ideas from stability and efficiency, which is also related to a majority-based stability notion that we will introduce.

The dynamical aspects of the coalition formation process have been studied in a series of very recent papers (Bilò et al., 2018; Hoefer et al., 2018; Carosi et al., 2019; Fanelli et al., 2021; Brandt et al., 2023; Boehmer et al., 2023; Bullinger and Suksompong, 2024). Most related is the work by Bilò et al. (2018), who consider Nash stability in fractional hedonic games, and by Brandt et al. (2023), who consider dynamics based on individual stability in several classes of hedonic games. Bullinger and Suksompong (2024) consider a generalization of additively separable hedonic games and a stability concept analogous to Nash stability. Hoefer et al. (2018); Carosi et al. (2019), and Fanelli et al. (2021)

a class of hedonic games with this property.

consider dynamics based on group deviations. Finally, Boehmer et al. (2023) propose a dynamical version of hedonic games where utilities are modified based on the history of the performed deviations. They study both single-agent and group stability. Similar dynamic processes have been studied in the domain of matchings (see, e.g., Roth and Vande Vate, 1990; Abeledo and Rothblum, 1995; Brandt and Wilczynski, 2019). Moreover, another dynamic aspect of hedonic games is captured by online hedonic games, where agents arrive in sequence, and have to be added to existing coalitions at their arrival. Online additively separable hedonic games have been introduced by Flammini et al. (2021) and subsequently been studied by Bullinger and Romen (2023, 2024). In particular, Bullinger and Romen (2024) study stability in the online model of additively separable hedonic games.

2. Preliminaries and Model

In this section, we introduce hedonic games and stability concepts. In the final part, we outline our general strategy to obtain computational hardness results. We use the notation $[k] = \{1, \dots, k\}$ for any positive integer k.

2.1. Hedonic Games

Throughout the paper, we consider settings with a set 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. A *partition* of N is 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 $C \cap D = \emptyset$. A nonempty subset of N is called a *coalition*. Hence, every element of a partition is a coalition, and given a partition π, we denote by $\pi(i)$ the coalition containing agent i. 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 strict part of an order $\succsim_i$ is denoted by $\succ_i$, i.e., $C \succ_i D$ if and only if $C \succsim_i D$ and not $D \succsim_i C$.

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 $v_i : N \to \mathbb{R}$ for every agent i. 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 $v_i : N \to \mathbb{R}$ such that $\pi(i) \succsim_i \pi'(i)$ if and only if $\sum_{j\in\pi(i)} v_i(j) \geq \sum_{j\in\pi'(i)} v_i(j)$. 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).

Every ASHG can be naturally represented 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 subclasses of ASHGs that allow a natural interpretation in terms of friends and enemies. An agent $j \in N$ is called a *friend* (or *enemy*) of agent $i \in N$ if $v_i(j) > 0$ (or $v_i(j) < 0$). An ASHG is called a *friends-and-enemies game* (FEG) if $v_i(j) \in \{-1, 1\}$ for every pair of agents $i, j \in N$. Further, following Dimitrov et al. (2006), an ASHG is called an *appreciation-of-friends game* (AFG) (or an *aversion-to-enemies game* (AEG)) if $v_i(j) \in \{-1, n\}$ (or $v_i(j) \in \{-n, 1\}$). In all of these games, agents aim 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. Based on the friendship of agents, we define the *friendship relation* (or *enemy relation*) as the subset $R \subseteq N \times N$ where $(i, j) \in R$ if and only if $v_i(j) > 0$ (or $v_i(j) < 0$).

2.2. Stability Based on Single-Agent Deviations

We focus on stability notions that are concerned with the incentives of single agents to deviate. 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 $j \neq i$, it holds that $\pi(j) \setminus \{i\} = \pi'(j) \setminus \{i\}$. We write $\pi \xrightarrow{i} \pi'$ to denote a single-agent deviation performed by agent i transforming partition π to partition π'.

We consider myopic agents whose rationale is to only engage in a deviation if it immediately makes them better off. A *Nash deviation* is a single-agent deviation performed by agent i making her better off, i.e., $\pi'(i) \succ_i \pi(i)$. Any partition in which no Nash deviation is possible is said to be *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 $i \in N$ 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 C rather than outside, i.e., $F_{\text{in}}(C, i) = \{j \in C \setminus \{i\} \colon C \cup \{i\} \succ_j C \setminus \{i\}\}$. 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 C rather than inside, i.e., $F_{\text{out}}(C, i) = \{j \in C \setminus \{i\} \colon C \setminus \{i\} \succ_j C \cup \{i\}\}$.

An *individual deviation* (or *contractual deviation*) is a Nash deviation $\pi \xrightarrow{i} \pi'$ such that $F_{\text{out}}(\pi'(i), i) = \emptyset$ (or $F_{\text{in}}(\pi(i), i) = \emptyset$). Then, a partition is said to be *individually stable* (IS) or *contractually Nash-stable* (CNS) if it allows for no individual or contractual deviation, respectively. A related weakening of both stability concepts is contractual individual stability (CIS), based on deviations that are both individual and contractual deviations (Bogomolnaia and Jackson, 2002; Dimitrov and Sung, 2007).

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 a *majority-in deviation* (or *majority-out deviation*) if $|F_{\text{in}}(\pi'(i), i)| \geq |F_{\text{out}}(\pi'(i), i)|$ (or $|F_{\text{out}}(\pi(i), i)| \geq |F_{\text{in}}(\pi(i), i)|$). A single-agent deviation that is both a majority-in deviation and a majority-out deviation is called *separate-majorities deviation*. Similar to before, a partition is said to be *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 deviation, respectively. Majority-in and majority-out stability are special cases of the voting-based stability notions by Gairing and Savani (2019) for a threshold of 1/2.

Finally, it is possible to relax separate-majorities stability 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. Joint-majority stability is particularly interesting as it is a natural local version of popularity, an axiom recently studied in the context of hedonic games (Gärdenfors, 1975; Cseh, 2017; Brandt and Bullinger, 2022).[3]

Also note that while contractual individual stability is a refinement of Pareto optimality, there is no logical relationship between other (majority-based) stability concepts and Pareto optimality. For simple terminology, we use abbreviations like NS or IS both as an adjective or a noun to refer to the stability concept. In addition, we denote the set of stability concepts based on single-agent deviations by $\mathcal{C}$, i.e., $\mathcal{C} = \{\text{NS}, \text{IS}, \text{CNS}, \text{CIS}, \text{MIS}, \text{MOS}, \text{SMS}, \text{JMS}\}$. A taxonomy of our and related solution concepts is provided in Figure 1. We refer to deviations with respect to stability concept $\alpha \in \mathcal{C}$ as α *deviations*, e.g., IS deviations for $\alpha = \text{IS}$. Similarly, we speak of α *partitions* if a partition satisfies α.

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 $\alpha \in \mathcal{C}$, 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 $\alpha \in \mathcal{C}$, 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 exis-

[3]Informally speaking, a partition is popular if there is no other partition preferred by a majority of all agents. By contrast, partitions can only be challenged by other partitions evolving through Nash deviations under joint-majority stability.

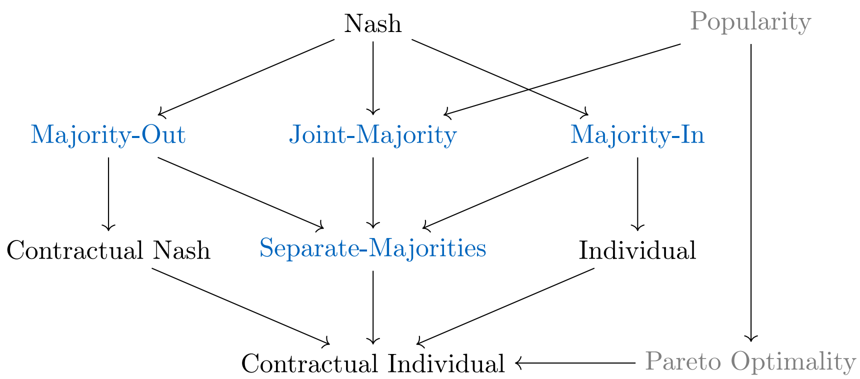

Figure 1: Logical relationships between stability notions and other solutions concepts. An arrow from concept α to concept β indicates that if a partition satisfies α, it also satisfies β. Majority-based stability notions are highlighted in blue, other single-agent based stability notions in black.

tence 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.

2.3. Computational Complexity

Our central question is to answer the following decision question regarding stable partitions for some stability concept $\alpha \in \mathcal{C}$.

> Given an additively separable hedonic game, possibly restricted to be from some subclass, does this game admit an α partition?

In this paper, we will derive a complete picture concerning the computational complexity of this question in all game classes introduced in Section 2.1 and all stability concepts in $\mathcal{C}$. In many cases, we will obtain computational intractabilities in the sense of NP-completeness. In this section, we want to describe the general scheme of our proof technique for these results.

First, note that for all of our stability notions, a stable partition is a polynomial-time verifiable certificate: one can simply check whether some agent can perform a deviation, and if no one can, then the partition is stable. Therefore, we omit the proof of membership in NP in all of our NP-completeness proofs.

For the NP-hardness of our problems, all our reductions are from the same NP-complete source problem, namely Exact Cover by 3-Sets (Karp, 1972). An instance of Exact Cover by 3-Sets (E3C) consists of a tuple $(R, \mathcal{S})$, where R is a ground set together with a set $\mathcal{S}$ of 3-element subsets of R. A Yes-instance is an instance such that there exists a subset $\mathcal{S}' \subseteq \mathcal{S}$ that partitions R. Given an instance $(R, \mathcal{S})$ of E3C, for every $r \in R$, we define $\mathcal{S}_r = \{S \in \mathcal{S} \colon r \in S\}$, i.e., $\mathcal{S}_r$ comprises the elements of $\mathcal{S}$ containing r, and $n_r = |\mathcal{S}_r|$.

While each of our reductions employs different ideas and gadgets specific to the considered subclass of ASHGs and stability concept, the reductions share the following central steps.[4]

1. Find a No-instance.
2. Encode the combinatorial structure of E3C.
3. Use the No-instance as a gadget.
4. Prove a correspondence of Yes-instances.

Of course, a reduction is a correspondence of Yes-instances and No-instances of the source problem and reduced problem, so finding No-instances is a vital step for making reductions work. In some cases, very easy No-instances exist. Consider, for instance, Nash stability, for which the run-and-chase example in the introduction can be reproduced in all of our considered game classes. However, this can become a nontrivial task for more restrictive solution concepts and game classes. For example, we have to put a large effort into designing FEGs without an MOS or MIS partition. The games that we construct in Propositions 6 and 8 contain a large number of agents, and it requires extensive arguments to verify that these games are, in fact, No-instances.

The second step usually consists of defining set gadgets for the 3-elementary subsets in $\mathcal{S}$ of a given instance $(R, \mathcal{S})$ of E3C. There are various models for this step (Sung and Dimitrov, 2010; Aziz et al., 2013; Brandt et al., 2023). Usually, within the gadget for a set $S \in \mathcal{S}$, there exist agents representing each of the elements in $\mathcal{S}$. Still, the actual gadgets usually need additional auxiliary agents.

Using No-instances for the design of gadgets is useful to mimic a covering of the ground set R of the source instance. One can introduce a gadget corresponding to a No-instance for each element in R, and it is then required that one agent from each of these gadgets forms a coalition with some agent outside of the gadget. With a clever design, one can enforce that this can only happen in very specific cases, namely when a coalition is formed with the agents representing corresponding sets in $\mathcal{S}$. We follow this idea in all reductions with complicated No-instances, but simpler ideas work for Nash stability in Theorems 1 and 2.

After coming up with the construction, a large part of the proofs is showing their correctness. The goal is to come up with a simple and concise structure of potentially stable partitions that combine the ideas of the gadget design discussed above. We manage to argue about large sets of possible partitions by proving structural properties of stable partitions.

3. Computational Boundaries for Nash and Contractual Nash Stability

In the next sections, we present our results. We start with new computational boundaries for classical solution concepts.

[4] We refer to the PhD thesis by Bullinger (2023, Chapter 4.3) for a more in-depth discussion of general ideas of hardness reductions for hedonic games.

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 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 nontrivial 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. Their reduction produces ASHGs 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 additionally cover other important cases. For instance, the hardness holds for utility values as in FEGs or AEGs.

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 E3C as introduced in Section 2.3.

Now, let $(R, \mathcal{S})$ 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 $(R, \mathcal{S})$ has an exact cover if and only if (N, v) has an NS partition. The reduction is illustrated in Figure 2.

Define the agent set as $N = \bigcup_{r\in R}\{b_i^r : i \in [n_r - 1]\} \cup \bigcup_{S\in\mathcal{S}} N_S \cup \{c\}$, where $N_S = \left\{a_{r_1}^S, a_{r_2}^S, a_{r_3}^S, a^S\right\}$ for $S = \{r_1, r_2, r_3\} \in \mathcal{S}$. Hence, the agent set consists of copies of the elements in R according to the frequency with which they occur in the sets of $\mathcal{S}$ minus 1, copies for the elements in sets of $\mathcal{S}$ together with one specific agent for each such set, and an auxiliary agent c. Now, define the following valuations v:

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

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

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

- First, we create coalitions corresponding to the cover. For each $S \in \mathcal{S}$, we take $N_S \in \pi$ if and only if $S \in \mathcal{S}'$.

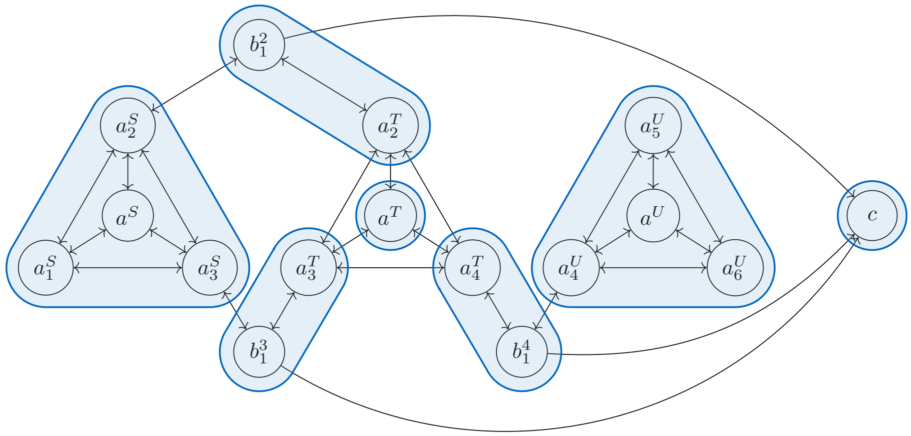

Figure 2: The reduction from the proof of Theorem 1 for the Yes-instance of E3C $(\{S, T, U\})$ with $S = \{1, 2, 3\}, T = \{2, 3, 4\}$ and $U = \{4, 5, 6\}$. Drawn edges have weight $f^+(n)$, omitted edges have weight $f^-(n)$. The partition corresponding to the exact cover $\{S, U\}$ is highlighted.

- This leaves for each $r \in R$ exactly $n_r - 1$ sets $S \in \mathcal{S}_r$ such that $N_S \notin \pi$. Arbitrarily number these sets $S_1, \dots, S_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalition $\left\{a_r^{S_i}, b_i^r\right\}$.
- All agents a^S with $N_S \notin \pi$ are in a singleton: $\pi(a^S) = \left\{a^S\right\}$.
- Agent c is also in a singleton: $\pi(c) = \{c\}$.

To see that this partition is Nash-stable, 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 $\pi(a) = N_S$ 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 $\pi(a_r^S) \neq N_S$ is in a pair with an agent b_i^r, and so are the other two agents $a_{r'}^S$ from N_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 $\pi(a^S) \neq N_S$ is alone, but all other agents $a_r^S \in N_S$ 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 Nash-stable.

$\Longleftarrow$ Suppose now that (N, v) contains an NS partition π. We show that there exists an exact cover $\mathcal{S}' \subseteq \mathcal{S}$ 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 $\pi(a^S) \cap N_{S'} = \emptyset$ for $S' \neq S$. Suppose for contradiction that there is an agent $a \in \pi(a^S) \cap N_{S'}$. Consider the sets $A = \{i \in \pi(a^S) \colon v_a(i) = f^+(n)\}$ and $A' = \{i \in \pi(a^S) \colon v_{a^S}(i) = f^+(n)\}$. 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 $\pi(a^S) \subseteq N_S$ for all $S \in \mathcal{S}$.

5. Now, consider an agent b_i^r. Define the sets $A = \{a_r^S \colon S \in \mathcal{S}_r\}$ and $B = \{b_j^r \colon j \in [n_r - 1]\}$. 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 $A \subseteq \{a_r^S \colon S \in \mathcal{S}_r\}, B \subseteq \{b_j^r \colon j \in [n_r - 1]\}$ and $|A| = |B|$. This leaves for each $r \in R$ exactly one agent a_r^S that is not in such a coalition. For these agents we have $\pi(a_r^S) = N_S$, yielding a cover $\mathcal{S}' = \{S \in \mathcal{S} \colon N_S \in \pi\}$. $\square$

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. The essential difference is that we now represent every element in the ground set of an E3C instance by a pair of agents.

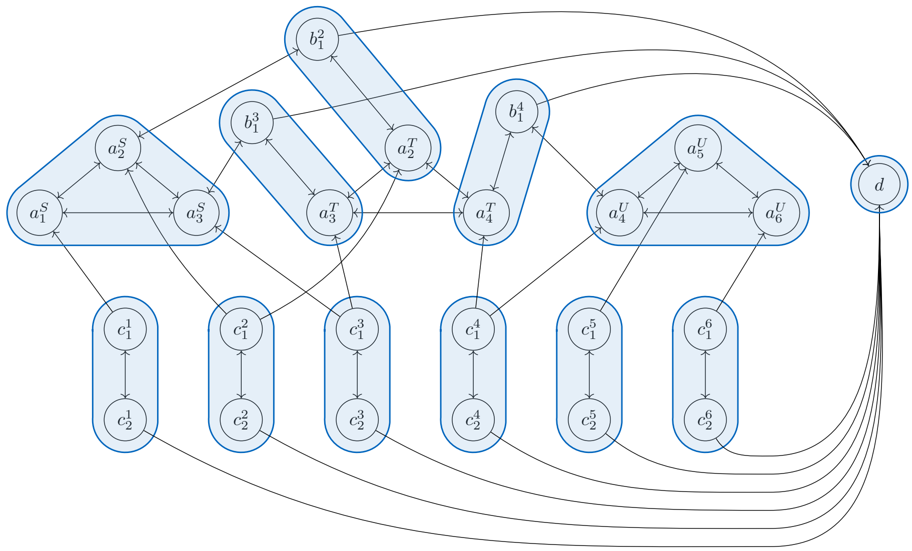

Figure 3: The reduction from the proof of Theorem 2 for the Yes-instance of E3C $(\{1, \dots, 6\}, \{S, T, U\})$ with $S = \{1, 2, 3\}, T = \{2, 3, 4\}$ and $U = \{4, 5, 6\}$. Drawn edges have weight n, and omitted edges have weight -1. The partition corresponding to the exact cover $\{S, U\}$ is highlighted.

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

Proof. We provide another reduction from E3C. The reduction is illustrated in Figure 3. Let $(R, \mathcal{S})$ be an instance of E3C. We produce an AFG (N, v) such that $(R, \mathcal{S})$ has an exact cover if and only if (N, v) has an NS partition. Define the agent set $N = \{d\} \cup \bigcup_{S \in \mathcal{S}} N_S \cup \bigcup_{r \in R}(\{c_1^r, c_2^r\} \cup \{b_i^r : i \in [n_r - 1]\})$, where $N_S = \left\{a_r^S : r \in S\right\}$ for $S \in \mathcal{S}$.

Also, define the following valuations v:

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

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

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

- First, we create coalitions corresponding to the cover, i.e., for each $S \in \mathcal{S}$: $N_S \in \pi$ if and only if $S \in \mathcal{S}'$.

- This leaves for each $r \in R$ exactly $n_r - 1$ sets $S \in \mathcal{S}_r$ such that $N_S \notin \pi$. 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 $r \in R$, 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 (except d who has no friends) 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 an incentive to deviate.

$\Longleftarrow$ Conversely, assume that (N, v) has an NS partition π. We show that there exists an exact cover $\mathcal{S}' \subseteq \mathcal{S}$ 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.

Define $\mathcal{S}' = \{S \in \mathcal{S} \colon \pi(a_r^S) \cap \{b_i^r \colon i \in [n_r - 1]\} = \emptyset \text{ for some } r \in S\}$. We claim that $\mathcal{S}'$ partitions R. First, we 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, and therefore not in a coalition with any agent from $\{b_i^r \colon i \in [n_r - 1]\}$. Hence, every agent in R is covered by $\mathcal{S}$.

Now, assume for contradiction that there are $S_1, S_2 \in \mathcal{S}$ with $S_1 \cap S_2 \neq \emptyset$. Let $j \in [2]$. Then, there exists $r_j \in S_j$ with $\pi(a_{r_j}^{S_j}) \cap \{b_i^{r_j} \colon i \in [n_r - 1]\} = \emptyset$. Since π is Nash-stable, it must be the case that $\pi(a_{r_j}^{S_j})$ contains at least one friend of $a_{r_j}^{S_j}$ and therefore $|\pi(a_{r_j}^{S_j}) \cap N_{S_j}| \geq 2$. Now, every agent in $N_{S_j} \setminus \pi(a_{r_j}^{S_j})$ can have at most one friend and would therefore perform an NS deviation to join $\pi(a_{r_j}^{S_j})$. Hence, there can be no such agent and therefore $A_{S_j} \subseteq \pi(a_{r_j}^{S_j})$. Hence, for $r \in S_1 \cap S_2$, at most $n_r - 2$ agents can be in pairs with agents b_i^r. This is a contradiction. Thus, the sets in $\mathcal{S}'$ are disjoint and therefore $\mathcal{S}'$ partitions R. $\square$

Our next result settles the computational complexity of contractual Nash stability in ASHGs. Generally, we follow the proof scheme described in Section 2.3. Still, since this is the first time that we see the entire scheme in practice, we start by describing how it works in this specific case.

Given an instance $(R, \mathcal{S})$ of E3C, the reduced instance consists of several gadgets. First, every element in R is represented by a subgame that does not contain a CNS partition. In principle, any such game can be used for a reduction, and we use a simple game identified by Sung and Dimitrov (2007). Moreover, we have further auxiliary gadgets that also consist of the same No-instance. The number of these auxiliary gadgets is equal to the number of sets in $\mathcal{S}$ that would remain after removing an exact cover of R, i.e., there are $|\mathcal{S}|-|R|/3$ such gadgets. By design, the agents in the subgames corresponding to No-instances have to form coalitions with agents outside of their subgame in every CNS partition. The only agents that can achieve this are agents in gadgets corresponding to elements in $\mathcal{S}$. A gadget corresponding to an element $S \in \mathcal{S}$ can either prevent nonstability caused by exactly one auxiliary gadget, or by the three gadgets corresponding to the elements $r \in R$ with $r \in S$. Hence, the only possibility to deal with all No-instances simultaneously is if there exists an exact cover of R by sets in $\mathcal{S}$. Then, the gadgets corresponding to elements in R can be dealt with by the cover, and there are just enough elements in $\mathcal{S}$ to additionally deal with the other auxiliary gadgets.

Theorem 3. *Deciding whether an ASHG contains a CNS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let $(R, \mathcal{S})$ be an instance of E3C and set $a = |\mathcal{S}| - |R|/3$ (this is the number of additional sets in $\mathcal{S}$ if removing some exact cover). Without loss of generality, $a \geq 0$. We define an ASHG (N, v) as follows. Let $N = N_{\mathcal{S}} \cup \bar{N}_S \cup N_R \cup N_A$ where

- $N_{\mathcal{S}} = \cup_{S\in\mathcal{S}} N_S$ with $N_S = \{a_r^S : r \in S\}$ for $S \in \mathcal{S}$,
- $\bar{N}_{\mathcal{S}} = \cup_{S\in\mathcal{S}} \bar{N}_S$ with $\bar{N}_S = \{\bar{a}_r^S : r \in S\}$ for $S \in \mathcal{S}$, and
- $N_R = \cup_{r\in R} N_r$ with $N_r = \{b_i^r : i \in [4]\}$ for $r \in R$,
- $N_A = \cup_{1\leq j\leq a} N^j$ with $N^j = \{x_i^j : i \in [4]\}$ for $1 \leq j \leq a$.

We define valuations v as follows:

- For each $r \in R$, $i \in [3]$: $v_{b_i^r}(b_4^r) = 1$.
- For each $r \in R$, $(i,j) \in \{(1,2),(2,3),(3,1)\}$: $v_{b_i^r}(b_j^r) = 0$.
- For each $1 \leq j \leq a$, $i \in [3]$: $v_{x_i^j}(x_4^j) = 1$.
- For each $1 \leq j \leq a$, $(i,k) \in \{(1,2),(2,3),(3,1)\}$: $v_{x_i^j}(x_k^j) = 0$.
- For each $S \in \mathcal{S}$, $r \in S$: $v_{a_r^S}(b_4^r) = 1$.
- For each $S \in \mathcal{S}$, $r \in S$, $1 \leq j \leq a$: $v_{a_r^S}(x_4^j) = v_{x_4^j}(a_r^S) = 0$.
- For each $S \in \mathcal{S}$, $r, r' \in S$: $v_{a_r^S}(a_{r'}^S) = 0$.
- For each $S \in \mathcal{S}$, $r, r' \in S$, $r \neq r'$, $z \in (N_{\mathcal{S}} \cup N_A) \setminus N_S$: $v_{\bar{a}_r^S}(a_r^S) = 3$, $v_{\bar{a}_r^S}(a_{r'}^S) = -2$, and $v_{\bar{a}_r^S}(z) = 0$.

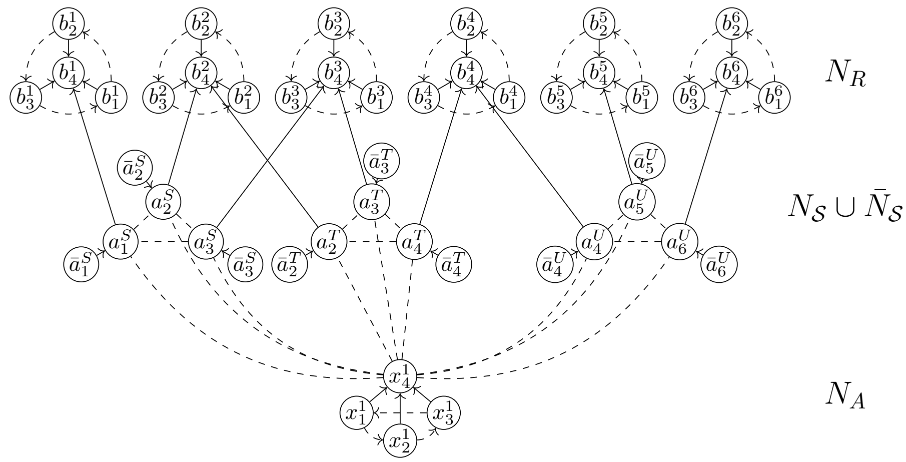

Figure 4: Schematic of the reduction from the proof of Theorem 3. We depict the reduced instance for the instance $(R, \mathcal{S})$ of E3C where $R = \{1, 2, 3, 4, 5, 6\}$, and $\mathcal{S} = \{S, T, U\}$, with $S = \{1, 2, 3\}$, $T = \{2, 3, 4\}$, and $U = \{4, 5, 6\}$. Fully drawn edges mean a positive utility, which is usually 1 except between agents of the types $\bar{a}_r^S$ and a_r^S, where $v_{\bar{a}_r^S}(a_r^S) = 3$. Dashed edges represent a utility of 0. For agents in $\bar{N}_{\mathcal{S}}$, only the single positive utility is displayed, but not the 0-utilities and negative utility of -2 to other agents in N^r. Other omitted edges represent a negative utility of -4.

- All other valuations are -4.

An illustration of the game is given in Figure 4. The agents in N_R in the reduced instance form gadgets consisting of a subgame without a CNS partition for every element in R. The agents in N_A constitute further such gadgets. The agents in $N_{\mathcal{S}}$ consist of triangles for every set in $\mathcal{S}$ and are the only agents who can bind agents in the gadgets in any CNS partition. Finally, agents in $\bar{N}_{\mathcal{S}}$ avoid having agents in $N_{\mathcal{S}}$ in separate coalitions to bind agents in N_A.

We claim that $(R, \mathcal{S})$ is a Yes-instance if and only if (N, v) contains a CNS partition.

$\Longrightarrow$ Suppose first that $\mathcal{S}' \subseteq \mathcal{S}$ partitions R. Consider any bijection $\phi \colon \mathcal{S} \setminus \mathcal{S}' \to [a]$. Define a partition π by taking the union of the following coalitions:

- For every $r \in R, i \in [3]$, form $\{b_i^r\}$.
- For $S \in \mathcal{S}', r \in S$, form $\{a_r^S, b_4^r\}$.
- For $S \in \mathcal{S} \setminus \mathcal{S}'$, form $\{a_r^S : r \in S\} \cup \{x_4^{\phi(s)}\}$.
- For $S \in \mathcal{S}, r \in S$, form $\{\bar{a}_r^S\}$.
- For $1 \leq j \leq a, i \in [3]$, form $\{x_i^j\}$.

We claim that π is contractually Nash-stable. We will show that no agent can perform a deviation.

- For $r \in R$, $i \in [3]$, it holds that $v_{b_i^r}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{b_i^r}(\pi(b_4^r) \cup \{b_i^r\}) = -3$.
- For $r \in R$, b_4^r is not allowed to leave her coalition.
- For $S \in \mathcal{S}'$, $r \in S$, it holds that $v_{a_r^S}(\pi) = 1$ and joining any other coalition results in a negative utility. The agent a_r^S is in a most preferred coalition.
- For $S \in \mathcal{S} \setminus \mathcal{S}'$, $r \in S$, it holds that $v_{a_r^S}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{a_r^S}(\pi(b_4^r) \cup \{a_r^S\}) = -3$.
- For $S \in \mathcal{S}'$, $r \in S$, the agent $\bar{a}_r^S$ obtains a nonpositive utility by joining any other coalition. In particular, $v_{\bar{a}_r^S}(\pi(a_r^S) \cup \{\bar{a}_r^S\}) = -1$.
- For $S \in \mathcal{S} \setminus \mathcal{S}'$, $r \in S$, the agent $\bar{a}_r^S$ obtains a nonpositive utility by joining any other coalition. In particular, $v_{\bar{a}_r^S}(\pi(a_r^S) \cup \{\bar{a}_r^S\}) = -1$.
- For $1 \leq j \leq a$, $i \in [3]$, it holds that $v_{x_i^j}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{x_i^j}(\pi(x_4^j) \cup \{x_i^j\}) = -11$.
- For $1 \leq j \leq a$, x_4^j is in a best possible coalition (achieving utility 0).

$\Longleftarrow$ Conversely, assume that (N, v) contains a CNS partition π. Define $\mathcal{S}' = \{S \in \mathcal{S} \colon \pi(a_r^S) \cap N_R \neq \emptyset \text{ for some } r \in S\}$. We will show first that $\mathcal{S}'$ covers all elements in R and then show that $|\mathcal{S}'| = |R|/3$.

Let $r \in R$. Then, for all $i \in [3]$, $\pi(b_i^r) \subseteq N_r$. This follows because there is no agent who favors b_i^r in her coalition. Therefore, she would leave any coalition with an agent outside N_r to receive nonnegative utility in a singleton coalition. Further, if there is no $S \in \mathcal{S}$ with $r \in S$ such that $b_4^r \in \pi(b_S^r)$, then $\pi(b_4^r) \subseteq N_r$. Indeed, if b_4^r forms any coalition except a singleton coalition, she will receive negative utility, and then there must exist an agent who favors her in the coalition. Consequently, if $b_4^r \notin \pi(b_S^r)$ for all $S \in \mathcal{S}$ with $r \in S$, then b_4^r is in a singleton coalition, or there exists $i \in [3]$ with $b_4^r \in \pi(b_i^r)$, for which we already know that $\pi(b_i^r) \subseteq N_r$.

Assume now that $\pi(b_4^r) \subseteq N_r$. For $i, i' \in [3]$, $b_i^r \notin \pi(b_{i'}^r)$ because then one of them would receive a negative utility and could perform a CNS deviation to form a singleton coalition. If $\{b_4^r\} \in \pi$, then b_1^r would deviate to join her. Hence, there exists exactly one $i \in [3]$ with $\{b_i^r, b_4^r\} \in \pi$. Suppose without loss of generality that $\{b_1^r, b_4^r\} \in \pi$. But then, b_3^r would perform a CNS deviation to join them, a contradiction. We can conclude that there exists $S \in \mathcal{S}$ with $r \in S$ such that $b_4^r \in \pi(b_S^r)$. Hence, $S \in \mathcal{S}'$ and we have shown that $\mathcal{S}'$ covers R.

To bound the cardinality of $\mathcal{S}'$, we will show that, for every $1 \leq j \leq a$, there exists $S \in \mathcal{S} \setminus \mathcal{S}'$ with $N_S \subseteq \pi(x_4^j)$. Let therefore $1 \leq j \leq a$ and let $C = \pi(x_4^j)$. Similar to the considerations about agents in N_r, we know that $\pi(x_i^j) \subseteq X^j$ for $i \in [3]$, and that it cannot happen that $C \subseteq X^j$, and therefore $C \cap X^j = \{x_4^j\}$. In particular, there must be an agent $y \in N \setminus X^j$ with $y \in C$. Since no agent in C favors x_4^j to be in her coalition, we know that $v_{x_4^j}(\pi) \geq 0$ and therefore

$C \subseteq \{x_4^j\} \cup N_{\mathcal{S}}$. Let $S \in \mathcal{S}$ and $r \in S$ with $a_r^S \in C$. As we already know that $\bar{a}_r^S \notin C$, it must hold that $N_S \subseteq C$ to prevent her from joining. It follows that $S \notin \mathcal{S}'$. Since $\pi(x_4^j) \cap \pi(x_4^{j'}) = \emptyset$ for $1 \leq j' \leq a$ with $j' \neq j$, we find an injective mapping $\phi\colon [a] \to S \setminus \mathcal{S}'$ such that, for every $1 \leq j \leq a$, $N_{\phi(j)} \subseteq \pi(x_4^j)$. Consequently, $|\mathcal{S}'| \leq |\mathcal{S}| - |\phi([a])| \leq |\mathcal{S}| - a = |R|/3$. Hence, $\mathcal{S}'$ covers all elements from R with (at most) $|R|/3$ sets and therefore is an exact cover. □

The reduction in the previous proof only uses a very limited number of different weights, namely the weights in the set $\{3, 1, 0, -2, -4\}$, where the weight -4 may be replaced by an arbitrary smaller weight. By contrast, we will see in the next section that CNS partitions always exist if the utility functions of an ASHG assume at most one nonpositive value, and can be computed efficiently in this case (cf. Theorem 5). This encompasses, for instance, FEGs, AFGs, and AEGs. Hence, the hardness result is close to the boundary of computational feasibility.

4. Deviation Lemma and Applications

In this section, we will present a unified approach for proving the existence of stable partitions. Even more, our technique allows us to prove convergence of dynamics. The goal is to apply 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:*

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

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 sum on 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 4. *The IS dynamics 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, then all IS deviations are singleton formations. Hence, 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 an 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

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

Telescoping and termwise application of the above inequality yields $\Phi(\pi_k) - \Phi(\pi_0) = \sum_{j\in[k]} \Phi(\pi_j) - \Phi(\pi_{j-1}) \geq \sum_{j\in[k]} \Delta + x\,(|\pi_j(i_j)| - |\pi_{j-1}(i_j)|) = k\Delta + x\sum_{j\in[k]}|\pi_j(i_j)| - |\pi_{j-1}(i_j)|$. 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 is 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, consider the instance with agent set $N = [5]$ and utility values $v_i(j) = 1, v_j(i) = 0$ for $(i,j) \in \{(1,2),(2,3),(3,4),(4,5),(5,1)\}$. All other values are -2. In this example, no IS partition exists (see also Bogomolnaia and Jackson, 2002, Example 5), and therefore, the dynamics is doomed to cycle. Notably, the example still works when replacing 0 and 1 by any two different nonnegative utility values x and y, and replacing the negative utility value by $-x-y-1$.

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 5. *The CNS dynamics converges in ASHGs with at most one nonpositive utility value.*

Theorems 4 and 5 use the Deviation Lemma to derive positive results for individual stability and contractual Nash stability, which only involve either the welcoming or the abandoned coalition. In a third application of the lemma, we show that this technique is also applicable for 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 6. *The JMS dynamics 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 4, 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 the 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

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

□

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 4, 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.[5]

Corollary 1. *The IS, CNS, and JMS dynamics converge in polynomial time in AFGs, AEGs, and FEGs.*

Inspecting Equations (3) and (4), one can immediately obtain a running time of $\mathcal{O}(n^3)$ for AFGs and AEGs, and of $\mathcal{O}(n^2)$ for FEGs. We conjecture that the former bound can be improved with a more refined analysis. We would, however, 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 NS dynamics 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).[6] □

While the previous proposition uses a nonconstructive argument and, therefore, does not identify 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, we 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.

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

[5]Their work even shows the existence of partitions satisfying properties stronger than IS.

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

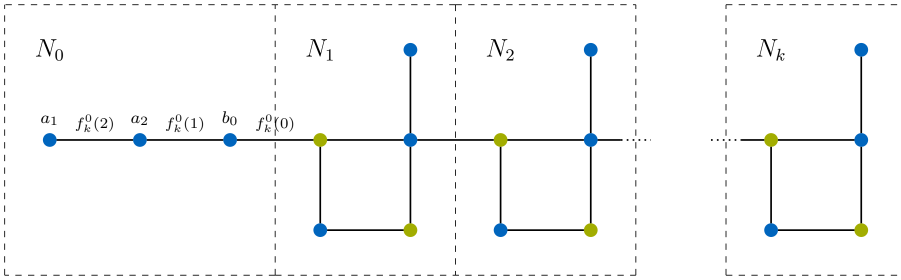

Figure 5: Exponential length IS dynamics inspired by Monien and Tscheuschner (2010). The starting partition into two coalitions is indicated by the coloring of the vertices.

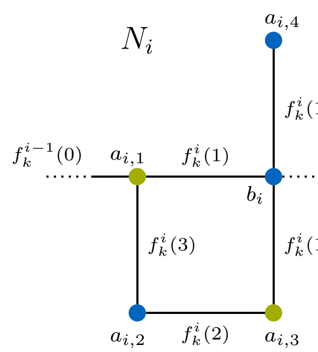

Figure 6: Utilities in the subgame induced by N_i

Proof. We define a class of ASHGs parametrized by $k \in \mathbb{N}$. Define the agent set as $N = N_0 \cup \bigcup_{i=1}^{k} N_i$ with $N_0 = \{a_1, a_2, b_0\}$ and for $1 \leq i \leq k$, define $N_i = \{b_i, a_{i,1}, a_{i,2}, a_{i,3}, a_{i,4}\}$. Consider the symmetric ASHG on this set of agents with utility values induced by the graph presented in Figure 5, where the weights of the game restricted to N_i are depicted in Figure 6.[7] More precisely, the weight function is given by $f_k^i(j) = j + 5(2^{k-i+1} - 1)$.[8] All weights on missing edges are 0.

The underlying combinatorial structure consists of a short path induced by N_0 together with k copies of the same graph with exponentially growing weights. The agent sets N_{i-1} and N_i are connected by a positive utility between b_{i-1} and $a_{i,1}$.

Consider the partition of N indicated by the blue and green vertices and defined by $\pi = \{\{a_1, b_0\} \cup \bigcup_{i=1}^{n}\{b_i, a_{i,2}, a_{i,4}\}, \{a_2\} \cup \bigcup_{i=1}^{n}\{a_{i,1}, a_{i,3}\}\}$. We claim

[7]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.

[8]Note that there is a typo in the weight function by Monien and Tscheuschner (2010). They probably meant to use a similar weight function as the one used here.

that there is an execution of the IS dynamics starting with π where agent b_i performs 2^{i+1} deviations for $i \in \{0, 1, \dots, k\}$.

We will recursively construct a sequence of deviating agents. These deviating agents perform deviations by joining the nonempty coalition which they are not part of. In the ith step of the recursion, agent b_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 change the coalition agent $a_{i+1,1}$ was part of when b_i performs an IS deviation. However, this is not a problem, because the IS deviations of b_i are valid independently of the coalition that $a_{i+1,1}$ is part of. For $i = 0$, consider the sequence of deviations performed by (b_0, a_2, b_0), where b_0 performs $2 = 2^{0+1}$ deviations.

Now, let $i \geq 1$ and assume that the sequence is constructed for $i - 1$. We extend the sequence of deviations by inserting suitable subsequences. Right before b_{i-1} performs her mth deviation, we insert

$$\begin{cases} (b_i, a_{i,3}, b_i, a_{i,2}, a_{i,3}, b_i, a_{i,1}) & \text{if } m \text{ odd} \\ (a_{i,2}, b_i, a_{i,1}) & \text{if } m \text{ even} \end{cases}$$

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 $a_{i,1}$. Whenever they perform a deviation, they leave the coalition of $a_{i,2}$ and b_i to join the coalition of b_{i-1}. Indeed, this yields an improvement in utility because $f_k^{i-1}(0) = 5(2^{k-i+2} - 1) > 4 + 10(2^{k-i+1} - 1) = f_k^i(3) + f_k^i(1)$. Note that after every even m, the subpartition of agents in N_k is the same as in the initial partition π. Moreover, the agent b_i performs 2^{i+1} deviations.

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

An interesting follow-up question is whether NS or IS dynamics may necessarily require exponential running time, i.e., take an exponential number of steps regardless of the sequence of deviations. We conjecture that such examples can be obtained by adapting similar examples for MaxCut (Monien and Tscheuschner, 2010; Michel and Scott, 2024).

5. Complexity of Stability under Majority Consent

In this section, we study stability under majority consent.

5.1. Aversion-To-Enemies Games

First, we contrast the existential results of Theorem 4 and Theorem 5 with the nonexistence of stable partitions in AEGs under the majority-based relaxations of the respective stability concepts.

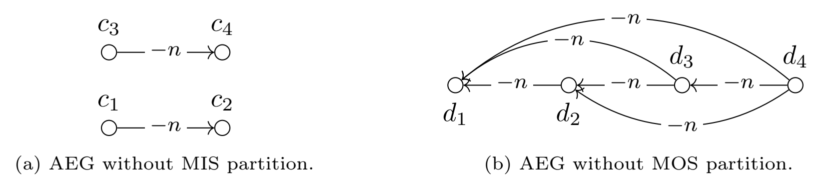

Figure 7: Aversion-to-enemies games without MIS and MOS partitions from the proof ofProposition 3. Omitted edges have weight 1.

Proposition 3. *There exists an AEG which contains no MIS (or 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 7a.

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, $|\pi(c_1)| \leq 1$ (and $|\pi(c_3)| \leq 1$), because otherwise, c_2 (or 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 7b.

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. $\square$

We can leverage the AEGs provided in the previous proposition as gadgets in reductions to show NP-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 4 and Theorem 5 for the special case of AEGs. We start with majority-in stability.

Theorem 7. *It is* NP*-complete to decide if there exists an MIS partition in AEGs.*

Proof. We perform another reduction from E3C. It is illustrated in Figure 8. Let $(R, \mathcal{S})$ be an instance of E3C. We produce an AEG (N, v) such that $(R, \mathcal{S})$ admits an exact cover if and only if (N, v) contains an MIS partition. Define $N = N_R \cup N_{\mathcal{S}}$ where

- $N_R = \bigcup_{r \in R} N_r$ where $N_r = \bigcup_{i=1}^{n_r - 1} B_i^r$ with $B_i^r = \left\{b_{i,j}^r : j \in [4]\right\}$ for $r \in R, i \in [n_r - 1]$, and
- $N_{\mathcal{S}} = \bigcup_{S \in \mathcal{S}} N_S$ with $N_S = \left\{a_{r_1}^S, a_{r_2}^S, a_{r_3}^S, a^S\right\}$ for $S = \{r_1, r_2, r_3\} \in \mathcal{S}$.

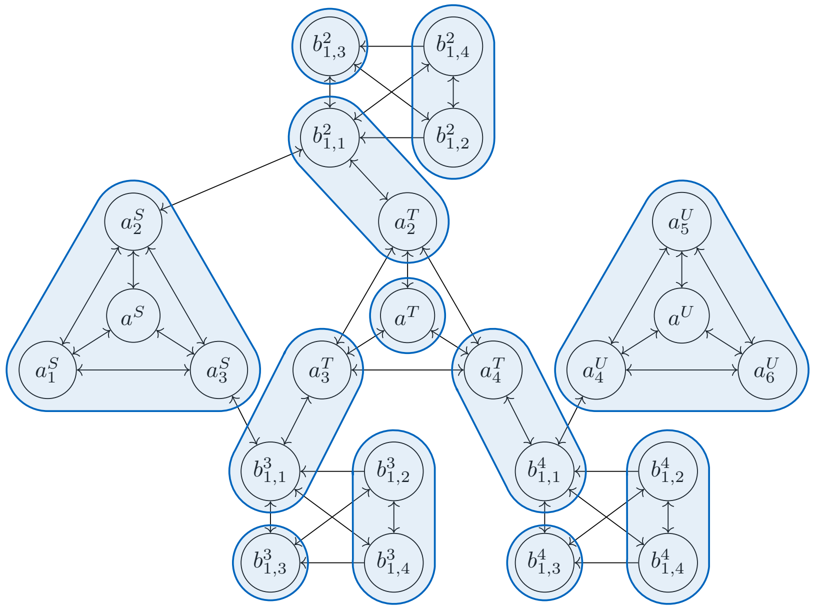

Figure 8: The reduction from the proof of Theorem 7 for the Yes-instance of E3C $(\{S, T, U\})$ with $S = \{1, 2, 3\}, T = \{2, 3, 4\}$ and $U = \{4, 5, 6\}$. Drawn edges have weight 1, omitted edges have weight $-n$. The partition corresponding to the exact cover $\{S, U\}$ is highlighted.

Define valuations v as follows.

- For each $S \in \mathcal{S}, a \neq a' \in N_S$: $v_a(a') = 1$.
- For each $r \in R, S \in \mathcal{S}_r, i \in [n_r - 1]$: $v_{a^S_r}(b^r_{i,1}) = v_{b^r_{i,1}}(a^S_r) = 1$.
- Each B^r_i has internal valuations as in the first example of Proposition 3, i.e., if v' denotes the valuations of this example, then $v_{b^r_{i,j}}(b^r_{i,k}) = v'_{c_j}(c_k)$, 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.

$\implies$ Suppose that $(R, \mathcal{S})$ has an exact cover $\mathcal{S}' \subseteq \mathcal{S}$. We construct an MIS partition π as follows.

- First, we create coalitions corresponding to the cover, i.e., for each $S \in \mathcal{S}$, we have $N_S \in \pi$ if and only if $S \in \mathcal{S}'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $S \in \mathcal{S}_r$ such that $N_S \notin \pi$. Arbitrarily number these sets $S_1, \ldots, S_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a^{S_i}_r, b^r_{i,1}\}$, $\{b^r_{i,2}, b^r_{i,4}\}$, and $\{b^r_{i,3}\}$.
- Finally, form $\{a^S\}$ for each $S \in \mathcal{S} \setminus \mathcal{S}'$.

In this partition, no agent can improve her utility by deviating, making the partition an NS partition and thus an MIS partition.

$\Longleftarrow$ Conversely, assume that (N, v) admits an MIS partition π. We construct an exact cover $\mathcal{S}' \subseteq \mathcal{S}$. 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 $\pi(a^S) \subseteq N_S$ 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 $\pi(a^S_r) = N_S$. Indeed, it is clear that we then must have $\pi(a^S_r) \subseteq N_S$. If $\pi(a^S_r) = \{a^S_r\}$, she would deviate to join $\pi(a^S)$. Then, $\left|\pi(a^S_r)\right| \geq 2$, and members from $N_S \setminus \pi(a^S_r)$ would have an incentive to join $\pi(a^S_r)$. It follows that $N_S \setminus \pi(a^S_r) = \emptyset$, and therefore $\pi(a^S_r) = N_S$. Hence, we obtain a cover $\mathcal{S}' = \{S \in \mathcal{S} \colon N_S \in \pi\}$. ☐

Note that it can be shown that a partition in the AEGs constructed in the reduction of Theorem 7 is Nash-stable if and only if it is majority-in stable. Hence, the theorem provides yet another proof of the respective statement about Nash stability first shown by Sung and Dimitrov (2010) (and already revisited in Theorem 1). We proceed with the analogous result for majority-out stability.

Theorem 8. *It is* NP*-complete to decide if there exists an MOS partition in AEGs.*

Proof. Again, we reduce from E3C. Let $(R, \mathcal{S})$ be an instance of E3C. The reduced instances are very similar to the instance in Figure 8, so we omit another illustration. We produce an AEG (N, v) as follows. Define $N = N_R \cup N_{\mathcal{S}}$ where

- $N_R = \bigcup_{r \in R} N_r$ where $N_r = \bigcup_{i=1}^{n_r - 1} B^r_i$ with $B^r_i = \left\{b^r_{i,j} \colon j \in [4]\right\}$ for $r \in R, i \in [n_r - 1]$, and
- $N_{\mathcal{S}} = \bigcup_{S \in \mathcal{S}} N_S$ with $N_S = \left\{a^S_{r_1}, a^S_{r_2}, a^S_{r_3}, a^S\right\}$ for $S = \{r_1, r_2, r_3\} \in \mathcal{S}$.

Define the following valuations v.

- For each $S \in \mathcal{S}, a \neq a' \in N_S$: $v_a(a') = 1$.
- For each $r \in R, S \in \mathcal{S}_r, i \in [n_r - 1]$: $v_{a^S_r}(b^r_{i,1}) = 1$.

- Each B_i^r 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 $v_{b_{i,j}^r}(b_{i,k}^r) = v'_{d_j}(d_k)$, where the negative valuations are adapted to the specific number of agents in the instance.

- All other valuations are $-n$.

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

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

- First, we create coalitions corresponding to the cover, i.e., for each $S \in \mathcal{S}$, we have $N_S \in \pi$ if and only if $S \in \mathcal{S}'$.

- This leaves for each $r \in R$ exactly $n_r - 1$ sets $S \in \mathcal{S}_r$ such that $N_S \notin \pi$. Arbitrarily number these sets $S_1, \dots, S_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalitions $\left\{a_r^{s_i}, b_{i,1}^r\right\}$, $\left\{b_{i,2}^r, b_{i,3}^r\right\}$, and $\left\{b_{i,4}^r\right\}$.

- Finally, form $\left\{a^S\right\}$ for each $S \in \mathcal{S} \setminus \mathcal{S}'$.

The only agents that have an NS deviation are agents of types $b_{i,1}^r$ and $b_{i,3}^r$. However, there is some $S \in \mathcal{S}$ such that $\pi(b_{i,1}^r) = \left\{b_{i,1}^r, a_r^S\right\}$, and a_r^S ensures that b_i^r cannot leave. Similarly, $\pi(b_{i,3}^r) = \left\{b_{i,3}^r, b_{i,2}^r\right\}$, and $b_{i,2}^r$ ensures that $b_{i,3}^r$ cannot leave. Note that agents a^S for $S \notin \mathcal{S}'$ cannot deviate, because all their friends form a coalition with an enemy. Hence, π is majority-out stable.

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

1. Agents $b_{i,2}^r$ must have $\pi(b_{i,2}^r) \subseteq B_i^r$. If there was an agent $a \in \pi(b_{i,2}^r) \setminus B_i^r$, then, as $v_{b_{i,2}^r}(a) = -n$, $b_{i,2}^r$ would rather be in a singleton, and could form one, as $\left|F_{\text{out}}(\pi(b_{i,2}^r), b_{i,2}^r)\right| \geq |\{a\}| = 1 = \left|\left\{b_{i,1}^r\right\}\right| \geq \left|F_{\text{in}}(\pi(b_{i,2}^r), b_{i,2}^r)\right|$.

2. Using Observation 1, we can conclude that agents $b_{i,3}^r$ must also have $\pi(b_{i,3}^r) \subseteq B_i^r$.

3. Using Observations 1 and 2, we can conclude that agents $b_{i,4}^r$ must also have $\pi(b_{i,4}^r) \subseteq B_i^r$.

4. Agents $a \in N_S$ and $a' \in N_{S'}$ with $S \neq S'$ satisfy $\pi(a) \neq \pi(a')$. For contradiction, suppose this is not the case, i.e., there are $a \in N_S$ and $a' \in N_{S'}$ with $S \neq S'$ such that $\pi(a) = \pi(a')$. Define $C = \pi(a)$. Clearly, both a and a' prefer to be in a singleton coalition. Further, we can assume without loss of generality that $|N_S \cap C| \leq |N_{S'} \cap C|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a)| \geq |N_{S'} \cap C| \geq |N_S \cap C| > |F_{\text{in}}(C, a)|$, 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 N_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 $\pi(a^S_r) \subseteq N_S$. Also, $\pi(a^S) \subseteq N_S$, as any agent outside would like to leave and there is at most 1 vote for her to stay. Consequently, $\left|\pi(a^S_r)\right| \geq 2$, and members from $N_S \setminus \pi(a^S_r)$ would have an incentive to join $\pi(a^S_r)$. Hence, $\pi(a^S_r) = N_S$, and we obtain a cover $\mathcal{S}' = \{S \in \mathcal{S} \colon N_S \in \pi\}$. □

5.2. Appreciation-Of-Friends Games

The utility restrictions in Theorems 7 and 8 are not as flexible as in the negative result for Nash stability in Theorem 1 or the positive results for unanimity-based dynamics in Theorems 4 and 5. In fact, the picture for majority-based notions is more diverse because we obtain another positive result for the class of AFGs.

Theorem 9. *When starting from the grand coalition, the MIS dynamics 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. $\forall j \in [k] : \left\{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\right\} = \emptyset$.

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 $\left\{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\right\} = \emptyset$. 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. □

By contrast, we show in our next result that MOS partitions need not exist in AFGs. In other words, despite their conceptual complementarity, majority-out and majority-in stability behave differently in a natural subclass of ASHGs. The constructed game has a sparse friendship relation in the sense that almost all agents only have a single friend. After discussing the counterexample, we show how requiring slightly more sparsity yields a positive result.

Proposition 4. *There exists an AFG without an MOS partition.*

Proof. We define the game formally. An illustration is given in Figure 9. Let $N = \{z\} \cup \bigcup_{x \in \{a,b,c\}} N_x$, where $N_x = \{x_i \colon i \in [5]\}$ for $x \in \{a, b, c\}$. In the whole proof, we read indices modulo 5, mapping to the respective representative in the set [5]. The utilities are given as:

- For all $i \in [5], x \in \{a, b, c\} \colon v_{x_i}(x_{i+1}) = n$.
- For all $x \in \{a, b, c\} \colon v_{x_1}(z) = n$.
- All other valuations are -1.

The AFG consists of 3 cycles with 5 agents each, together with a special agent that is liked by a fixed agent of each cycle and has no friends herself. The key insight to understanding why there exists no MOS partition is that agents of type x_1 where $x \in \{a, b, c\}$ have conflicting candidate coalitions in a potential MOS partition. Either, they want to be with z (a coalition that has to be small because z prefers to stay alone) or they want to be with x_2 which requires a rather large coalition containing their cycle.

Before going through the proof that this game has no MOS partition, it is instructional to verify that, for cycles of 5 agents, the unique MOS partition is the grand coalition, i.e., the unique MOS partition of the game restricted to N_x is $\{N_x\}$, where $x \in \{a, b, c\}$. This is a key idea of the construction and is implicitly shown in Case 2 of the proof for $x = b$.

Assume for contradiction that the defined AFG admits an MOS partition π. To derive a contradiction, we perform a case distinction over the coalition sizes of z.

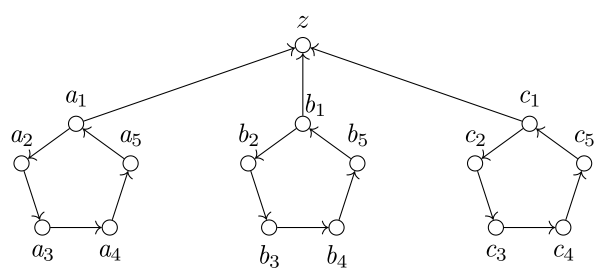

Figure 9: AFG without an MOS partition. The depicted (directed) edges represent friends, i.e., a utility of n, whereas missing edges represent a utility of -1.

Case 1: Assume that $|\pi(z)| = 1$.

In this case, it holds that $\pi(z) = \{z\}$. Then, $\pi(a_1) \in \{\{a_1, a_2\}, \{a_1, a_5\}\}$. Indeed, if $\pi(a_1) \neq \{a_1, a_2\}$, then a_1 has an NS deviation to join z, and is allowed to perform it unless $\pi(a_1) = \{a_1, a_5\}$. We may therefore assume that $\{a_i, a_{i+1}\} \in \pi$ for some $i \in \{1, 5\}$. Then, $\pi(a_{i-1}) = \{a_{i-1}, a_{i-2}\}$. Otherwise, a_{i-1} can perform an MOS deviation to join $\{a_i, a_{i+1}\}$. But then a_{i+2} can perform an MOS deviation to join $\pi(a_{i-1})$. This is a contradiction and concludes the case that $|\pi(z)| = 1$.

Case 2: Assume that $|\pi(z)| > 1$.

Define $F = \{a_1, b_1, c_1\}$, i.e., F is the set of agents that have z as a friend. Note that z can perform an NS deviation by forming a singleton coalition. Hence, as π is majority-out stable, $|F \cap \pi(v)| \geq |\pi(z)|/2$. In particular, there exists an $x \in \{a, b, c\}$ with $\pi(z) \cap N_x = \{x_1\}$. We may assume without loss of generality that $\pi(z) \cap N_a = \{a_1\}$. Then, $\pi(a_5) = \{a_4, a_5\}$. Otherwise, a_5 has an MOS deviation to join $\pi(z)$. Similarly, $\pi(a_3) = \{a_2, a_3\}$ (because of the potential deviation of a_3 who would like to join $\{a_4, a_5\}$). Now, note that $v_{a_1}(\{a_1, a_2, a_3\}) = n - 1$. We can conclude that $|\pi(z)| \leq 3$ as a_1 would join $\{a_2, a_3\}$ by an MOS deviation, otherwise. Hence, we find $x \in \{b, c\}$ with $N_x \cap \pi(z) = \emptyset$. Assume without loss of generality that $x = b$ has this property.

Assume first that $\pi(b_1) = \{b_1, b_5\}$. Then, $\pi(b_4) = \{b_3, b_4\}$. Otherwise, b_4 has an MOS deviation to join $\{b_1, b_5\}$. But then b_2 has an MOS deviation to join $\{b_3, b_4\}$, a contradiction. Hence, $\pi(b_1) \neq \{b_1, b_5\}$. Note that we have now excluded the only case where b_1 is not allowed to perform an MOS deviation, because of the vote of her abandoned coalition. In all other cases, no majority of agents prefers her to stay in the coalition. We can conclude that $b_2 \in \pi(b_1)$ because otherwise, b_1 can perform an MOS deviation to join $\pi(z)$. If $b_5 \notin \pi(b_1)$, then $\pi(b_5) = \{b_4, b_5\}$ (to prevent a potential deviation by b_5). But then b_3 has an MOS deviation to join them. Hence, $b_5 \in \pi(b_1)$. Similarly, if $b_4 \notin \pi(b_1)$, then $\pi(b_4) = \{b_3, b_4\}$ and b_2 has an MOS deviation to join $\{b_3, b_4\}$ (which is permissible because $b_5 \in \pi(b_1)$). Hence $\{b_1, b_2, b_4, b_5\} \subseteq \pi(b_1)$, and therefore even $N_b \subseteq \pi(b_1)$. Hence, b_1 has an MOS deviation to join $\pi(z)$ (recall that $|\pi(z)| \leq 3$). This is the final contradiction, and we can conclude that π is not majority-out stable. □

Note that most agents in the previous example have at most 1 friend (only three agents have 2 friends). By contrast, if every agent has at most one friend, MOS partitions are guaranteed to exist. This is interesting because it covers in particular directed cycles, which include the run-and-chase example in the introduction, which is a No-instance for Nash stability. The constructive proof of the following proposition can be directly converted into a polynomial-time algorithm.

Proposition 5. *Every AFG in which every agent has at most one friend admits an MOS partition.*

Proof. We prove the statement by induction over n. Clearly, the grand coalition is majority-out stable for $n = 1$. Now, assume that (N, v) is an AFG with $n \geq 2$ agents such that every agent has at most one friend. Consider the underlying directed graph $G = (N, A)$ where $(x, y) \in A$ if and only if $v_x(y) > 0$, i.e., y is a friend of x. By assumption, G has a maximum out-degree of 1, hence it can be decomposed into connected components containing a directed cycle and a directed acyclic graph.

Assume first that there exists $C \subseteq N$ such that C induces a directed cycle in G. We call an agent y *reachable* by agent x if there exists a directed path in G from x to y. Let $c \in C$ and define $R = \{x \in N \colon c \text{ reachable by } x\}$. Note that $C \subseteq R$ and that R is identical to the set of agents that can reach *any* agent in C. By induction, there exists an MOS partition π' of the subgame of (N, v) induced by $N \setminus R$ that is majority-out stable. Define $\pi = \pi' \cup \{R\}$. We claim that π is majority-out stable. Let $x \in N \setminus R$. By our assumptions on π', there exists no MOS deviation of x to join $\pi(y)$ for $y \in N \setminus R$. In particular, if x is allowed to perform a deviation, then x must have a nonnegative utility (otherwise, she can form a singleton coalition contradicting that π' is majority-out stable). So her only potential deviations are to a coalition where she has a friend. Note that x has no friend in R. Indeed, if y was a friend of x in R, then c is reachable for x in G through the concatenation of (x, y) and the path from y to c. Hence, x has no MOS deviation. Now, let $x \in R$. Then, $v_x(\pi) > 0$ because she forms a coalition with her unique friend. By assumption, x has no friend in any other coalition. Therefore, x has no MOS deviation either.

We may therefore assume that G is a directed acyclic graph. Hence, there exists an agent $x \in N$ with in-degree 0. If x has no friend, let $T = \{x\}$. If x has a friend y, we claim that there exists an agent w such that (i) w is the friend of at least one agent and (ii) every agent that has w as a friend has in-degree 0, i.e., such agents are not the friend of any agent. We provide a simple linear-time algorithm that finds such an agent. Starting with $w = y$, we will maintain a tentative agent w that will continuously fulfill (i) and update w until this agent also fulfills (ii). Note that the starting agent $w = y$ fulfills (i) because y is a friend of x. Now, consider any tentative agent w. If w is the friend of some agent z that is herself the friend of some other agent, update $w = z$. For the finiteness (and efficient computability) of this procedure, consider a topological order σ of the agents N in the directed acyclic graph G (Kahn, 1962), i.e., a function $\sigma \colon N \to [n]$ such that $\sigma(a) < \sigma(b)$ whenever $(a, b) \in A$. Note that if

w is replaced by the agent z in the procedure, then $\sigma(z) < \sigma(w)$. Hence, w is replaced at most n times, and our procedure finds the desired agent w after a linear number of steps. Now, define $T = \{a \in N : w \text{ reachable by } a\}$, i.e., T contains precisely w and all agents that have w as a friend.

We are ready to find the MOS partition. By induction, we find a partition π' that is majority-out stable for the subgame induced by $N \setminus T$. Consider $\pi = \pi' \cup \{T\}$. Then, $a \in T \setminus \{w\}$ has no incentive to deviate, because she has no friend in any other coalition and has w as a friend. Also, w is not allowed to perform a deviation, because the nonempty set of agents $T \setminus \{w\}$ unanimously prevents that. Possible deviations by agents in $N \setminus T$ can be excluded as in the first part of the proof because these agents have no friend in T. Together, we have completed the induction step and found an MOS partition. □

On the other hand, it is NP-complete to decide whether an AFG contains an MOS partition. For a proof, we use the game constructed in Proposition 4 as a gadget in a greater game. The difficulty is to preserve bad properties about the existence of MOS partitions because the larger game might allow for new possibilities to create coalitions with the agents in the counterexample.

Theorem 10. *Deciding whether an AFG contains an MOS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let $(R, \mathcal{S})$ be an instance of E3C. We define an ASHG (N, v) as follows. Let $N = N_R \cup N_\mathcal{S}$ where

- $N_R = \cup_{r \in R} N_r$ with $N_r = \{a_i^r, b_i^r, c_i^r : i \in [5]\} \cup \{z^r\}$ for $r \in R$ and
- $N_\mathcal{S} = \cup_{S \in \mathcal{S}} N_S$ with $N_S = \{e_r^S : r \in S\} \cup \{e_0^S\}$ for $S \in \mathcal{S}$.

In the whole proof, we read indices of agents a_i^r, b_i^r, and c_i^r modulo 5, mapping to the representative in the set [5].

We define utilities v as follows:

- For all $S \in \mathcal{S}$, $r \in S$: $v_{e_r^S}(e_0^S) = n$.
- For all $S \in \mathcal{S}$, $r, r' \in S$, $r \neq r'$: $v_{e_r^S}(e_{r'}^S) = n$.
- For all $S \in \mathcal{S}$, $r \in S$: $v_{e_r^S}(a_1^r) = n$.
- For all $r \in R$, $i \in [5]$, and $x \in \{a, b, c\}$: $v_{x_i^r}(x_{i+1}^r) = n$.
- For all $r \in R$, $x \in \{a, b, c\}$: $v_{x_1^r}(z^r) = n$.
- All other valuations are -1.

An illustration of the reduction is provided in Figure 10. As in previous reductions, the reduced instance consists of two types of gadgets. The elements in the ground set R are represented by R-gadgets which are subgames identical to the counterexample in Proposition 4. The sets in $\mathcal{S}$ are represented by $\mathcal{S}$-gadgets consisting of a triple of agents representing its elements in R that are each linked

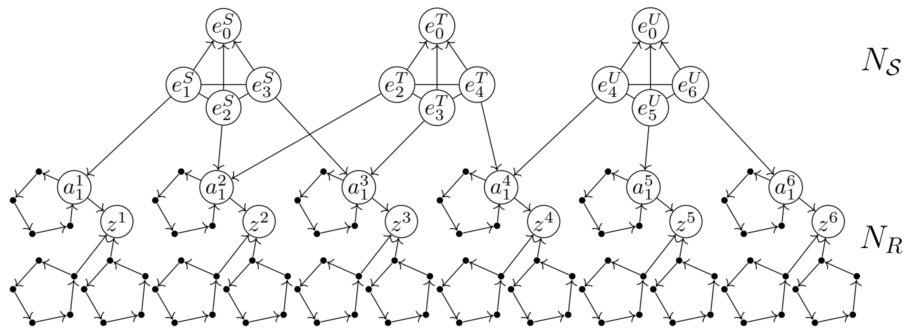

Figure 10: Schematic of the reduction from the proof of Theorem 10. We depict the reduced instance for the instance $(R, \mathcal{S})$ of E3C where $R = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{S} = \{S, T, U\}$ with $S = \{1, 2, 3\}$, $T = \{2, 3, 4\}$, and $U = \{4, 5, 6\}$. Directed edges indicate a utility of n, undirected edges a mutual utility of n, and missing edges a utility of -1. Every element in R is represented by a gadget identical to the game in Proposition 4.

to their respective R-gadget. Furthermore, in every $\mathcal{S}$-gadget, there is one special agent without any friends attracting the other agents in this $\mathcal{S}$-gadget.

We claim that $(R, \mathcal{S})$ is a Yes-instance if and only if the reduced AFG contains an MOS partition.

$\Longrightarrow$ Suppose first that $\mathcal{S}' \subseteq \mathcal{S}$ partitions R. We define a partition π by taking the union of the following coalitions:

- For $r \in R$, $x \in \{a, b, c\}$, form $\{x_2^r, x_3^r\}$, $\{x_4^r, x_5^r\}$, and $\{b_1^r, c_1^r, z^r\}$.
- For $S \in \mathcal{S}', r \in S$, form $\{e_r^S, a_1^r\}$.
- For $S \in \mathcal{S}'$, form $\{e_0^S\}$.
- For $S \in \mathcal{S} \setminus \mathcal{S}'$, form N_S.

We prove that π is majority-out stable by performing a case analysis to show that no agent can perform a deviation.

- For $r \in R$ and $x \in \{a, b, c\}$, the agents x_3^r and x_5^r are not allowed to perform an MOS deviation. Moreover, the agents x_2^r and x_4^r are in their most preferred coalitions and have, therefore, no incentive to perform a deviation.
- For $r \in R$, the agents a_1^r and z^r are not allowed to perform an MOS deviation.
- For $r \in R$ and $x \in \{b, c\}$, the agent x_1^r has no incentive to deviate. It holds that $v_{x_1^r}(\pi) = n - 1$, whereas no deviation increases her utility. In particular, joining $\pi(x_2^r)$ only yields the same utility.
- For $S \in \mathcal{S}$ and $r \in S$, the agent e_r^S has at most one friend after any possible deviation. However, she has at least two friends in π, and therefore no incentive to perform a deviation.

- For $S \in \mathcal{S}'$, the agent e_0^S is in her most preferred coalition and has no incentive to perform a deviation. Finally, for $S \in \mathcal{S} \setminus \mathcal{S}'$, the agent e_0^S is not allowed to perform an MOS deviation.

$\Longleftarrow$ Conversely, assume that the reduced instance contains an MOS partition π. We show that it originates from a Yes-instance. We split the proof into several claims.

Claim 1. *For all $S \in \mathcal{S}$, it holds that $\pi(e_0^S) = \{e_0^S\}$ or $N_S \subseteq \pi(e_0^S)$.*

Proof. Let $S \in \mathcal{S}$, say $S = \{u, w, x\}$, and define $C = \pi(e_0^S)$ and $D = \{e_u^S, e_w^S, e_x^S\}$. Assume that $C \supsetneq \{e_0^S\}$. Note that since e_0^S has no friends, she would prefer to stay in a singleton coalition. Hence, $C \supsetneq \{e_0^S\}$ implies that some agent has to be against her potential MOS deviation to a single coalition, and therefore $C \cap D \neq \emptyset$, say $e_u^S \in C$.

Assume for contradiction that $D \setminus C \neq \emptyset$, say $e_w^S \notin C$. Then, $e_x^S \in \pi(e_w^S)$. Indeed, if $e_x^S \notin \pi(e_w^S)$, then e_w^S has at most one friend in her coalition, and no agent would prevent her from performing an MOS deviation to join C. Hence, $e_x^S \in \pi(e_w^S)$. Then, $C = \{e_0^S, e_u^S\}$, as e_0^S could leave her coalition to form a singleton coalition if any other agent was part of it. But then, e_u^S has an incentive to join $\pi(e_w^S)$, and could perform a valid MOS deviation to do so. This is a contradiction and, therefore, $D \subseteq C$. $\lhd$

In the next claim, we improve upon Claim 1 and show that there are in fact only two possible coalitions for e_0^S.

Claim 2. *For all $S \in \mathcal{S}$, it holds that $\pi(e_0^S) = \{e_0^S\}$ or $\pi(e_0^S) = N_S$.*

Proof. Let $S \in \mathcal{S}$ and define $C = \pi(e_0^S)$. Assume that $C \supsetneq \{e_0^S\}$. By Claim 1, it holds that $N_S \subseteq C$ and since e_0^S has an NS deviation to form a singleton coalition, even $|C| \leq 6$. This means, in particular, that every agent $y \in C \setminus N_S$ must have a friend in C. Indeed, if this was not the case, then such an agent y would like to deviate to form a singleton coalition, and this is an MOS deviation as it is supported by at least three agents in N_S. Hence, $C \setminus N_S \neq \emptyset$ can only happen if there are two more agents in C who are a friend of each other. By the design of the utilities, the only possibility for this to happen is that there exists $t \in \mathcal{S}$ with $t \neq s$ and $u, v \in t$ with $C = N_S \cup \{t_u, t_v\}$. Then, by Claim 1, $\{t_0\} \in \pi$, implying that t_u has an MOS deviation to join t_0. This is a contradiction, and we can therefore conclude that $\pi(e_0^S) = N_S$. $\lhd$

Next, we consider the coalitions of other agents in gadgets related to sets in $\mathcal{S}$.

Claim 3. *For all $S \in \mathcal{S}$ and $r \in R$, it holds that $\pi(e_r^S) = \{e_r^S, a_1^r\}$ or $N_S \setminus \{e_0^S\} \subseteq \pi(e_r^S)$.*

Proof. Let $S \in \mathcal{S}$, say $S = \{r, u, w\}$, and define $C = \pi(e_r^S)$. If $e_0^S \in C$, then $C = N_S$ by Claim 2 and the assertion is true. Suppose, therefore, that $e_0^S \notin C$. Assume now that there is $x \in S$ with $e_x^S \notin C$, say $e_u^S \notin C$. If $e_w^S \notin C$, then no agent in C has e_r^S as a friend and could therefore vote against a deviation.

Moreover, since the deviation of e_r^S to join e_0^S is not an MOS deviation, it must be the case that $v_{e_r^S}(\pi) = n$, which can, under the given assumptions, only be the case if $\pi(e_r^S) = \{e_r^S, a_1^r\}$.

It remains to consider the case that $e_w^S \in C$. But then, e_u^S is in a coalition with at most one friend (note that it is excluded that $e_0^S \in \pi(e_u^S)$ by Claim 2) and no agent in her coalition has her as a friend. Hence, e_u^S has an MOS deviation to join C, a contradiction. Together, we have shown that if there is $x \in S$ with $e_x^S \notin C$, then $\pi(e_r^S) = \{e_r^S, a_1^r\}$, which proves this claim. $\lhd$

In the next claim, we gain even more insight on the coalitions of agents of the type e_r^S.

Claim 4. *For all $S \in \mathcal{S}$, $r \in S$, and $u \in R$, it holds that if $\pi(e_r^S) \cap N_u \neq \emptyset$, then $r = u$ and $\pi(e_r^S) = \{e_r^S, a_1^u\}$.*

Proof. Let $S \in \mathcal{S}$, $r \in S$, and $u \in R$. The assertion is true if $\pi(e_r^S) = \{e_r^S, a_1^r\}$. Hence, by Claim 3, we may assume that $N_S \setminus \{e_0^S\} \subseteq C$. We will show that $\pi(e_r^S) \cap N_u = \emptyset$. First, note that since z^u has an NS deviation to form a singleton coalition whenever she is not in such a coalition already and because only three agents have z^u as a friend, it holds that z^u forms a coalition with at most two agents that have her as an enemy. This implies in particular that $z^u \notin C$ and that $|\pi(z^u)| \leq 6$.

Assume for contradiction that there exists an agent $y \in N_u \cap C$. We already know that $y \neq z^u$. Next, if $y \neq a_1^u$, then y must have a friend in C. Indeed, at most one agent in C can have y as a friend, but the three agents in $N_S \setminus \{e_0^S\}$ favor y to leave. Hence, y could perform an MOS deviation to form a singleton coalition, otherwise. In addition, if $y = a_1^u$, then y must also have a friend in C. Note that at most two agents in $(N_u \cup N_S) \cap C$ favor her to stay while all other agents in $(N_u \cup N_S) \cap C$ (of which there are at least 2 agents) favor her to leave. The only possibility that there is another agent who favors a_1^u to stay is if there exists $t \in \mathcal{S}$ with $u \in t$ and $t_u \in C$. But then, Claim 3 implies that $N_t \setminus \{t_0\} \subseteq C$, a majority of which favors a_1^u to leave. Together, a_1^u is favored to leave C by a (weak) majority of agents. Therefore, she must not have an incentive to form a singleton coalition and, therefore, has a friend in C.

Now, assume that there exists $x \in \{a, b, c\}$ and $i \in [5]$ with $x_i^u \in C$. Then, our previous observation implies that $\{x_i^u : i \in [5]\} \subseteq C$. Hence, $|C| \geq 8$ and therefore $v_{x_1^u}(\pi) \leq n - 6 < n - 5 \leq v_{x_1^u}(\pi(z^u) \cup \{x_1^u\})$. Hence, x_1^u could perform an MOS deviation. This is a contradiction. Therefore, we have shown that $\pi(e_r^S) \cap N_u = \emptyset$. $\lhd$

Now, we show that coalitions of agents in different sets of the type N_r are disjoint.

Claim 5. *For all $r, u \in R$ and agents $w \in N_r, y \in N_u$, it holds that $\pi(w) \cap \pi(y) = \emptyset$.*

Proof. Let $r, u \in R$ and assume for contradiction that there exist agents $w \in N_r$ and $y \in N_u$ with $\pi(w) = \pi(y)$. Define $C = \pi(w)$. By Claim 2 and Claim 4, it holds that $C \cap N_S = \emptyset$ for all $S \in \mathcal{S}$. We may assume without loss of generality that $|C \cap N_r| \leq |C \cap N_u|$. Since every agent in $C \cap N_r$ is preferred to leave by a majority of agents in C, it holds that $z^r \notin C$ and every agent in $C \cap N_r$ must have a friend in C. The remaining proof of this step is similar to the proof of Claim 4. Let $x \in \{a, b, c\}$ and $i \in [5]$ with $x_i^r \in C$. Then, $\{x_i^r : i \in [5]\} \subseteq C$ and therefore $|C| \geq 10$. As in the previous claim, $|\pi(z^r)| \leq 6$. Hence, $v_{x_1^r}(\pi) \leq n - 8 < n - 5 \leq v_{x_1^r}(\pi(z^r) \cup \{x_1^r\})$, a contradiction. $\lhd$

Finally, we can conclude the proof by showing that there exists $\mathcal{S}' \subseteq \mathcal{S}$ partitioning R. Therefore, let $\mathcal{S}' = \{S \in \mathcal{S} : \pi(e_r^S) = \{e_r^S, a_1^r\}$ for some $r \in S\}$. We show that $\mathcal{S}'$ partitions R by showing that it covers all elements from R and that its elements are disjoint sets.

For the first part, let $r \in R$. By the proof of Proposition 4, if $\pi(y) \subseteq N_r$ for all $y \in N_r$, then the partition π is not MOS. Hence, some agent in N_r must form a coalition with an agent outside of N_r. Combining Claim 2, Claim 4, and Claim 5, this can only be the case if there exists $S \in \mathcal{S}$ with $r \in S$ and $\pi(e_r^S) = \{e_r^S, a_1^r\}$. Consequently, $\mathcal{S}'$ covers R.

For the second part, assume for contradiction that some element in R is covered at least twice by sets in $\mathcal{S}'$. Then, there exists $S \in \mathcal{S}'$ with $r \in S$ and $\{e_r^S, a_1^r\} \notin \pi$. By Claim 3, $N_S \setminus \{e_0^S\} \subseteq \pi(e_r^S)$. But then, according to the definition of $\mathcal{S}'$, it follows that $S \notin \mathcal{S}'$, a contradiction. Hence, the elements of $\mathcal{S}'$ are disjoint sets. This completes the proof. $\square$

5.3. Friends-And-Enemies Games

We have already seen that friends-and-enemies games contain efficiently computable stable coalition structures with respect to the unanimity-based concepts of individual stability and contractual Nash stability (cf. Corollary 1). In this section, we will see that the transition to majority-based consent crosses the boundary of tractability. The closeness to this boundary is also emphasized by the fact that it is surprisingly difficult to even construct No-instances for majority-out and majority-in stability, i.e., FEGs which do not contain an MOS or MIS partition, respectively. Indeed, the smallest such games that we can construct are games with 23 and 183 agents, respectively. We will start by considering majority-out stability.

Proposition 6. *There exists an FEG without an MOS partition.*

Proof. We start by sketching the FEG constructed for the proof and by giving some intuition why it contains no MOS partition. The instance is illustrated in Figure 11. It consists of a triangle of agents together with 4 sets of agents whose friendship relation is complete and transitive,[9] together with one additional agent each that gives a temptation for the agent of the transitive substructures with the most friends.

[9]Completeness means that for any pair of agents, at least one agent is the friend of the other agent. Transitivity means that if b is a friend of a and c is a friend of b, then c is a friend of a.

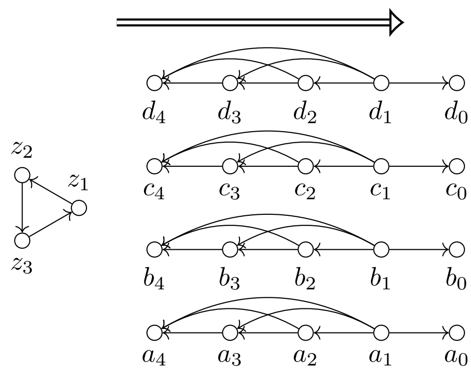

Figure 11: FEG without an MOS partition. The depicted (directed) edges represent friends. The double arrow denotes that every agent to the left of the tail of the arrow has every agent below the arrow as a friend.

An important reason for the nonexistence of MOS partitions is that there is a high incentive for the transitive structures to form coalitions. This gives incentive to agents z_i to join them. If z_1, z_2, and z_3 are in disjoint coalitions, then they would chase each other according to their cyclic structure. If they are all in the same coalition, then agents x_0 for $x \in \{a, b, c, d\}$ prevent the complete transitive structures to be part of this coalition and other transitive structures are more attractive.

We now provide the formal proof. Let $N = N_z \cup N_a \cup N_b \cup N_c \cup N_d$, where $N_z = \{z_1, z_2, z_3\}$ and $N_x = \{x_0, x_1, x_2, x_3, x_4\}$ for $x \in \{a, b, c, d\}$. Utilities are given as

- $v_x(y) = 1$ if $(x, y) \in \{(z_1, z_2), (z_2, z_3), (z_3, z_1)\}$,
- $v_{x_i}(x_j) = 1$ if $x \in \{a, b, c, d\}, i, j \in [4], i < j$,
- $v_{x_1}(x_0) = 1$ if $x \in \{a, b, c, d\}$,
- $v_{z_i}(x_j) = 1$ if $x \in \{a, b, c, d\}, i \in [3], j \in [4]$, and
- all other valuations are -1.

Assume for contradiction that this FEG admits an MOS partition π. We will derive a contradiction in 4 steps. First, Claim 6 describes possible coalitions of agents x_0 where $x \in \{a, b, c, d\}$. Second, Claim 7 establishes that coalitions from agents of different sets of N_x, $x \in \{a, b, c, d\}$, are disjoint. Then, Claim 8 excludes that all agents in N_z are in a joint coalition. Finally, we complete the proof by performing a case analysis for two disjoined coalitions containing different agents from N_z.

Claim 6. *It holds that* $\pi(x_0) \in \{\{x_0\}, \{x_0, x_1\}\}$ *for* $x \in \{a, b, c, d\}$.

Proof. Let $x \in \{a, b, c, d\}$ and suppose that $|\pi(x_0)| > 1$. Then, x_0 has an NS deviation to form a singleton. The claim follows because the only agent that prevents her to leave the coalition is x_1. $\lhd$

Claim 7. *It holds that $x_i \notin \pi(y_j)$ for $x, y \in \{a, b, c, d\}, x \neq y$, and $i, j \in [4]$.*

Proof. Assume for contradiction that there exist $x, y \in \{a, b, c, d\}, x \neq y$, and $i, j \in [4]$ with $x_i \in \pi(y_j)$. Without loss of generality, $x = a$ and $y = b$. Define $\Gamma = \pi(b_j)$. Again, without loss of generality, we may assume that $|\Gamma \cap N_a| \geq |\Gamma \cap N_b|$. Let $j^* = \min\{j \in [4] \colon b_j \in \Gamma\}$.

By Claim 6, $x_0 \notin \Gamma$ for $x \in \{a, b, c, d\}$. Hence, b_{j^*} wants to perform an NS deviation to form a singleton and is only favored to stay by agents in N_z. As $a_i \in F_{\text{out}}(\Gamma, b_{j^*})$, at least two agents must favor b_{j^*} to stay. We conclude that

- $|\Gamma \cap N_z| \geq 2$ and $(*)$
- $|\Gamma \setminus N_z| \leq 3$. $(**)$

There, $(**)$ follows because at most 3 agents favor b_{j^*} to stay, and she can therefore have at most two enemies. To conclude this step, we distinguish two cases.

<u>Case 1</u>: It holds that $|N_z \cap \Gamma| = 3$, i.e., $N_z \subseteq \Gamma$.

We now consider the agents in N_c. By Claim 6, $(*)$, and $N_z \subseteq \Gamma$, we derive that $\pi(c_i) \subseteq N_c \setminus \{c_0\}$ for $i = 2, 3, 4$, and $\pi(c_1) \subseteq N_c$. If $\pi(c_1) = \{c_0, c_1\}$, then there is a coalition of size at least 2 consisting of agents in $C \setminus \{c_0, c_1\}$, and c_1 could perform an MOS deviation to join them. Hence, using Claim 6, it follows that $\pi(c_1) \subseteq C \setminus \{c_0\}$.

Let $\Phi \subseteq C \setminus \{c_0\}$ be a coalition of the largest size. Note that $C \setminus \{c_0\}$ cannot contain (at least) 2 singleton coalitions. Then, the singleton with the lower index would join the other singleton. If $|\Phi| = 2$, then $C \setminus \{c_0\}$ consists of two pairs, and c_1 has an MOS deviation to join the other pair. Next, assume that $|\Phi| = 3$. If c_1 or c_2 remain as a singleton, they would join Φ. If c_3 or c_4 remain as a singleton, then c_2 performs an MOS deviation to join her. This leaves only the case $|\Phi| = 4$ and we can conclude that $C \setminus \{c_0\} \in \pi$. But then, by $(**)$, z_k has an MOS deviation to join $C \setminus \{c_0\}$ for $k \in [3]$, a contradiction. This concludes Case 1.

<u>Case 2</u>: It holds that $|N_z \cap \Gamma| = 2$.

Then, $|\Gamma \setminus N_z| \leq 2$ which means that $\Gamma \setminus N_z = \{a_i, b_j\}$ and it follows that $\Gamma \cap N_c = \Gamma \cap N_d = \emptyset$. Let $k^* \in [3]$ be the unique index with $z_{k^*} \notin \Gamma$, say without loss of generality $k^* = 1$. Using $(*)$, it must also be the case that $\pi(z_1) \cap N_c = \emptyset$ or $\pi(z_1) \cap N_d = \emptyset$, say without loss of generality $\pi(z_1) \cap N_c = \emptyset$. The identical arguments as in the previous case show that $C \setminus \{c_0\} \in \pi$. But then z_3 could perform an MOS deviation to join $C \setminus \{c_0\}$, a contradiction. This concludes Case 2 and, therefore, the proof of the claim. $\lhd$

Claim 8. *There exists no $\Gamma \in \pi$ with $N_z \subseteq \Gamma$.*

Proof. Assume for contradiction that there exists $\Gamma \in \pi$ with $N_z \subseteq \Gamma$. By Claim 6 and Claim 7, there exists $x \in \{a, b, c, d\}$ with $\Gamma \subseteq N_z \cup N_x$. Without loss of generality, assume that $\Gamma \subseteq N_z \cup N_a$. By Claim 6, $a_0 \notin \Gamma$. We claim that $|\Gamma \cap N_a| \leq 3$. For the contrary, assume that $|\Gamma \cap N_a| = 4$. Then, Claim 6 implies that $\{a_0\} \in \pi$. Also, $v_{a_1}(\pi) = 0$ and $|F_{\text{in}}(\Gamma, a_1)| = |N_z| = |\{a_2, a_3, a_4\}| = |F_{\text{out}}(\Gamma, a_1)|$. Hence, a_1 can perform an MOS deviation to join $\{a_0\}$, a contradiction. Thus, $|\Gamma \cap N_a| \leq 3$, as claimed.

As in the proof of Claim 7, we can show that $B \setminus \{b_0\} \in \pi$. But then z_k has an MOS deviation to join this coalition for every $k \in [3]$, a contradiction. This concludes the proof of this claim. $\lhd$

We are ready to obtain a final contradiction. By Claim 8, there exist $i, j \in [3]$ with $z_i \notin \pi(z_j)$. Without loss of generality, we may assume that $i = 2$ and $j = 1$.

<u>Case 1</u>: It holds that $z_3 \in \pi(z_2)$.

By Claim 6, $v_{z_k}(x) = 1$ for all $k \in [3], x \in (\pi(z_1) \cup \pi(z_2)) \setminus N_z$. Let $m_1 = |\pi(z_2)| - 2 = |\pi(z_2) \setminus N_z|$ and $m_2 = |\pi(z_1)| - 1 = |\pi(z_1) \setminus N_z|$.

If $m_2 \geq m_1$, then z_3 can perform an NS deviation to join $\pi(z_1)$. This is also an MOS deviation unless $\pi(z_2) = \{z_2, z_3\}$. But in this case, we find a coalition of the form $N_x \setminus \{x_0\}$ for some $x \in \{a, b, c, d\}$ as in the previous steps. Then, z_2 has an MOS deviation to join this coalition.

On the other hand, if $m_2 < m_1$, then z_1 can perform an MOS deviation to join $\pi(z_2)$. This concludes Case 1. By symmetry, this covers even all cases where at least two agents from N_z are in the same coalition. Hence, it remains one final case.

<u>Case 2</u>: The agents in N_z are in pairwise disjoint coalitions.

Let $p_k = |\pi(z_k)|$ for $k \in [3]$ and $k^* = \arg\max_{k \in [3]} p_k$. Without loss of generality, $k^* = 1$. As in the previous case, it follows from Claim 6 that $v_{z_k}(x) = 1$ for all $k \in [3], x \in \bigcup_{l \in [3]} \pi(z_l) \setminus N_z$. But then z_3 has an MOS deviation to join $\pi(z_1)$. This is the final contradiction and completes the proof. $\square$

In the previous proof, it was particularly useful to establish disjoint coalitions of groups such that agents from one group dislike the agents of all other groups. By contrast, if we make the further assumption that one agent from every pair of agents likes the other agent, then this does not work anymore, and the grand coalition is majority-out stable. More formally, we say that the friendship relation of an ASHG (N, v) is *complete* if, for every pair of agents $i, j \in N$ with $i \neq j$, it holds that $v_i(j) > 0$ or $v_j(i) > 0$. Note that the next proposition is not true for other stability concepts, such as Nash stability or even individual stability.

Proposition 7. *The grand coalition is majority-out stable in every FEG with complete friendship relation.*

Proof. Let (N, v) be an FEG with complete friendship relation, and let π be the grand coalition. We claim that π is majority-out stable. Suppose that there is an agent $x \in N$ who can perform an NS deviation to form a singleton.

Then, $v_x(N) < 0$ and therefore $|\{y \in N \setminus \{x\}\colon v_x(y) = -1\}| > \{y \in N \setminus \{x\}\colon v_x(y) = 1\}|$. Hence,

$$\begin{aligned} |F_{\text{in}}(N,x)| &\geq |\{y \in N \setminus \{x\}\colon v_x(y) = -1\}| \\ &> |\{y \in N \setminus \{x\}\colon v_x(y) = 1\}| \\ &\geq |F_{\text{out}}(N,x)|. \end{aligned}$$

In the first inequality, we use that x is a friend of all of her enemies. In the final inequality, we use that x can only be an enemy of her friends. Hence, x is not allowed to perform an MOS deviation. □

Still, the nonexistence of MOS partitions in FEGs shown in Proposition 6 can be leveraged to prove an intractability result. Towards the hardness reduction, we start with a useful lemma. It lets us separate the agent set into subsets such that agents from different subsets cannot form joint coalitions within an MOS partition.

Lemma 3. *Consider an FEG (N, v) with an MOS partition π. Let $i, j \in N$ be two agents with $v_i(j) = v_j(i) = -1$ and assume that, for every agent $k \in N \setminus \{i, j\}$, it holds that*

- $v_i(k) = -1$ *or* $v_j(k) = -1$*,*
- $v_k(i) = -1$ *or* $v_k(j) = -1$*,*
- $v_k(i) = -1$ *whenever* $v_j(k) = 1$*, and*
- $v_k(j) = -1$ *whenever* $v_i(k) = 1$*.*

Then, $i \notin \pi(j)$.

Proof. Let an FEG (N, v) be given together with an MOS partition π, and let $i, j \in N$ be two agents satisfying the assumptions of the lemma. Assume for contradiction that $i \in \pi(j)$, and define $C = \pi(j)$. We will use the first assumption of the lemma to show that either i or j can perform an NS deviation to form a singleton coalition, and the other conditions to argue that there is even a valid MOS deviation. First, note that the first assumption implies that, for every agent $k \in N \setminus \{i, j\}$, it holds that $v_i(k) + v_j(k) \leq 0$. Hence,

$$v_i(\pi) + v_j(\pi) = -2 + \sum_{k \in \pi(j) \setminus \{i,j\}} v_i(k) + v_j(k) \leq -2.$$

Therefore, $v_i(\pi) < 0$ or $v_j(\pi) < 0$, and thus either i or j can perform an NS deviation to form a singleton coalition.

In addition, our second assumption implies that, for every agent $k \in N \setminus \{i, j\}$, it holds that $k \in F_{\text{out}}(C, i)$ or $k \in F_{\text{out}}(C, j)$. Hence, a similar averaging argument as the previous consideration shows that $|F_{\text{out}}(C, i)| > |C|/2$ or $|F_{\text{out}}(C, j)| > |C|/2$.

Assume first that $v_i(\pi) < 0$ and $v_j(\pi) < 0$. Then, our second observation implies that one of i and j can perform an MOS deviation to form a singleton coalition, a contradiction. Hence, we may assume without loss of generality that $v_i(\pi) < 0$ and $v_j(\pi) \geq 0$. Then,

$$
\begin{aligned}
&|F_{\text{in}}(C,i)| - |F_{\text{out}}(C,i)| \\
&= |\{l \in C \setminus \{i\} \colon v_l(i) = 1\}| - |\{l \in C \setminus \{i\} \colon v_l(i) = -1\}| \\
&\leq |\{l \in C \setminus \{i\} \colon v_j(l) = -1\}| - |\{l \in C \setminus \{i\} \colon v_j(l) = 1\}| = -v_j(\pi) \leq 0.
\end{aligned}
$$

In the inequality, we have used the third assumption of the lemma (the fourth assumption is necessary for the symmetric case where i and j swap roles). Hence, agent i can perform an MOS deviation to form a singleton coalition. This is a contradiction and we can conclude that $i \notin \pi(j)$. □

We are ready for the proof of the hardness result. Interestingly, in contrast to the proofs of Theorem 3 and Theorem 10, the next theorem merely uses the existence of an FEG without an MOS partition to design a gadget and does not exploit the specific structure of a known counterexample.

Theorem 11. *Deciding whether an FEG contains an MOS partition is NP-complete.*

Proof. We provide another reduction from E3C. Let $(R, \mathcal{S})$ be an instance of E3C. We define a reduced FEG (N, v) as follows. By Proposition 6, there exists an FEG without an MOS partition and we may assume that (N', v') is such an FEG with the additional property that there exists an agent $x \in N'$ such that the FEG restricted to $N' \setminus \{x\}$ contains an MOS partition π'. Indeed, an FEG with the additional property can be obtained simply by removing agents until the property is satisfied.

Now, let $N = N_{\mathcal{S}} \cup N_R$ where

- $N_{\mathcal{S}} = \cup_{S \in \mathcal{S}} N_S$ with $N_S = \{a_0^S\} \cup \{a_r^S \colon r \in S\}$ for $S \in \mathcal{S}$ and
- $N_R = \cup_{r \in R} N_r$ with $N_r = \{b^r \colon b \in N'\}$ for $r \in R$.

Specifically, we denote the agent corresponding to the special agent $x \in N'$ by x^r. Agents of the type a_r^S will receive a positive utility from forming a coalition with x^r, and therefore have the capability of forcing x^r to stay in a coalition of size 2 with them.

We define utilities v as follows:

- For all $S \in \mathcal{S}$, $b, c \in N_S$: $v_b(c) = 1$.
- For all $S \in \mathcal{S}$, $r \in S$: $v_{a_r^S}(x^r) = 1$.
- For all $r \in R$ and $b, c \in N'$: $v_{b^r}(c^r) = v'_b(c)$, i.e., the internal valuations for agents in N_r are identical to the valuations in the counterexample (N', v').
- All other valuations are -1.

The reduced instances look very similar to the ones of Theorem 10 depicted in Figure 10. The key differences are that the agents in every N_S now form cliques, and that we replace the gadgets formed by the agents in N_R by the No-instance (N', v').

We claim that $(R, \mathcal{S})$ is a Yes-instance if and only if the reduced FEG contains an MOS partition.

$\Longrightarrow$ Suppose first that $\mathcal{S}' \subseteq \mathcal{S}$ partitions R. We define a partition π as follows.

- For $S \in \mathcal{S} \setminus \mathcal{S}'$, we have $N_S \in \pi$ and for $S \in \mathcal{S}'$, we have $\{a_0^S\} \in \pi$.
- For $S \in \mathcal{S}'$, $r \in S$, we have $\{a_r^S, x^r\} \in \pi$.
- For $r \in R$ and $b \in N' \setminus \{x\}$, we have $\pi(b^r) = \{b^r : b \in \pi'(x)\}$.

Recall that π' is the MOS partition in (N', v') after removing x.

We claim that the partition π is majority-out stable.

- Let $r \in R$ and consider an agent $b \in N' \setminus \{x\}$. Then, b^r cannot perform an MOS deviation to join $\pi(c^r)$ for any $c \in N' \setminus \{x\}$, because π' restricted to $N' \setminus \{x\}$ is an MOS partition. Moreover, joining $\pi(c)$ for any $c \in N \setminus N_r$ yields utility at most 0 (in fact, the only such coalition that b^r could join to obtain a utility of 0 is $\pi(x^r)$). Hence, if this constituted an MOS deviation, then forming a singleton coalition would also be an MOS deviation, contradicting the fact that π' is an MOS partition.
- Let $r \in R$. Then, x^r is not allowed to leave her coalition by means of an MOS deviation.
- Let $S \in \mathcal{S}'$. Then $v_{a_0^S}(\pi) = 0$ and joining any other coalition yields utility at most 0. In particular, $v_{a_0^S}(\pi(a_r^S) \cup \{a_0^S\}) = 0$ for all $r \in S$. Moreover, for $r \in S$, $v_{a_r^S}(\pi) = 1$ and joining any other coalition yields utility at most 1. In particular, $v_{a_r^S}(\pi(a_0^S) \cup \{a_r^S\}) = 1$.
- Let $S \in \mathcal{S} \setminus \mathcal{S}'$. Then, $\pi(a_0^S)$ is the most preferred coalition of a_0^S and she has no incentive to perform an MOS deviation. Moreover, for $r \in S$, $v_{a_r^S}(\pi) = 3$ and joining any other coalition yields a utility of at most 0.

Together, we have shown that π is an MOS partition.

$\Longleftarrow$ For the reverse implication, assume that π is an MOS partition for the reduced instance (N, v). We start by determining the coalitions of agents of the type a_0^S.

Claim 9. *Let $S \in \mathcal{S}$. Then, $\pi(a_0^S) = \{a_0^S\}$ or $\pi(c) \subseteq N_S$ for all $c \in N_S$.*

Proof. Let $S \in \mathcal{S}$ and define $C = \pi(a_0^S)$. A close inspection of the utilities in the definition of the reduced instance lets us apply Lemma 3 multiple times to conclude that

- for all $S' \in \mathcal{S} \setminus \{S\}$, $C \cap N_{S'} = \emptyset$,

- for all $r \in R \setminus S$, $C \cap N_r = \emptyset$, and
- for all $r \in S$, $C \cap N_r \subseteq \{x^r\}$.

Together, $C \subseteq N_S \cup \{x^r : r \in S\}$. Even more, for $r \in S$, if $x^r \in C$, then $v_{x^r}(\pi) < 0$. In addition, $F_{\text{in}}(C, x^r) \subseteq \{a_r^S\}$ and $a_0^S \in F_{\text{out}}(C, x^r)$. Hence, x^r could perform an MOS deviation to form a singleton coalition. We can therefore conclude that $C \subseteq N_S$.

Assume that $C \supsetneq \{a_0^S\}$. If $|C| = 3$, then there exists a unique $r \in S$ with $a_r^S \notin C$. Since a_r^S has only one friend outside C, this would imply that $v_{a_r^S}(\pi) \leq 1$ whereas $v_{a_r^S}(C \cup \{a_r^S\}) = 3$ and $F_{\text{in}}(\pi(a_r^S), a_r^S) = \emptyset$. Hence, a_r^S could perform an MOS deviation to join C, a contradiction. Hence, $|C| = 2$ or $|C| = 4$. As the latter case corresponds to the situation that $C = N_S$, we only need to consider the former case.

Suppose that $S = \{r_1, r_2, r_3\}$ and that $C = \{a_0^S, a_{r_1}^S\}$. Then, it holds that $a_{r_3}^S \in \pi(a_{r_2}^S)$, as otherwise $v_{a_{r_2}^S}(\pi) \leq 1$ whereas $v_{a_r^S}(C \cup \{a_{r_2}^S\}) = 3$ and $F_{\text{in}}(\pi(a_{r_2}^S), a_{r_2}^S) = \emptyset$. But then, $\pi(a_{r_2}^S) = \{a_{r_2}^S, a_{r_3}^S\}$. Any other agent would only have enemies in $\pi(a_{r_2}^S)$ and is allowed to leave by a weak majority. This concludes the proof of the claim. $\triangleleft$

Our next claim investigates elements $S \in \mathcal{S}$ for which $\{a_0^S\} \in \pi$.

Claim 10. *Let $S \in \mathcal{S}$ such that $\{a_0^S\} \in \pi$. Then, for every $r \in S$, it holds that $\pi(a_r^S) = \{a_r^S, x^r\}$.*

Proof. Let $S \in \mathcal{S}$ with $\{a_0^S\} \in \pi$ and consider any $r \in S$. Define $C = \pi(a_r^S)$ and assume for contradiction that there exists $r' \in S$ with $r' \neq r$ and $a_{r'}^S \in C$. We can combine the following observations:

- Claim 9 shows that $a_0^{S'} \notin C$ for every $S' \in \mathcal{S} \setminus \{S\}$.
- Let $\hat{r} \in R$. We can apply Lemma 3 for a_r^S (or $a_{r'}^S$) and an agent in $N_{\hat{r}}$ to show that $C \cap N_{\hat{r}} = \emptyset$ if $\hat{r} \neq r$ (or if $\hat{r} = r$).
- Let $S' \in \mathcal{S}$ and $\hat{r} \in S'$. We can apply Lemma 3 for a_r^S (or $a_{r'}^S$) and $a_{\hat{r}}^{S'}$ to show that $a_{\hat{r}}^{S'} \notin C$ if $\hat{r} \neq r$ (or $\hat{r} = r$).

Together, the observations show that $C \subseteq N_S$. But then, $v_{a_0^S}(C \cup \{a_0^S\}) \geq 2$, and a_0^S could perform an MOS deviation to join C. This is a contradiction and we can conclude that $C \cap N_S = \{a_r^S\}$.

This means in particular, that $F_{\text{in}}(C, a_r^S) = \emptyset$. Since $v_{a_r^S}(\{a_0^S, a_r^S\}) = 1$, it must hold that $v_{a_r^S}(\pi) = 1$. Since the unique friend of a_r^S outside N_S is x^r, we can conclude that $\pi(a_r^S) = \{a_r^S, x^r\}$. $\triangleleft$

We are ready to finish the proof. Therefore, let $\mathcal{S}' = \{S \in \mathcal{S} : \{a_0^S\} \in \pi\}$. We show that $\mathcal{S}'$ partitions R in two steps. First, the sets in $\mathcal{S}'$ are disjoint. Indeed, if $S, S' \in \mathcal{S}'$ with $S \neq S'$ and $r \in S \cap S'$, then Claim 10 implies that $\{a_r^S, x^r\} \in \pi$ and $\{a_r^{S'}, x^r\} \in \pi$, contradicting the fact that π is a partition.

It remains to show that all elements of R are covered by a set in $\mathcal{S}'$. Therefore, consider an arbitrary $r \in R$ and let $b \in N'$. By Lemma 3, $\pi(b^r) \cap N^{r'} = \emptyset$ for all $r' \in R$ with $r' \neq r$. Moreover, Claim 9 and Claim 10 imply that $\pi(b^r) \cap N_S = \emptyset$ for all $S \in \mathcal{S}$ with $r \notin S$. Assume for contradiction that there exists no $S \in \mathcal{S}'$ with $r \in S$. Then, Claim 9 implies that $\pi(b^r) \cap N_S = \emptyset$ for all $S \in \mathcal{S}$ with $r \in S$. Together, $\pi(b^r) \subseteq N_r$. This means that π restricted to the agents in N_r is an MOS partition, contradicting the fact that such a partition does not exist. Hence, r must be covered by some set in $\mathcal{S}'$. □

Our next goal is to construct an FEG without an MIS partition. Despite a lot of structure, the game that we find, which is depicted in Figure 12, is quite large, encompassing 183 agents. Before we describe and analyze the game, we prove a useful lemma showing that certain agents in cliques of mutual friendship playing identical roles have to be in joint coalitions in every MIS partition. This will concern the agents in the sets K_i and B_i^j for $i, j \in [5]$ (cf. Figure 12).

Lemma 4. *Consider an FEG (N, v) with an MIS partition π. Let $W \subseteq N$ such that the following conditions hold:*

1. *For all $i, j \in W$: $v_i(j) = 1$.*
2. *For all $i, j \in W$, $k \in N$: $v_i(k) = v_j(k)$.*
3. *For all $i \in W$, $k \in N$: $v_i(k) = 1$ implies $v_k(i) = 1$.*

Then, there exists a coalition $C \in \pi$ with $W \subseteq C$.

Proof. Let an FEG (N, v) be given together with an MIS partition π, and let $W \subseteq N$ be a subset of agents that fulfills the three conditions of the assertion. Assume for contradiction that there exist $i, j \in W$ with $\pi(i) \neq \pi(j)$. We may assume without loss of generality that $v_i(\pi) \geq v_j(\pi)$. Consider the deviation where agent j joins $\pi(i)$. Then,

$$v_j(\pi(i) \cup \{j\}) \overset{(1),(2)}{=} 1 + v_i(\pi) > v_j(\pi).$$

Hence, this constitutes an NS deviation. Moreover, since π is majority-in stable, it holds that $v_i(\pi) \geq 0$ and therefore, because the game is an FEG,

$$|\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = 1\}| \geq |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = -1\}|. \qquad (*)$$

It follows that

$$\begin{aligned}
|F_{\text{in}}(\pi(i), j)| &\overset{(1)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_x(j) = 1\}| + 1 \\
&\overset{(3)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_j(x) = 1\}| + 1 \overset{(2)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = 1\}| + 1 \\
&\overset{(*)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = -1\}| + 1 \overset{(2)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_j(x) = -1\}| + 1 \\
&\overset{(3)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_x(j) = -1\}| + 1 = |F_{\text{out}}(\pi(i), j)| + 1 > |F_{\text{out}}(\pi(i), j)|.
\end{aligned}$$

Hence, this is even an MIS deviation, a contradiction. □

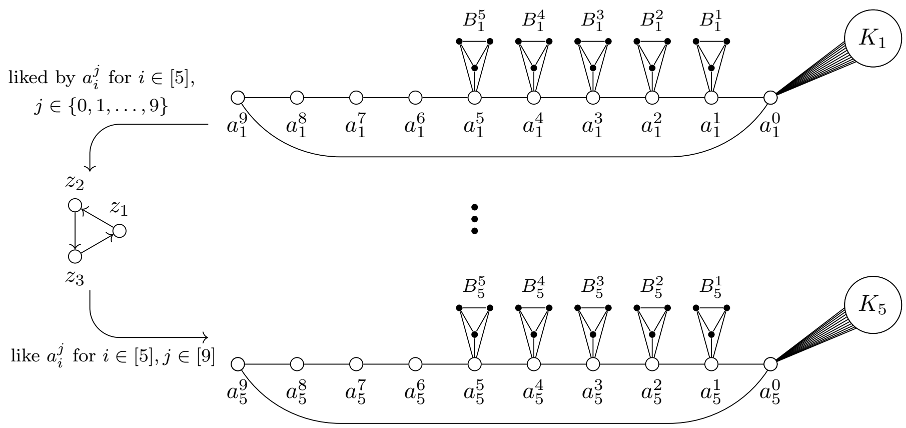

Figure 12: FEG without an MIS partition. The depicted edges represent friends. Undirected edges represent mutual friendship. For $i \in [5]$, some of the edges of agents in A_i are omitted. In fact, these agents form cliques. Also, each K_i represents a clique of 11 agents. Finally, the agents $\{z_1, z_2, z_3\}$ have edges from and towards agents a_i^j for $i \in [5]$ and $j \in [9]$. Moreover, there are edges towards them from a_i^0 but not vice versa.

We are ready to prove the next statement.

Proposition 8. *There exists an FEG without an MIS partition.*

Proof. The game we construct is illustrated in Figure 12. Before the formal description and analysis, we briefly discuss some key features. Similar to the game in Proposition 6, the central element is a directed cycle of three agents. These agents are connected to five copies of the same gadget. This gadget consists of a main clique $\{a_i^0, \ldots, a_i^9\}$ of 10 mutual friends and further cliques that cause certain temptations for agents in the main clique. Cliques are linked by agents that have an incentive to be part of two cliques, which are part of disjoint coalitions. Since it is impossible to avoid all of these potential deviations, the instance does not admit an MIS partition.

Formally, the agent set is given by $N = Z \cup \bigcup_{i \in [5]} (A_i \cup B_i \cup K_i)$, where the exact definitions and interpretation of the subsets of agents is as follows:

- The set of agents $Z = \{z_1, z_2, z_3\}$ forms a directed triangle.
- For $i \in [5]$, the sets $A_i = \{a_i^j : j = \{0, 1, \ldots, 9\}$ form cliques which are liked by agents in Z, except for the special agent a_i^0. In turn, all of them like the agents in Z.
- For $i \in [5]$, the sets $K_i = \{k_i^j : j \in [11]\}$ form cliques not liked by agents in Z, but a_i^0 likes these agents.
- For $i \in [5]$, define $B_i = \bigcup_{j=1}^5 B_i^j$, where $B_i^j = \{b_i^{j,l} : l \in [3]\}$. The set B_i^j forms a small clique which tries to tempt agent a_i^j to join if B_i^j is a coalition.

The utilities are defined as

- $v_x(y) = 1$ if $(x, y) \in \{(z_1, z_2), (z_2, z_3), (z_3, z_1)\}$,
- $v_{z_i}(a^l_j) = 1$ if $i \in [3]$, $j \in [5]$, and $l \in [9]$,
- $v_{a^j_i}(a^l_i) = 1$ if $i \in [5]$, $j, l \in \{0, 1, \dots, 9\}$,
- $v_{a^j_i}(z_l) = 1$ if $i \in [5]$, $j \in \{0, 1, \dots, 9\}$, and $l \in [3]$,
- $v_{a^0_i}(k^j_i) = v_{k^j_i}(a^0_i) = 1$ if $i \in [5]$, $j \in [11]$,
- $v_{a^j_i}(b^{j,l}_i) = 1$ if $i, j \in [5]$, $l \in [3]$,
- $v_{b^{j,l}_i}(b^{j,l'}_i) = 1$ if $i, j \in [5]$, $l, l' \in [3]$,
- $v_{k^j_i}(k^l_i) = 1$ if $i \in [5]$, $j, l \in [11]$, and
- all other valuations are -1.

Assume for contradiction that π is an MIS partition for this game. The following observation is helpful in various places:

$$\text{Every agent receives nonnegative utility in } \pi \text{, i.e., } v_i(\pi) \geq 0 \text{ for all } i \in N. \quad (*)$$

The observation is true because every agent of negative utility could perform an MIS deviation to form a singleton coalition. We will now derive a contradiction proving several claims. The first one is a direct application of Lemma 4 for the agents in sets K_i for $i \in [5]$.

Claim 11. *For all $i \in [5]$, there exists $C \in \pi$ with $K_i \subseteq C$.*

The next claim improves upon the previous claim.

Claim 12. *If $i \in [5]$, then $K_i \in \pi$ or $K_i \cup \{a^0_i\} \in \pi$.*

Proof. Let $i \in [5]$ and assume for contradiction that there exists $C \in \pi$ with $K_i \subseteq C$ and $C \setminus (K_i \cup \{a^0_i\}) \neq \emptyset$. By $(*)$, $v_{k^1_i}(\pi) \geq 0$ and therefore $|C \setminus (K_i \cup \{a^0_i\})| \leq |K_i \cup \{a^0_i\}| - 1 = 11$. Let $x \in C \setminus (K_i \cup \{a^0_i\})$. Then, $a^0_i \in C$, $|C \setminus (K_i \cup \{a^0_i\})| = 11$, and $v_x(y) = 1$ for all $y \in C \setminus (K_i \cup \{a^0_i\})$. Otherwise, x has at most 10 friends leading to $v_x(\pi) \leq 10 - |K_i| < 0$, contradicting $(*)$. Consequently, the agents $C \setminus (K_i \cup \{a^0_i\})$ form a set of 11 mutual friends which all have a^0_i as a friend. Such a set of agents does not exist, and we derive a contradiction. ◁

The next two claims make similar structural observations for the agent sets B^j_i. First, we can apply Lemma 4 again for a statement analogous to Claim 11.

Claim 13. *For all $i, j \in [5]$, there exists $C \in \pi$ with $B^j_i \subseteq C$.*

We also refine this claim.

Claim 14. *If $i, j \in [5]$, then $B_i^j \in \pi$ or $B_i^j \cup \{a_i^j\} \in \pi$.*

Proof. Let $i, j \in [5]$ and assume for contradiction that there exists $C \in \pi$ with $B_i^j \subseteq C$ and $C \setminus (B_i^j \cup \{a_i^j\}) \neq \emptyset$. If $|C \setminus (B_i^j \cup \{a_i^j\})| < 3 = |B_i^j|$, then $x \in C \setminus (B_i^j \cup \{a_i^j\})$ has a negative utility, contradicting $(*)$. If $|C \setminus (B_i^j \cup \{a_i^j\})| > 3$, then $b_i^{j,1}$ has negative utility, contradicting $(*)$. Hence, $|C \setminus (B_i^j \cup \{a_i^j\})| = 3$. Moreover, then $a_i^j \in C$ as an agent in $C \setminus (B_i^j \cup \{a_i^j\})$ would have at most two friends but three enemies, and therefore a negative utility, otherwise. For similar reasons, the agents in $C \setminus (B_i^j \cup \{a_i^j\})$ have to form a clique of friends of a_i^j.

We will exclude all possible agents in $C \setminus (B_i^j \cup \{a_i^j\})$. First note that the structure we obtained so far holds for arbitrary i and j. Hence, if $a_i^{j'} \in C$ for $j' \in [5] \setminus \{j\}$, then the assertion of Claim 14 is already true for i and j' and therefore $B_i^{j'} \in \pi$. But then, $a_i^{j'}$ can perform an MIS deviation to join $B_i^{j'}$, a contradiction. Thus, since the agents in Z are no mutual friends, there exist $l, l' \in \{6, 7, 8, 9\}$ with $a_i^l \in C$ and $a_i^{l'} \notin C$. By $(*)$, $v_{a_i^{l'}}(\pi) \geq 0$. Moreover, since a_i^l and $a_i^{l'}$ have identical friends in $N \setminus \{a_i^l, a_i^{l'}\}$ and $a_i^{l'}$ is also a friend of a_i^l, it holds that $v_{a_i^l}(\pi(a_i^{l'}) \cup \{a_i^l\}) \geq 1$. Since $v_{a_i^l}(\pi) = 0$, this is an NS deviation. Also, since all friends of $a_i^{l'}$ and $a_i^{l'}$ herself favor a_i^l to join their coalition, this is even an MIS deviation. Hence, we obtain a contradiction. ◁

The next claim establishes a relationship between agents in Z and A_i.

Claim 15. *For $i \in [5]$, if $Z \cap \pi(a_i^j) = \emptyset$ for all $j \in [9]$, then $A_i \setminus \{a_i^0\} \in \pi$.*

Proof. Let $i \in [5]$ such that $Z \cap \pi(a_i^j) = \emptyset$ for all $j \in [9]$. First, we show that then $\pi(a_i^j) \subseteq A_i$ for $j = 6, 7, 8, 9$. Let therefore $j \in \{6, 7, 8, 9\}$ and assume for contradiction that $\pi(a_i^j) \setminus A_i \neq \emptyset$. By Claim 12, Claim 14, and the initial assumptions of this claim, $\pi(a_i^j) \subseteq \bigcup_{l \in [5]} A_l$. Consider $x \in \pi(a_i^j) \setminus A_i$. If $|\pi(a_i^j) \setminus A_i| \leq |\pi(a_i^j) \cap A_i|$, then $v_x(\pi) < 0$, contradicting $(*)$. On the other hand, if $|\pi(a_i^j) \setminus A_i| \geq |\pi(a_i^j) \cap A_i|$, then $v_{a_i^j}(\pi) < 0$, also contradicting $(*)$. We derived a contradiction in both cases and can therefore conclude that $\pi(a_i^j) \subseteq A_i$.

As in previous steps, we can use the symmetry of the agents in $\{a_i^j : j = 6, 7, 8, 9\}$ to show that there exists a coalition $C \in \pi$ with $\{a_i^j : j = 6, 7, 8, 9\} \subseteq C \subseteq A_i$. Indeed, otherwise, one of these agents could join the coalition of another such agent of at least as high utility by an MIS deviation. Hence, $B_i^j \cup \{a_i^j\} \notin \pi$ for $j \in [5]$ as a_i^j would perform an MIS deviation to join C, otherwise. But then, similarly as above, $\pi(a_i^j) \subseteq A_i$ for $j \in [5]$, and therefore even $A_i \setminus \{a_i^0\} \subseteq C$. Finally, if $a_0^i \in C$, then $v_{a_i^0} = 9$. However, by Claim 12, $K_i \in \pi$ and therefore a_i^0 could perform an MIS deviation to join K_i. Hence, $C = A_i \setminus \{a_i^0\}$. ◁

We have now collected enough structural results to consider the agents in Z. The next two claims will yield the final contradiction.

Claim 16. *There exists no $C \in \pi$ with $Z \subseteq C$.*

Proof. Assume for contradiction that there exists $C \in \pi$ with $Z \subseteq C$. By Claim 12 and Claim 14, $C \subseteq Z \cup \bigcup_{i\in[5]} A_i$. Define $I = \{i \in [5] \colon A_i \cap C \neq \emptyset\}$ and let

$$i^* \in \arg\min_{i\in I}\{|A_i \cap C|\}. \tag{$**$}$$

Let $x \in A_{i^*} \cap C$.

<u>Case 1</u>: $|I| = 5$.

In this case, we obtain a contradiction to $(*)$ because

$$\begin{aligned} v_x(\pi) &= 3 + (|A_{i^*} \cap C| - 1) - \sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \\ &\overset{(**)}{\leq} 2 - (|I| - 2)|A_{i^*} \cap C| \leq -1 < 0. \end{aligned}$$

<u>Case 2</u>: $|I| = 4$.

As in the previous case, $0 \overset{(*)}{\leq} v_x(\pi) \leq 2 + |A_{i^*} \cap C| - \sum_{i\in I\setminus\{i^*\}} |A_i \cap C|$. Therefore,

$$3|A_{i^*} \cap C| \leq \sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \leq 2 + |A_{i^*} \cap C|.$$

Consequently, $|A_{i^*} \cap C| = 1$ and $|A_i \cap C| = 1$ for $i \in I \setminus \{i^*\}$. Let $l \in [3]$. Then, $v_{z_l}(\pi) \leq 4$. By Claim 15, it holds that $A_{i'} \setminus \{a^0_{i'}\} \in \pi$, where $i' \in [5] \setminus I$. Hence, z_l has an MIS deviation, a contradiction.

<u>Case 3</u>: $|I| = 3$.

As in Case 2,

$$2|A_{i^*} \cap C| \leq \sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \leq 2 + |A_{i^*} \cap C|.$$

Hence, $|A_{i^*} \cap C| \leq 2$ and thus $\sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \leq 4$. Therefore, $v_{z_l}(\pi) \leq 6$ if $l \in [3]$, and an analogous MIS deviation is possible as in the previous case.

<u>Case 4</u>: $|I| = 2$.

Let $i' \in I \setminus \{i^*\}$ be the unique second index in I. We claim that $a^j_i \notin C$ for $i \in I$ and $j \in [5]$. Let $j \in [5]$. First, if $a^j_{i^*} \in C$, then $v_{a^j_{i^*}}(\pi) \leq 3 + (|A_{i^*} \cap C| - 1) - |A_{i'} \cap C| \leq 2$. Moreover, by Claim 14, $B^j_{i^*} \in \pi$ and $a^j_{i^*}$ could perform an MIS deviation to join $B^j_{i^*}$.

Second, assume that $a^j_{i'} \in C$. Then, again by Claim 14, $B^j_{i'} \in \pi$ and since π is majority-in stable, $u_{a^j_{i'}}(\pi) \geq 3$. Let $j' \in [9] \setminus \{j\}$ and assume for contradiction that $a^{j'}_{i'} \notin C$. Since $a^{j'}_{i'}$ has at least as many friends in C as $a^j_{i'}$ (recall that $B^j_{i'} \in \pi$), $v_{a^{j'}_{i'}}(\pi) \geq v_{a^j_{i'}}(\pi) + 1 \geq 4$. Using Claim 14, this means in particular that $B^{j'}_{i'} \cap \pi(a^{j'}_{i'}) = \emptyset$ if $j' \in [5]$. Therefore, $v_{a^{j'}_{i'}}(\pi(a^j_{i'}) \cup \{a^{j'}_{i'}\}) \geq v_{a^j_{i'}}(\pi) + 1$ and

$v_{a^j_{i'}}(\pi(a^{j'}_{i'}) \cup \{a^j_{i'}\}) \geq v_{a^{j'}_{i'}}(\pi) + 1$. Hence, either $a^{j'}_{i'}$ has an MIS deviation to join $\pi(a^j_{i'})$ or $a^j_{i'}$ has an MIS deviation to join $\pi(a^{j'}_{i'})$, a contradiction. Consequently, $a_{i'}(j') \in C$ and therefore $A_{i'} \setminus \{a^0_{i'}\} \subseteq C$.

Recall that we already know that $|A_{i^*} \cap C| \leq 5$ because $a^l_{i^*} \notin C$ for $l \in [5]$. We obtain a contradiction to $(*)$ because

$$v_x(\pi) \leq 3 + (\underbrace{|A_{i^*} \cap C|}_{\leq 5} - 1) - \underbrace{|A_{i'} \cap C|}_{\geq 9} \leq -2 < 0.$$

Hence, we can conclude that $a^j_{i'} \notin C$ for $j \in [5]$. But then, for $l \in [3]$, $v_{z_l} \leq |(A_{i^*} \setminus \{a^0_{i^*}\}) \cap C| + |(A_{i'} \setminus \{a^0_{i'}\}) \cap C| \leq 8$. Hence, z_l can perform an MIS deviation to join $A_i \setminus \{a^0_i\}$ for $i \in [5] \setminus I$, as in the previous two cases.

Case 5: $|I| = 1$.

If $C \neq Z \cup (A_{i^*} \setminus \{a^0_{i^*}\})$, then, for $l \in [3]$, $v_{z_l}(\pi) \leq 8$, and an analogous MIS deviation as in the previous cases is possible. Hence, $C = Z \cup (A_{i^*} \setminus \{a^0_{i^*}\})$. But then $v_{a^0_{i^*}}(\pi) \leq 11$, whereas $v_{a^0_{i^*}}(C \cup \{a^0_{i^*}\}) \geq 12$. Hence, $a^0_{i^*}$ has an MIS deviation to join C (which is favored by all agents in $A_{i^*} \setminus \{a^0_{i^*}\}$). This is a contradiction, and concludes the proof of the claim. $\lhd$

For a final contradiction, it remains to lead the case to a contradiction that the agents in Z are part of different coalitions.

Claim 17. *There exists $C \in \pi$ with $Z \subseteq C$.*

Proof. Assume for contradiction that there exists $C \in \pi$ with $Z \cap C \neq \emptyset$ and $Z \not\subseteq C$.

Assume first that $|Z \cap C| = 2$ and suppose without loss of generality that $z_1, z_2 \in C$. Note that $v_{z_3}(C \cup \{z_3\}) = v_{z_2}(\pi) + 1$. Hence, if $v_{z_3}(\pi) \leq v_{z_2}(\pi)$, then z_3 can perform an NS deviation to join C. This is even an MIS deviation as $v_{z_2}(\pi) \geq 0$ and z_2 favors her to join. On the other hand, $v_{z_2}(\pi(z_3) \cup \{z_2\}) = v_{z_3}(\pi) + 1$. Hence, if $v_{z_2}(\pi) < v_{z_3}(\pi)$, then z_2 has an NS deviation to join $\pi(z_3)$. Note that z_3 is opposed to that. However, as $v_{z_3}(\pi) > v_{z_2}(\pi) \geq 0$, and every friend of z_3 in $\pi(z_3)$ favors to let z_2 join, it holds that

$$\begin{aligned}|F_{\text{in}}(\pi(z_3), z_2)| &= |\{y \in \pi(z_3) \colon u_{z_3}(y) = 1\}| \\ &\geq |\{y \in \pi(z_3) \colon u_{z_3}(y) = -1\}| + 1 \\ &\geq |F_{\text{out}}(\pi(z_3), z_2)|.\end{aligned}$$

Hence, this is even an MIS deviation.

Finally, assume that $\pi(z_l) \cap Z = \{z_l\}$ for all $l \in [3]$. Let $l \in [3]$ and $i \in [5]$. Then, $a^0_i \notin \pi(z_l)$. Indeed, if $a^0_i \in \pi(z_l)$, then u^0_i can have at most 10 friends in her coalition. By Claim 12, $K_i \in \pi$ and a^0_i would perform an MIS deviation to join this coalition. By this observation and using Claim 12 and Claim 14, z_l forms a coalition with friends only (and these do additionally also have all agents in Z as a friend).

Let $l^* \in \arg\min_{l\in[3]}\{v_{z_l}(\pi)\}$. Without loss of generality, we may assume that $l^* = 1$. Then, z_1 has an NS deviation to join $\pi(z_2)$. This is also an MIS deviation unless $\pi(z_2) = \{z_2\}$. Then, z_2 has an NS deviation to join $\pi(z_3)$, which in turn is an MIS deviation unless $\pi(z_3) = \{z_3\}$. By the minimality assumption on l^*, it must then also hold that $\pi(z_1) = \{z_1\}$. But then, using Claim 15, $A_1 \setminus \{a_1^0\} \in \pi$ and z_1 could perform an MIS deviation to join this coalition. This contradiction concludes the proof of the claim. ◁

As the combination of Claim 16 and Claim 17 directly leads to a contradiction, we have shown that the constructed FEG has no MIS partition. □

Similar to Proposition 7, it is easy to see that the singleton partition is majority-in stable in every FEG with complete enemy relation. Indeed, then an agent either has no incentive to join another agent or the other agent will deny her consent. Hence, also majority-in stability guarantees the existence of a stable partition in the run-and-chase example discussed in the introduction, for which no NS partition exists.

Still, we obtain a hardness for general FEGs based on the game constructed in Proposition 8. First, we prove another useful lemma, which excludes that enemies can be in a joint coalition of an MIS partition if they do not have a common friend in their coalition.

Lemma 5. *Consider an FEG (N, v) with an MIS partition π. Let $i, j \in N$ be two agents with $v_i(j) = v_j(i) = -1$ such that, for every agent $k \in \pi(i) \setminus \{i, j\}$, it holds that $v_i(k) = -1$ or $v_j(k) = -1$. Then, $i \notin \pi(j)$.*

Proof. Let an FEG (N, v) be given together with an MIS partition π, let $i, j \in N$ be two agents satisfying the assumptions of the lemma. Assume for contradiction that $i \in \pi(j)$. Our assumptions imply in particular that, for every agent $k \in N \setminus \{i, j\}$, it holds that $v_i(k) + v_j(k) \le 0$. Hence,

$$v_i(\pi) + v_j(\pi) = -2 + \sum_{k\in\pi(j)\setminus\{i,j\}} v_i(k) + v_j(k) \le -2.$$

Therefore $v_i(\pi) < 0$ or $v_j(\pi) < 0$, a contradiction. □

We are ready to prove hardness of deciding on the existence of MIS partitions in FEGs.

Theorem 12. *Deciding whether an FEG contains an MIS partition is NP-complete.*

Proof. We provide another reduction from E3C. Let $(R, \mathcal{S})$ be an instance of E3C. We now define the reduced FEG (N, v), which is illustrated in Figure 13.

Let $N = N_{\mathcal{S}} \cup N_R$ where

- $N_{\mathcal{S}} = \cup_{S\in\mathcal{S}} N_S$ with $N_S = V^S \cup \bigcup_{r\in S} V_r^S$ for $S \in \mathcal{S}$ and
- $N_R = \cup_{r\in R} N_r$.

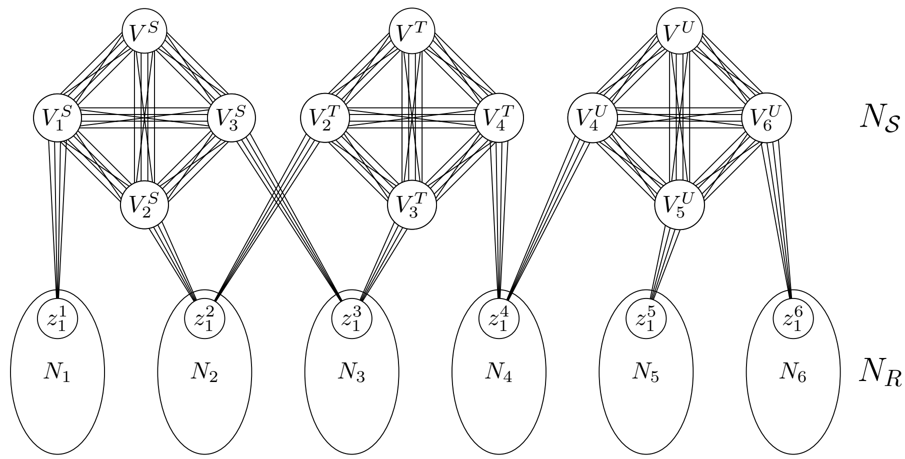

Figure 13: Schematic of the reduction from the proof of Theorem 12. We depict the reduced instance for the instance $(R, \mathcal{S})$ of E3C where $R = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{S} = \{S, T, U\}$ with $S = \{1, 2, 3\}$, $T = \{2, 3, 4\}$, and $U = \{4, 5, 6\}$. In contrast to earlier reductions, $N_{\mathcal{S}}$ now contains multiple copies for each of the agents in previous reductions. The edge bundles indicate mutual friendships of all involved agents. Every element in R is represented by a gadget identical to the game in Proposition 8, which is indicated by the ellipses at the bottom. Figure 12 provides an illustration of this gadget.

We define, for $S \in \mathcal{S}$, $V^S = \{c_i^S : i \in [10]\}$, and for $S \in \mathcal{S}$ and $r \in S$, $V_r^S = \{c_{r,i}^S : i \in [10]\}$. To define the sets N_r, assume that (N', v') is the FEG constructed in the proof of Proposition 8. Then, for $r \in R$, we define $N_r = \{x^r : x \in N'\}$. Specifically, we denote the agent corresponding to z_1 by z_1^r. Agents of this type will be linked to agents in V_r^S by means of a positive utility correspondence. We define utilities v as follows:

- For all $S \in \mathcal{S}$, $x, y \in N_S$: $v_x(y) = 1$.
- For all $S \in \mathcal{S}$, $r \in S$, and $x \in V_r^S$: $v_x(z_1^r) = v_{z_1^r}(x) = 1$.
- For all $r \in R$ and $x, y \in N'$: $v_{x^r}(y^r) = v'_x(y)$ i.e., the internal valuations for agents in N_r are identical to the valuations in the counterexample defined in the proof of Proposition 8.
- All other valuations are -1.

We claim that $(R, \mathcal{S})$ is a Yes-instance if and only if the reduced FEG contains an MIS partition.

$\Longrightarrow$ Suppose first that $\mathcal{S}' \subseteq \mathcal{S}$ partitions R. We define a partition π based on a partition π' of the agent set $N' \setminus \{z_1\}$ in the game (N', v') from the proof of Proposition 8. The partition π' is given as follows.

- We have $\{z_2, z_3\} \cup A_1 \in \pi'$ and $K_1 \in \pi'$.
- For $i, j \in [5]$, we have $B_i^j \in \pi'$.

- For $i \in \{2,3,4,5\}$, we have $A_i \setminus \{a_i^0\} \in \pi'$ and $K_i \cup \{a_i^0\} \in \pi'$.

Based on this partition, we can define the partition π as follows.

- For $S \in \mathcal{S} \setminus \mathcal{S}'$, we have $N_S \in \pi$ and for $S \in \mathcal{S}'$, we have $V^S \in \pi$.
- For $S \in \mathcal{S}'$, $r \in S$, we have $V_r^S \cup \{z_1^r\} \in \pi$.
- For $r \in R$ and $x \in N' \setminus \{z_1\}$, we have $\pi(x^r) = \{y^r : y \in \pi'(x)\}$.

Showing that π is majority-in stable follows from a lengthy, but straightforward case analysis.

- For every $r \in R$ and $x \in N' \setminus \{z_1\}$, agent x^r has utility $v_{x^r}(\pi) > 0$, and therefore x^r cannot join a coalition containing an agent outside N_r as this would give her negative utility. Moreover, also deviations within N_r cannot improve her utility:
  - For $i, j \in [5]$, and $l \in [3]$, if $x = b_i^{j,l}$, then $v_{x^r}(\pi) = 2$, but x^r can have at most one friend in any other coalition.
  - For $i \in [5]$ and $j \in [11]$, if $x = k_i^j$, then $v_{x^r}(\pi) \geq 10$, but x^r can have at most one friend in any other coalition.
  - If $x = a_1^0$, then $v_{x^r}(\pi) = 11$, and the only possible deviation that gives x^r positive utility, i.e., joining K_1, would not increase her utility.
  - For $i \in \{2,3,4,5\}$, if $x = a_i^0$, then $v_{x^r}(\pi) = 11$, and the only possible deviation that gives x^r positive utility, i.e., joining $A_i \setminus \{a_i^0\}$ would decrease her utility.
  - If $x = z_2$ or $x = z_3$, then $v_{x^r}(\pi) \geq 9$, and the only possible deviations, i.e., joining a coalition $A_i \setminus \{a_i^0\}$ for $i \in \{2,3,4,5\}$ would not increase her utility.
- For $r \in R$, $v_{z_1^r}(\pi) = 10$, and joining any other coalition does not increase her utility.
- For $S \in \mathcal{S} \setminus \mathcal{S}'$ and $x \in N_S$, $v_x(\pi) = 39$, and joining any other coalition does not give agent x positive utility.
- For $S \in \mathcal{S}'$ and $x \in V^S$, $v_x(\pi) = 9$, and joining any other coalition does not give her a better utility. In particular, joining $V_r^S \cup \{z_1^r\}$ for $r \in S$ would also give her a utility of 9.
- For $S \in \mathcal{S}'$, $r \in S$, and $x \in V_r^S$, $v_x(\pi) = 10$, and no other coalition gives her a better utility. In particular, joining V^S would also give her a utility of 10.

Together, we have shown that π is an MIS partition (we have even shown that it is an NS partition).

$\Longleftarrow$ Conversely, assume that the reduced FEG contains an MIS partition π.

Note that the assumptions of Lemma 5 are in particular satisfied for two agents $i, j \in N$ with $v_i(j) = v_j(i) = -1$ such that, for *every* agent $k \in N \setminus \{i, j\}$, it holds that $v_i(k) = -1$ or $v_j(k) = -1$. Therefore, we can apply Lemma 5 multiple times to obtain the following facts:

1. For $r, r' \in R$ with $r \neq r'$, $x \in N_r$, and $y \in N^{r'}$, it holds that $y \notin \pi(x)$.

2. For every $S, S' \in \mathcal{S}$, $S \neq S'$, $x \in V^S$, and $y \in N_{S'}$, it holds that $y \notin \pi(x)$.

3. For every $S \in \mathcal{S}$, $r \in R \setminus S$, $x \in N_S$, and $y \in N_r$, it holds that $y \notin \pi(x)$.

4. For every $S \in \mathcal{S}$, $r \in S$, and $x \in V^S$, it holds that $\pi(x) \cap N_r \subseteq \{z_1^r\}$.

Next, we can apply Lemma 4 to obtain the next two facts.

5. For every $S \in \mathcal{S}$, there exists a coalition $C \in \pi$ with $V^S \subseteq C$.

6. For every $S \in \mathcal{S}$, $r \in \mathcal{S}$, there exists a coalition $C \in \pi$ with $V_r^S \subseteq C$.

Moreover, combining Lemma 5 with Fact 6 allows us to further refine Fact 4 yielding the fact

7. For every $S \in \mathcal{S}$, $r \in S$, and $x \in V^S$, it holds that $V_r^S \subseteq \pi(x)$ whenever $z_1^r \in \pi(x)$.

We are ready to restrict the coalitions of agents in sets V^S to two possibilities.

Claim 18. *For all $S \in \mathcal{S}$, it holds that $V^S \in \pi$ or $N_S \in \pi$.*

Proof. Let $S \in \mathcal{S}$ and $x \in V^S$, and define $C = \pi(x)$. By Fact 5, $V^S \subseteq C$. Furthermore, by Fact 2, Fact 3, and Fact 4, it holds that $C \subseteq N_S \cup \{z_1^r : r \in S\}$.

Suppose that $V^S \subsetneq C$. We have to show that $C = N_S$. By Fact 7, there exists $r \in S$ with $V_r^S \subseteq C$. Assume for contradiction that $z_1^r \in C$. Since all agents in C except the agents in N_S^r are enemies of z_1^r, it holds that $v_{z_1^r}(\pi) < 0$ if $C \supsetneq V^S \cup V_r^S \cup \{z_1^r\}$. This would contradict that π is an MIS partition and therefore $C = V^S \cup V_r^S \cup \{z_1^r\}$. In particular, every agent $y \in N_S \setminus C$ has to satisfy $v_y(\pi) \geq 19$. Otherwise, this agent could perform an MIS deviation to join C. Hence, there exists a coalition $D \in \pi$ with $N_S \setminus C \subseteq D$. Assume that $S = \{r, r', r''\}$. Let $y' \in V_{r'}^S$ and $y'' \in V_{r''}^S$. If there exists an agent $q \in N \setminus (V_{r'}^S \cup V_{r''}^S)$, then either $v_{y'}(q) = -1$ or $v_{y''}(q) = -1$. Assume without loss of generality that the former case holds. Then, $z_1^{r'} \in D$. Otherwise, $v_{y'}(\pi) \leq 18$ and y' would deviate to join C. But then also $z_1^{r''} \in D$ (due to the utility of y''), and it must hold that $D = V_{r'}^S \cup V_{r''}^S \cup \{z_1^{r'}, z_1^{r''}\}$. But then, $v_{z_1^{r'}}(\pi) = -1$, a contradiction. Hence, $D = V_{r'}^S \cup V_{r''}^S$. but then, any agent in V^S has an MIS deviation to join D, a contradiction. We can conclude that $z_1^r \notin C$.

Since the previous argument is valid for every $r \in S$ with $V_r^S \subseteq C$, we can conclude that $C \subseteq N_S$. Assume for contradiction that there exists an agent $y \in N_S \setminus C$, say without loss of generality that $y \in V_{r'}^S$. Note that $v_y(C \cup \{y\}) \geq 20$, and therefore, it must hold that $v_y(\pi) \geq 20$. Hence, $V_{r'}^S \cup V_{r''}^S \cup \{z_1^{r'}\} \subseteq \pi(y)$. Therefore, even $z_1^{r''} \in \pi(y)$ because otherwise, an agent in $V_{r''}^S$ would perform an MIS deviation to join C. But then, as in the previous argument, $z_1^{r'}$ has a negative utility, a contradiction. Hence, $C = N_S$. This concludes the proof of the claim. $\lhd$

Our next goal is to pinpoint the coalitions of agents in sets of the type V_r^S.

Claim 19. *For all $S \in \mathcal{S}$ and $r \in S$, it holds that $V_r^S \cup \{z_1^r\} \in \pi$ or $N_S \in \pi$.*

Proof. For $S \in \mathcal{S}$ and $r \in S$ consider an agent $x \in V_r^S$ and define $C = \pi(x)$. Assume that $C \neq N_S$. We have to show that $C = V_r^S \cup \{z_1^r\}$. By Claim 18, we know then that $V^S \in \pi$. By Fact 3, we know that $C \subseteq N_{\mathcal{S}} \cup \bigcup_{t \in S} N^t$. Assume that $S = \{r, r', r''\}$.

Assume for contradiction that there exists an agent $y \in (V_{r'}^S \cup V_{r''}^S) \cap C$. Then, $C \cap N^t \subseteq \{z_1^t\}$ for $t \in S$. Indeed, if there is $t \in S$ and an agent $q \in (N^t \setminus \{z_1^t\}) \cap C$, then we derive a contradiction by applying Lemma 5 for q and one of x and y. A similar argument shows that $N_{\mathcal{S}} \cap C \subseteq N_S$. Hence, $C \subseteq N_S \cup \bigcup_{t \in S}\{z_1^t\}$.

By Fact 6 and our assumptions, we know that in addition $V_r^S \cup V_t^S \subseteq C$ for $t \in S$ with $y \in V_t^S$. Hence, $v_p(C \cup \{p\}) \geq 17 > 9 = v_p(\pi)$ for every $p \in V^S$. Hence, such an agent p could perform an MIS deviation, a contradiction. We can, therefore, conclude that $C \cap N_S = V_r^S$. Since $V^S \in \pi$, it must hold that $v_x(\pi) \geq 10$. Since we already know that $C \subseteq N_S \cup (N_{\mathcal{S}} \setminus N_S) \cup \bigcup_{t \in S} N^t$, this is only possible if $C = V_r^S \cup \{z_1^r\}$. $\lhd$

We are ready to prove that $(R, \mathcal{S})$ is a Yes-instance. Define $\mathcal{S}' = \{S \in \mathcal{S} \colon N_S \notin \pi\}$. First, note that the sets in $\mathcal{S}'$ are disjoint. Indeed, let $S \in \mathcal{S}'$ and consider $r \in S$. By Claim 19, $V_r^S \cup \{z_1^r\} \in \pi$. Hence, for every $S' \in \mathcal{S} \setminus \{S\}$ with $r \in S'$, it cannot be the case that $V_{S'}^r \cup \{z_1^r\} \in \pi$. Hence, another application of Claim 19 yields $N_{S'} \in \pi$, and therefore $S' \notin \mathcal{S}'$.

It remains to show that $\mathcal{S}'$ covers all elements in R. Therefore, let $r \in R$. By Fact 1, Claim 18, and Claim 19, it holds that $\pi(x) \subseteq N_r$ for all $x \in N_r \setminus \{z_1^r\}$ and $\pi(z_1^r) \subseteq N_r$ or $\pi(z_1^r) = V_r^S \cup \{z_1^r\}$ for some $S \in \mathcal{S}$. In the former case, $\pi(x) \subseteq N_r$ for all $x \in N_r$, which contradicts the fact that π is an MIS partition because, according to the proof of Proposition 8, the game restricted to N_r contains no MIS partition. Hence, the latter case must be true, i.e., $\pi(z_1^r) = V_r^S \cup \{z_1^r\}$ for some $S \in \mathcal{S}$. Then, $S \in \mathcal{S}'$, and therefore r is covered by an element in $\mathcal{S}'$. $\square$

5.4. Joint-Majority and Separate-Majorities Stability

The computational boundaries encountered so far only hold for one-sided stability notions where only the welcoming or the abandoned coalition takes a vote. In addition, Theorem 6 shows that these are opposed by tractabilities under two-sided majority consent.

However, for general utilities, the existence of SMS (and therefore JMS) partitions is not guaranteed anymore.

Proposition 9. *There exists an ASHG without SMS partition.*

Proof. Let $N = [5]$ and consider the utilities according to Table 1 below.

See Figure 14 for a graphical representation of this example. We show that no partition can be separate-majorities stable by an exhaustive case analysis.

Table 1: 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

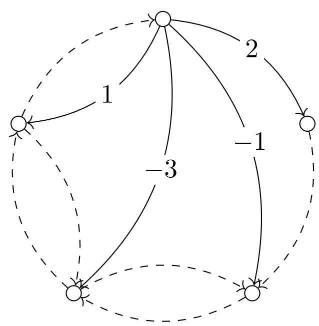

Figure 14: The ASHG without SMS partition from Proposition 9. Outgoing edges with weights have been drawn explicitly only for one agent, they are the same for each agent (up to rotation).

Let $+_{[5]}$ denote addition modulo 5, mapping to the representative in the set [5]. Assume for contradiction that π is separate-majorities stable, 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.

□

We can leverage the game constructed in the proof of Proposition 9 to oppose Theorem 6 with a hardness result in general ASHGs.

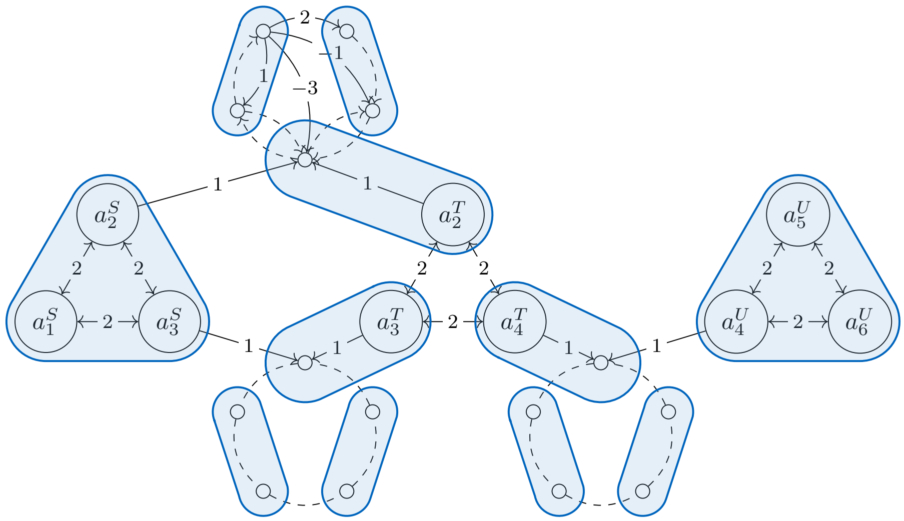

Figure 15: Schematic of the reduction from the proof of Theorem 13 for the Yes-instance of E3C $(\{1,\ldots,6\},\{s,t,u\})$ with $S=\{1,2,3\}, t=\{2,3,4\}$ and $u=\{4,5,6\}$. Some edges have been omitted for clarity. The indicated partition is both SMS and JMS.

Theorem 13. *Deciding whether an ASHG contains an SMS (or JMS) partition is* NP*-complete.*

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

- $A^S=\left\{a_r^S\colon r\in S\right\}$ for $S\in\mathcal{S}$ and
- $B_i^r=\left\{b_{i,j}^r\colon j\in[5]\right\}$ for $r\in R, i\in[n_r-1]$.

Also, define utilities v as follows:

- For each $S\in\mathcal{S}, a\neq a'\in A^S: v_a(a')=2$.
- For each $r\in R, S\in\mathcal{S}_r, i\in[n_r-1]: v_{a_r^S}(b_{i,1}^r)=1, v_{b_{i,1}^r}(a_r^S)=0$.
- For each $r\in R, i\in[n_r-1]$, each B_i^r has internal utilities as in the example constructed in Proposition 9, i.e., if v' are the utilities in the example, then $v_{b_{i,j}^r}(b_{i,k}^r)=v_j'(k)$ for all $j,k\in[5]$.
- All other utilities are $-M$, where $M=|\mathcal{S}|+5$ (can be thought of as $-\infty$).

The reduction is visualized in Figure 15. Note that it can be performed in polynomial time, as there are at most $3|\mathcal{S}|+5|R||\mathcal{S}|$ agents. We proceed with the proof of the correctness of the reduction and show that if $(R, \mathcal{S})$ 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 $(R, \mathcal{S})$.

$\implies$ Suppose $(R, \mathcal{S})$ has an exact cover $\mathcal{S}' \subseteq \mathcal{S}$. We construct a stable partition π.

- First, we create coalitions corresponding to the cover, i.e., for each $S \in \mathcal{S}$, we have $A^S \in \pi$ if and only if $S \in \mathcal{S}'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $S \in \mathcal{S}_r$ 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 joint-majority stable and separate-majorities stable. To see this, note that the only agents that have an 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 $b_{i,1}^r$ and consider $S \in \mathcal{S}$ with $a_r^S \in \pi(b_{i,1}^r)$. Then, a_r^S is against $b_{i,1}^r$ leaving, so the partition is majority-out stable 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.

$\impliedby$ Conversely, assume that there is a partition π that is joint-majority stable or separate-majorities stable. We show that then there must be an exact cover $\mathcal{S}' \subseteq \mathcal{S}$ of R. We begin with some observations:

1. Agents $b_{i,j}^r$ with $j \in \{2, \ldots, 5\}$ must have $\pi(b_{i,j}^r) \subseteq B_i^r$. For contradiction, suppose this is not so. Consider first the case that there is exactly one outside agent $a \in \pi(b_{i,j}^r) \backslash B_i^r$. Then, as $v_a(b_{i,j}^r) = -M$, a has an incentive to form a singleton coalition, and this is a valid SMS deviation (and therefore JMS deviation). The other case is that there are at least two agents $a, a' \in \pi(b_{i,j}^r) \setminus B_i^r$ with $a \neq a'$. Then, as $v_{b_{i,j}^r}(a) = -M$ and $|F_{\text{out}}(\pi(b_{i,j}^r), b_{i,j}^r)| \geq |\{a, a'\}| = 2 = \left|\left\{b_{i,j+_{[5]}1}^r, b_{i,j+_{[5]}4}^r\right\}\right| \geq |F_{\text{in}}(\pi(b_{i,j}^r), b_{i,j}^r)|$, $b_{i,j}^r$ can form a singleton coalition.
2. Agents a_r^S and $a_{r'}^{S'}$ with $S \neq S'$ have $\pi(a_r^S) \neq \pi(a_{r'}^{S'})$. For contradiction, suppose the contrary, i.e., suppose that there are a_r^S and $a_{r'}^{S'}$ with $S \neq S'$, but $\pi(a_r^S) = \pi(a_{r'}^{S'})$. Define $C = \pi(a_r^S)$. As $v_{a_r^S}(a_{r'}^{S'}) = v_{a_{r'}^{S'}}(a_r^S) = -M$, both would rather be in a singleton coalition. Further, we can assume without loss of generality that $|A^S \cap C| \leq |A^{S'} \cap C|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a_r^S)| \geq |A^{S'} \cap C| \geq |A^S \cap C| > |F_{\text{in}}(C, a_r^S)|$, it holds that a_r^S can deviate to form a singleton coalition.
3. Agents $b_{i,1}^r$ must be in a pair with exactly one agent a_r^S. Fix such an agent $b_{i,1}^r$. First, due to the first observation, she cannot be alone, and no

Table 2: Overview of our computational results. A red cell denotes the existence of games without a 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 results from this paper, even by executing a dynamics.
[a]: Aziz and Brandl (2012), [b]: Aziz et al. (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. 4)	FP[a] (Th. 4)	FP[c] (Th. 4)
CNS	NP-c (Th. 3)	FP (Th. 5)	FP (Th. 5)	FP (Th. 5)
CIS	FP[b]	FP[b]	FP[b]	FP[b]
MIS	NP-c (Th. 7)	NP-c (Th. 12)	NP-c (Th. 7)	FP (Th. 9)
MOS	NP-c (Th. 8)	NP-c (Th. 11)	NP-c (Th. 8)	NP-c (Th. 10)
JMS	NP-c (Th. 13)	FP (Th. 6)	FP (Th. 6)	FP (Th. 6)
SMS	NP-c (Th. 13)	FP (Th. 6)	FP (Th. 6)	FP (Th. 6)

other agents from B_i^r can be in her coalition, as the example constructed in Proposition 9 has no SMS partition. Consequently, she must form a coalition with at least one agent outside of B_i^r, and no agents from B_i^r. Next, due to the second observation, she can be together with at most one agent of type a_r^S. If there was another member from A^S (other than a_r^S), $b_{i,1}^r$ could deviate to a singleton coalition.

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 $b_{i,1}^r$. This leaves exactly one agent a_r^S not in a pair. We claim that for these agents we have $\pi(a_r^S) = A^S$, yielding a cover $\mathcal{S}' = \{S \in \mathcal{S} \colon A^S \in \pi\}$.

Suppose that a_r^S is such an agent not in a pair. Then, $\pi(a_r^S) \subseteq A^S$. If the other two agents from A^S form a pair, then a_r^S has an incentive to join them. Otherwise, the other two agents would have an incentive to join a_r^S. In any case, the only stable situation is $\pi(a_r^S) = A^S$. □

6. Conclusion and Discussion

We studied stability based on single-agent deviations in additively separable hedonic games. Our results complete the complexity picture of eight stability notions in additively separable hedonic games under four utility assumptions. Table 2 summarizes our results and compares them with related results from the literature. All stability notions we consider are based on deviations by single agents, where deviations might be subject to the consent of the abandoned or welcoming coalition of a deviation. The consent can either be unanimous or by a majority vote. Apart from unrestricted utility functions, we consider three restrictions that can be naturally interpreted in terms of friends and enemies.

Our work identifies several interesting complexity dichotomies. First, Nash stability yields computational boundaries that cannot be crossed for severe utility restrictions. On the other hand, unanimous consent leads to positive results for all three utility restrictions based on friends and enemies. 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 mostly negative results if only one of these coalitions is considered.

Notably, we obtain all of our positive results through the convergence of simple and natural dynamics. The convergence of dynamics is a strong certificate of existence because it does not only answer which outcomes are desirable but it also helps to understand how and why certain outcomes form. It would be interesting to see whether dynamics based on stability lead to reasonable predictions of the outcomes of real coalition formation processes.

Moreover, our results obtained by dynamics extend previously known results about individual stability. 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 et al. (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 the convergence of dynamics is the Deviation Lemma, a general combinatorial insight that allows us to study dynamics from a global perspective.

By contrast, we have determined strong boundaries to the efficient computability of stable partitions. First, we resolve the computational complexity of computing CNS partitions, which considers the last open unanimity-based stability notion in unrestricted ASHGs. Second, our intractability concerning AFGs for majority-out stability is a contrast to the positive results for all other consent-based stability notions, and can also be circumvented by considering AFGs with a sparse friendship relation. Finally, we provide sophisticated hardness proofs for majority-based stability concepts in FEGs. However, these turn into computational feasibilities when transitioning to unanimity-based stability, or under further assumptions to the structure of the friendship graph.

A key step of all hardness results in restricted classes of ASHGs was to construct the first No-instances, that is, games that do not admit stable partitions for the respective stability notion. This is no trivial task as can be seen from the complexity of the constructed games. Once No-instances are found, we can leverage them as gadgets of hardness reductions, which is a typical approach for complexity results about hedonic games. We have provided both reductions where the explicit structure of the determined No-instances is used as well as reductions where the mere existence of No-instances is sufficient and can be used as a black box.

Our results give a complete picture of the computational complexity for all considered stability notions and game classes. Still, majority-based stability notions deserve further attention because they offer a natural degree of consent to perform deviations. Majority-based decisions can likely explain agreements to form a coalition where the strong demands of unanimous consent fail. Their thorough investigation in other classes of hedonic games might lead to interesting discoveries. Another intriguing direction is to establish further applications of the combinatorial insights from the Deviation Lemma.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/11-2 and BR 2312/12-1, and by the AI Programme of The Alan Turing Institute. We are grateful for the helpful comments by the anonymous reviewers.

References

Abeledo, H., Rothblum, U.G., 1995. Paths to marriage stability. Discrete Applied Mathematics 63, 1–12.

Ahmadi, S., Awasthi, P., Khuller, S., Kleindessner, M., Morgenstern, J., Sukprasert, P., Vakilian, A., 2022. Individual preference stability for clustering, in: Proceedings of the 39th International Conference on Machine Learning (ICML), pp. 197–246.

Alcalde, J., Revilla, P., 2004. Researching with whom? Stability and manipulation. Journal of Mathematical Economics 40, 869–887.

Aziz, H., 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), pp. 763–770.

Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., Peters, D., 2019. Fractional hedonic games. ACM Transactions on Economics and Computation 7, 1–29.

Aziz, H., Brandt, F., Seedig, H.G., 2013. Computing desirable partitions in additively separable hedonic games. Artificial Intelligence 195, 316–334.

Aziz, H., Savani, R., 2016. Hedonic games, in: Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D. (Eds.), Handbook of Computational Social Choice. Cambridge University Press. chapter 15.

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

Banerjee, S., Konishi, H., 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., Moscardelli, L., 2018. Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. Journal of Artificial Intelligence Research 62, 315–371.

Boehmer, N., Bullinger, M., Kerkmann, A.M., 2023. Causes of stability in dynamic coalition formation, in: Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI), pp. 5499–5506.

Bogomolnaia, A., Jackson, M.O., 2002. The stability of hedonic coalition structures. Games and Economic Behavior 38, 201–230.

Brandt, F., Bullinger, M., 2022. Finding and recognizing popular coalition structures. Journal of Artificial Intelligence Research 74, 569–626.

Brandt, F., Bullinger, M., Tappe, L., 2022. Single-agent dynamics in additively separable hedonic games, in: Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI), pp. 4867–4874.

Brandt, F., Bullinger, M., Wilczynski, A., 2023. Reaching individually stable coalition structures. ACM Transactions on Economics and Computation 11, 4:1–65.

Brandt, F., Wilczynski, A., 2019. On the convergence of swap dynamics to Pareto-optimal matchings, in: Proceedings of the 15th International Conference on Web and Internet Economics (WINE), Springer-Verlag. pp. 100–113.

Bullinger, M., 2020. Pareto-optimality in cardinal hedonic games, in: Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 213–221.

Bullinger, M., 2022. Boundaries to single-agent stability in additively separable hedonic games, in: Proceedings of the 47th International Symposium on Mathematical Foundations of Computer Science (MFCS), pp. 26:1–26:15.

Bullinger, M., 2023. Computing Desirable Outcomes in Coalition Formation. Ph.D. thesis. Technische Universität München.

Bullinger, M., Kober, S., 2021. Loyalty in cardinal hedonic games, in: Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), pp. 66–72.

Bullinger, M., Romen, R., 2023. Online coalition formation under random arrival or coalition dissolution, in: Proceedings of the 31st Annual European Symposium on Algorithms (ESA), pp. 27:1–27:18.

Bullinger, M., Romen, R., 2024. Stability in online coalition formation, in: Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI), pp. 9537–9545.

Bullinger, M., Suksompong, W., 2024. Topological distance games. Theoretical Computer Science 981, 114238.

Carosi, R., Monaco, G., 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), pp. 574–582.

Cechlárová, K., 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. AI Access. chapter 6.

Dimitrov, D., Borm, P., Hendrickx, R., Sung, S.C., 2006. Simple priorities and core stability in hedonic games. Social Choice and Welfare 26, 421–433.

Dimitrov, D., Sung, S.C., 2007. On top responsiveness and strict core stability. Journal of Mathematical Economics 43, 130–134.

Drèze, J.H., Greenberg, J., 1980. Hedonic coalitions: Optimality and stability. Econometrica 48, 987–1003.

Elkind, E., Fanelli, A., Flammini, M., 2020. Price of pareto optimality in hedonic games. Artificial Intelligence 288, 103357.

Elkind, E., Wooldridge, M., 2009. Hedonic coalition nets, in: Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 417–424.

Fanelli, A., Monaco, G., Moscardelli, L., 2021. Relaxed core stability in fractional hedonic games, in: Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), pp. 182–188.

Feldman, M., Lewin-Eytan, L., Naor, J., 2015. Hedonic clustering games. ACM Transactions on Parallel Computing (TOPC) 2, 1–48.

Flammini, M., Monaco, G., Moscardelli, L., Shalom, M., Zaks, S., 2021. On the online coalition structure generation problem. Journal of Artificial Intelligence Research 72, 1215–1250.

Gairing, M., Savani, R., 2019. Computing stable outcomes in symmetric additively separable hedonic games. Mathematics of Operations Research 44, 1101–1121.

Gärdenfors, P., 1975. Match making: Assignments based on bilateral preferences. Behavioral Science 20, 166–173.

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

Kahn, A.B., 1962. Topological sorting of large networks. Communications of the ACM 5, 558–562.

Karp, R.M., 1972. Reducibility among combinatorial problems, in: Miller, R.E., Thatcher, J.W. (Eds.), Complexity of Computer Computations. Plenum Press, pp. 85–103.

Meade, J.E., 1972. The theory of labour-managed firms and of profit sharing. The Economic Journal 82, 402–428.

Michel, L., Scott, A., 2024. Superpolynomial smoothed complexity of 3-FLIP in Local Max-Cut. Technical Report. https://arxiv.org/abs/2310.19594.

Monien, B., 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), Springer-Verlag. pp. 264–275.

Olsen, M., 2012. On defining and computing communities, in: Proceedings of the 18th Computing: The Australasian Theory Symposium (CATS), pp. 97–102.

Roth, A.E., Vande Vate, J.H., 1990. Random paths to stability in two-sided matching. Econometrica 58, 1475–1480.

Saad, W., Han, Z., Basar, T., Debbah, M., Hjorungnes, A., 2011. Hedonic coalition formation for distributed task allocation among wireless agents. IEEE Transactions on Mobile Computing 10, 1327–1344.

Schäffer, A.A., Yannakakis, M., 1991. Simple local search problems that are hard to solve. SIAM Journal on Computing 20, 56–87.

Suksompong, W., 2015. Individual and group stability in neutral restrictions of hedonic games. Mathematical Social Sciences 78, 1–5.

Sung, S.C., Dimitrov, D., 2007. On myopic stability concepts for hedonic games. Theory and Decision 62, 31–45.

Sung, S.C., Dimitrov, D., 2010. Computational complexity in additive hedonic games. European Journal of Operational Research 203, 635–639.