Boundaries to Single-Agent Stability in Additively Separable Hedonic Games

Martin Bullinger ✉
Technical University of Munich, Germany

Abstract

Coalition formation considers the question of how to partition a set of agents into coalitions with respect to their preferences. Additively separable hedonic games (ASHGs) are a dominant model where cardinal single-agent values are aggregated into preferences by taking sums. Output partitions are typically measured by means of stability, and we follow this approach by considering stability based on single-agent movements (to join other coalitions), where a coalition is defined as stable if there exists no beneficial single-agent deviation. Permissible deviations should always lead to an improvement for the deviator, but they may also be constrained by demanding the consent of agents involved in the deviations, i.e., by agents in the abandoned or welcoming coalition. Most of the existing research focuses on the unanimous consent of one or both of these coalitions, but more recent research relaxes this to majority-based consent. Our contribution is twofold. First, we settle the computational complexity of the existence of contractually Nash-stable partitions, where deviations are constrained by the unanimous consent of the abandoned coalition. This resolves the complexity of the last classical stability notion for ASHGs. Second, we identify clear boundaries to the tractability of stable partitions under majority-based stability concepts by proving elaborate hardness results for restricted classes of ASHGs. Slight further restrictions lead to positive results.

2012 ACM Subject Classification Computing methodologies → Multi-agent systems; Theory of computation → Design and analysis of algorithms

Keywords and phrases Coalition Formation, Hedonic Games, Stability

Funding *Martin Bullinger*: Deutsche Forschungsgemeinschaft, grants BR 2312/11-2, BR 2312/12-1.

Acknowledgements I would like to thank Felix Brandt and Leo Tappe for the helpful discussions.

1 Introduction

Coalition formation is a vibrant topic in multi-agent systems at the intersection of theoretical computer science and economic theory. Given a set of agents, e.g., humans or machines, the central concern is to determine a coalition structure, or partition, of the agents into subsets, or so-called coalitions. Agents have preferences over coalition structures, and therefore coalition formation naturally generalizes the matching problem under preferences [22]. As in the special case of matchings, a common assumption is that externalities outside one's own coalition play no role, i.e., agents are only concerned about the coalition they are part of. This assumption leads to the popular framework of hedonic games [18].

In contrast to matchings, the number of coalitions an agent can be part of is not polynomially bounded in coalition formation, and therefore, a lot of effort has been put into identifying reasonable and succinct classes of hedonic games (see, e.g., [2, 5, 8, 20]). In many such classes, agents extract cardinal preferences from a weighted and possibly directed graph by some aggregation method. Probably the most natural and thoroughly researched 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 concept of additively separable hedonic games (ASHGs) [8]. This paper continues to investigate this class of hedonic games.

The desirability of an output, i.e., of a coalition structure, is frequently measured with respect to stability, which captures the prospect of agents maintaining their coalitions. A coalition structure is stable if no single agent or group of agents has an incentive to deviate by leaving their coalitions and joining other coalitions or forming new coalitions. Depending on the requirements that deviators need to meet, one can define various specific stability notions. In this paper, we focus on stability based on single-agent deviations. This means that a deviation consists of a single agent that abandons her current coalition to join another existing coalition or to form a new coalition of her own.

In this case, a reasonable minimum requirement is that a deviating agent should improve her coalition. If no such deviation is possible, then a coalition structure is said to be Nash-stable. However, this leads to an immensely strong stability concept because the deviation is only constrained weakly. As a consequence, Nash-stable outcomes hardly ever exist. For instance, consider a game with two agents x and y where x prefers to form a coalition with y over staying alone, whereas y prefers to stay alone. Then, x always has an incentive to join y whenever she is in a coalition of her own, whereas y would always leave x. Such run-and-chase situations occur in most classes of hedonic games.[1]

Therefore, various weakenings of Nash stability have been proposed. These restrict the possible deviations by adding further requirements on other agents involved in the deviation. Typically, two types of constraints are considered, namely the demanding of some kind of consent from the abandoned or the welcoming coalition. Most of the research has focused on the unanimous consent of these coalitions. This leads to the concepts of contractual Nash stability and individual stability where all agents in the abandoned or welcoming coalition have to approve the deviation. Still, unanimous consent of involved coalitions is a strong requirement. Hence, a reasonable compromise is to merely demand partial consent. Therefore, we also study stability where deviations are constrained by the approval of a majority vote of the abandoned or welcoming coalition.

1.1 Contribution

Our contribution is twofold. First, we settle the complexity of the existence problem of contractually Nash-stable coalition structures. Despite knowing for quite long that No-instances, i.e., additively separable hedonic games which do not admit a contractually Nash-stable coalition structure, exist [28], detailed computational investigations of single-agent stability during the last decade have left this problem open [10, 29]. Hence, we complete the picture of the complexity of unanimity-based single-agent stability concepts in ASHGs.

Second, we investigate majority-based stability concepts. We will show that, even under significant weight restrictions, stable coalition structures need not exist and we can leverage No-instances to obtain computational intractabilities. This complements very recent results by Brandt et al. [10] and resolves problems left open by this work. In particular, we completely pinpoint the complexity of majority-based stability notions in friends-and-enemies games and appreciation-of-friends games.

These results are in line with the repeatedly observed theme in hedonic games research that the existence of counterexamples is the key to computational intractabilities (see, e.g., [3, 10, 11, 16, 29]).[2] On the other hand, we demonstrate that the observed intractabilities lie at the computational boundary by carving out further weak restrictions that lead to the existence and efficient computability of stable states.

---

[1] Notably, Nash-stable coalition structures always exist in ASHGs if the input graph is symmetric [8], and in a generalization of this class of games called subset-neutral hedonic games [27].

[2] A notable exception is provided by Bullinger and Kober [13] who identify a class of hedonic games where partitions in the core always exist, but are still hard to compute.

1.2 Related Work

The study of hedonic games was initiated by Drèze and Greenberg [18] but was only popularized two decades later by Banerjee et al. [6], Cechlárová and Romero-Medina [15], and Bogomolnaia and Jackson [8]. Aziz and Savani [4] review many important concepts in their survey. 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 [5, 20] are ordinal and fully expressive, i.e., they can, in principle, express every preference relation over coalitions. Still, representing certain preference relations requires exponential space. These representations are contrasted by cardinal representations based on weighted graphs [2, 8, 26], which are not fully expressive but only require polynomial space (except when weights are artificially large). Apart from the already discussed additively separable hedonic games, important aggregation methods consider the average of weights leading to the classes fractional hedonic games [2] and modified fractional hedonic games [26]. Additively separable hedonic games have important subclasses where the focus lies in distinguishing friends and enemies, and therefore only two different weights are present in the underlying graph [16].

The computational properties of hedonic games have been extensively studied and we focus on literature related to additively separable hedonic games. Various versions of stability have been investigated [1, 3, 10, 16, 29, 21]. The closest to our work are the detailed studies of single-agent stability by Sung and Dimitrov [29] and Brandt et al. [10]. Gairing and Savani [21] settle the complexity of single-agent stability for symmetric input graphs. Majority-based stability has only received little attention thus far [10, 21]. Apart from stability, other desirable axioms concern efficiency and fairness. Aziz et al. [3] cover a wide range of axioms, whereas Elkind et al. [19] and Bullinger [12] focus on Pareto optimality, and Brandt and Bullinger [9] investigate popularity, an axiom combining ideas from stability and efficiency which is also related to certain majority-based stability notions [10]. Finally, a recent trend in the research on coalition formation is to complement the static view of existence problems by considering dynamics based on stability concepts (see, e.g., [7, 10, 11, 14, 23]).

2 Preliminaries

In this section, we formally introduce hedonic games and our considered stability concepts.

2.1 Hedonic Games

Let $N = [n]$ be a set of $n \in \mathbb{N}$ agents, where we define $[n] = \{1, \dots, n\}$. The output of a coalition formation problem is a coalition structure, that is, 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 $C \cap D = \emptyset$. An element of a partition is called a *coalition* and, given a partition π, the unique coalition containing agent i is denoted by $\pi(i)$. 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 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 $\mathcal{N}_i$ representing the preferences of the respective agent i. Hence, as mentioned before, agents express preferences only over the coalitions of which they are part 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 $D \succsim_i C$.

Additively separable hedonic games assume that every agent is equipped with a cardinal utility function that is aggregated by taking the sum of single-agent values. Formally, following [8], an *additively separable hedonic game* (ASHG) (N, v) consists of an agent set N and a tuple $v = (v_i)_{i\in N}$ of utility functions $v_i \colon N \to \mathbb{R}$ such that $\pi \succsim_i \pi'$ 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. When we specify ASHG utilities, we neglect, without loss of generality, $v_i(i)$ because the preferences do not depend on it and we implicitly assume that it is set to an appropriate constant if an ASHG has to fit into a certain subclass of games.

Every ASHG can be naturally represented by a complete directed graph $G = (N, E)$ with weight $v_i(j)$ on arc (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 $i, j \in N$ [10]. Further, following [16], 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 such games, agents seek to maximize their number of friends while minimizing their number of enemies, where these goals have a different priority in each case. 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 Single-Agent Stability

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

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

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

We consider 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).

Following [10], we introduce consent-based stability concepts via 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 [8, 17].

$$\begin{tikzcd}[column sep=small]
 & \text{Majority-Out Stability} \arrow[r] & \text{Contractual Nash Stability} \arrow[dr] & \\
\text{Nash Stability} \arrow[ur] \arrow[dr] & & & \text{Contractual Individual Stability} \\
 & \text{Majority-In Stability} \arrow[r] & \text{Individual Stability} \arrow[ur] &
\end{tikzcd}$$

Figure 1 Logical relationships between stability notions. An arrow from concept S to concept S' indicates that if a partition satisfies S, it also satisfies S'. Conversely, this means that every S' deviation is also an S deviation.

Finally, we define hybrid stability concepts according to [10] where the consent of the abandoned or welcoming coalition is decided by a majority vote. 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 $|F_{\text{out}}(\pi(i), i)| \geq |F_{\text{in}}(\pi(i), i)|$). Similar to before, a partition is said to be *majority-in stable* (MIS) or *majority-out stable* (MOS) if it allows for no majority-in or majority-out deviation, respectively. The concepts MIS and MOS are special cases of the voting-based stability notions by Gairing and Savani [21] for a threshold of $1/2$. Brandt et al. [10] also consider stability concepts that require voting-based consent by both the abandoned and welcoming coalition, similar to CIS.

For a stability concept $S \in \{\text{NS}, \text{IS}, \text{CNS}, \text{MIS}, \text{MOS}\}$, we denote the deviation corresponding to S as S *deviation*, e.g., CNS deviation for a contractual deviation. A taxonomy of our related solution concepts is provided in Figure 1.

3 Contractual Nash Stability

Our first result settles the computational complexity of contractual Nash stability in ASHGs. All of our reductions in this and the subsequent sections are from the NP-complete problem Exact3Cover (E3C) [25]. An instance of E3C consists of a tuple (R, S), where R is a ground set together with a set S of 3-element subsets of R. A Yes-instance is an instance such that there exists a subset $S' \subseteq S$ that partitions R.

Before giving the complete proof, we briefly describe the key ideas. Given an instance (R, S) of E3C, the reduced instance consists of three types of 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 the game identified by Sung and Dimitrov [28]. 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 S that would remain after removing an exact cover of R, i.e., there are $|S| - |R|/3$ 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 S. A gadget corresponding to an element $s \in S$ can either prevent non-stability 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 S. Then, the gadgets corresponding to elements in R can be dealt with by the cover and there are just enough elements in S to additionally deal with the other auxiliary gadgets.

▶ **Theorem 1.** *Deciding whether an ASHG contains a CNS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C and set $a =$

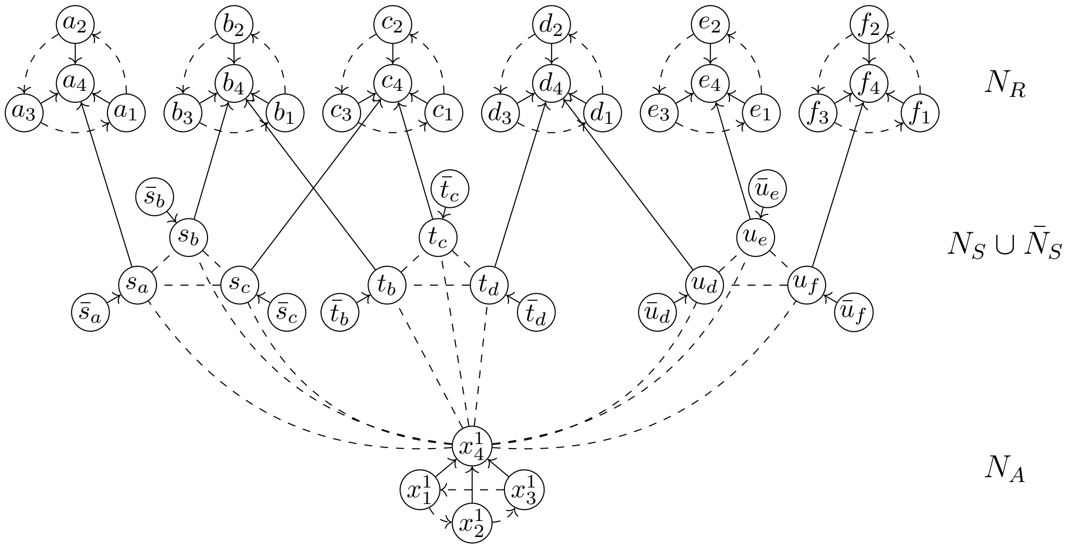

Figure 2 Schematic of the reduction from the proof of Theorem 1. We depict the reduced instance for the instance (R, S) of E3C where $R = \{a, b, c, d, e, f\}$, and $S = \{s, t, u\}$, with $s = \{a, b, c\}$, $t = \{b, c, d\}$, and $u = \{d, e, f\}$. Fully drawn edges mean a positive utility, which is usually 1 except between agents of the types $\bar{s}_r$ and s_r, where $v_{\bar{s}_r}(s_r) = 3$. Dashed edges represent a utility of 0. For agents in $\bar{N}_S$, only the single positive utility is displayed. Other omitted edges represent a negative utility of -4.

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

- $N_R = \cup_{r \in R} N_r$ with $N_r = \{r_i : i \in [4]\}$ for $r \in R$,
- $N_S = \cup_{s \in S} N_s$ with $N_s = \{s_r : r \in s\}$ for $s \in S$,
- $\bar{N}_S = \cup_{s \in S} \bar{N}_s$ with $\bar{N}_s = \{\bar{s}_r : r \in s\}$ for $s \in S$, and
- $N_A = \cup_{1 \leq j \leq a} N^j$ with $N^j = \{x_i^j : i \in [4]\}$ for $1 \leq j \leq a$.

We define valuations v as follows:

- For each $r \in R$, $i \in [3]$: $v_{r_i}(r_4) = 1$.
- For each $r \in R$, $(i, j) \in (1, 2), (2, 3), (3, 1)$: $v_{r_i}(r_j) = 0$.
- 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 $s \in S$, $r \in s$: $v_{s_r}(r_4) = 1$.
- For each $s \in S$, $r \in s$, $1 \leq j \leq a$: $v_{s_r}(x_4^j) = v_{x_4^j}(s_r) = 0$.
- For each $s \in S$, $r, r' \in s$: $v_{s_r}(s_{r'}) = 0$.
- For each $s \in S$, $r, r' \in s$, $r \neq r'$, $z \in (N_S \cup N_A) \setminus N_s$: $v_{\bar{s}_r}(s_r) = 3$, $v_{\bar{s}_r}(s_{r'}) = -2$, and $v_{\bar{s}_r}(z) = 0$.
- All other valuations are -4.

An illustration of the game is given in Figure 2. The agents in N_R in the reduced instance form gadgets consisting of a subgame without CNS partition for every element in R. The agents in N_A constitute further such gadgets. The agents in N_S consist of triangles for every set in S and are the only agents who can bind agents in the gadgets in any CNS partition. Finally, agents in $\bar{N}_S$ avoid having agents in N_S in separate coalitions to bind agents in N_A.

We claim that (R,S) is a Yes-instance if and only if (N,v) contains a CNS partition. Suppose first that $S' \subseteq S$ partitions R. Consider any bijection $\phi\colon S \setminus S' \to [a]$. Define a partition π by taking the union of the following coalitions:

- For every $r \in R, i \in [3]$, form $\{r_i\}$.
- For $s \in S', r \in s$, form $\{s_r, r_4\}$.
- For $s \in S \setminus S'$, form $\{s_r : r \in s\} \cup \{x_4^{\phi(s)}\}$.
- For $s \in S, r \in s$, form $\{\bar{s}_r\}$.
- For $1 \leq j \leq a, i \in [3]$, form $\{x_i^j\}$.

We claim that π is CNS. We will show that no agent can perform a deviation.

- For $r \in R$, $i \in [3]$, it holds that $v_{r_i}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{r_i}(\pi(r_4) \cup \{r_i\}) = -3$.
- For $r \in R$, r_4 is not allowed to leave her coalition.
- For $s \in S'$, $r \in s$, it holds that $v_{s_r}(\pi) = 1$ and joining any other coalition results in a negative utility. The agent s_r is in a most preferred coalition.
- For $s \in S \setminus S'$, $r \in s$, it holds that $v_{s_r}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{s_r}(\pi(r_4) \cup \{s_r\}) = -3$.
- For $s \in S'$, $r \in s$, the agent $\bar{s}_r$ obtains a non-positive utility by joining any other coalition. In particular, $v_{\bar{s}_r}(\pi(s_r) \cup \{\bar{s}_r\}) = -1$.
- For $s \in S \setminus S'$, $r \in s$, the agent $\bar{s}_r$ obtains a non-positive utility by joining any other coalition. In particular, $v_{\bar{s}_r}(\pi(s_r) \cup \{\bar{s}_r\}) = -1$.
- 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).

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

Let $r \in R$. Then, for all $i \in [3]$, $\pi(r_i) \subseteq N_r$. This follows because there is no agent who favors r_i in her coalition. Therefore, she would leave any coalition with an agent outside N_r to receive non-negative utility in a singleton coalition. Further, if there is no $s \in S$ with $r \in s$ such that $r_4 \in \pi(r_s)$, then $\pi(r_4) \subseteq N_r$. Indeed, if r_4 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 $r_4 \notin \pi(r_s)$ for all $s \in S$ with $r \in s$, then r_4 is in a singleton coalition, or there exists $i \in [3]$ with $r_4 \in \pi(r_i)$, for which we already know that $\pi(r_i) \subseteq N_r$.

Assume now that $\pi(r_4) \subseteq N_r$. For $i, i' \in [3]$, $r_i \notin \pi(r_{i'})$ because then one of them would receive a negative utility and could perform a CNS deviation to form a singleton coalition. If $\{r_4\} \in \pi$, then r_1 would deviate to join her. Hence, there exists exactly one $i \in [3]$ with $\{r_i, r_4\} \in \pi$. Suppose without loss of generality that $\{r_1, r_4\} \in \pi$. But then, r_3 would perform a CNS deviation to join them, a contradiction. We can conclude that there exists $s \in S$ with $r \in s$ such that $r_4 \in \pi(r_s)$. Hence, $s \in S'$ and we have shown that S' covers R.

To bound the cardinality of S', we will show that, for every $1 \leq j \leq a$, there exists $s \in S \setminus S'$ 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 $C \subseteq \{x_4^j\} \cup N_S$. Let $s \in S$ and $r \in s$ with $s_r \in C$. As we already know that $\bar{s}_r \notin C$, it must hold that $N_s \subseteq C$ to prevent her from joining. It follows that $s \notin S'$. 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 $\phi\colon [a] \to S \setminus S'$ such that, for every $1 \leq j \leq a$, $N_{\phi(j)} \subseteq \pi(x_4^j)$. Consequently, $|S'| \leq |S| - |\phi([a])| \leq |S| - a = |R|/3$. Hence, 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 $\{1, 0, -2, -4\}$, where the weight -4 may be replaced by an arbitrary smaller weight. By contrast, 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 [10, Theorem 4]. This encompasses for instance FEGs, AFGs, and AEGs. Hence, the hardness result is close to the boundary of computational feasibility.

4 Appreciation-of-Friends Games

In this section, we consider appreciation-of-friends games. Typically, these games behave well with respect to stability. In particular, IS, CNS, and MIS partitions always exist and can be computed efficiently, while it is only known that NS leads to non-existence and computational hardness among single-agent stability concepts [10, 16]. By contrast, we show in our next result that MOS partitions need not exist in AFGs. In other words, despite their conceptual complementarity, the stability concepts MOS and MIS lead to very different behavior in a natural class 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. Complete proofs for all omitted and sketched proofs can be found in the appendix.

▶ **Proposition 2.** *There exists an AFG without an MOS partition.*

Proof. We define the game formally. An illustration is given in Figure 3. 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 [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.

<u>Case 1</u>: $|\pi(z)| = 1$.

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

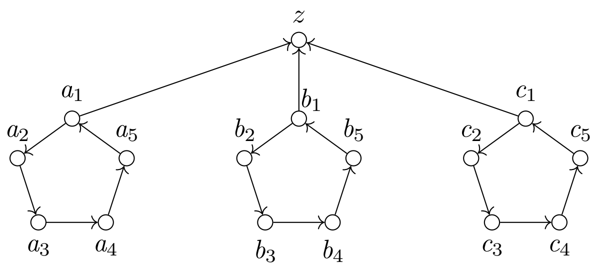

Figure 3 AFG without an MOS partition. The depicted (directed) edges represent friends, i.e., a utility of n, whereas missing edges represent a utility of -1.

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, $\pi(a_{i-1}) = \{a_{i-1}, a_{i-2}\} =: C$. 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 C. This is a contradiction and concludes the case that $|\pi(z)| = 1$.

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

Let $F := \{a_1, b_1, c_1\}$, i.e., the set of agents that have z as a friend. Note that z can perform an NS deviation to be a singleton. Hence, as π is MOS, $|F \cap \pi(v)| \geq |\pi(z)|/2$. 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 NS deviation. 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 $\pi(v)$ (recall that $|\pi(v)| \leq 3$). This is the final contradiction, and we can conclude that π is not MOS. ◀

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 cause problems for Nash stability. The constructive proof of the following proposition can be directly converted into a polynomial-time algorithm.

▶ **Proposition 3.** *Every AFG where 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 MOS for $n = 1$. Now, assume that (N, v) is an AFG with $n \geq 2$ 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 directed cycles 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 $R = \{x \in N : c \text{ reachable by } x\}$. 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 MOS. Define $\pi = \pi' \cup \{R\}$. We claim that π is MOS. 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 non-negative utility (otherwise, she can form a singleton coalition contradicting that π' is MOS). 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. We will maintain a tentative agent w that will continuously fulfill (i) and update w until this agent also fulfills (ii). Start with $w = y$. Note that this agent w fulfills (i) because y is a friend of x. 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 [24], i.e., a function $\sigma : N \to [n]$ 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 MOS 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 non-empty 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 2 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 4.** *Deciding whether an AFG contains an MOS partition is NP-complete.*

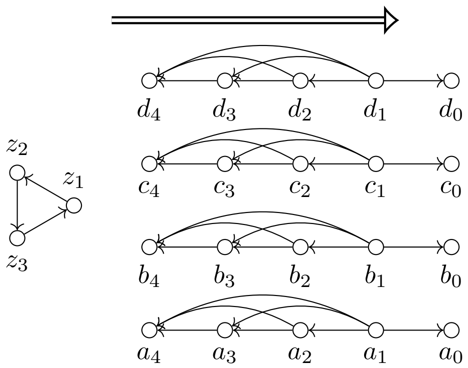

Figure 4 FEG without an MOS partition. The depicted (directed) edges represent friends. The double arrow means that every agent to the left of the tail of the arrow has every agent below the arrow as a friend.

5 Friends-and-Enemies Games

Friends-and-enemies games always contain efficiently computable stable coalition structures with respect to the unanimity-based stability concepts IS and CNS [10]. 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 MOS and MIS, 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 MOS.

▶ **Proposition 5.** *There exists an FEG without an MOS partition.*

Proof sketch. We only give a brief overview of the instance by means of the illustration in Figure 4. The FEG consists of a triangle of agents together with 4 sets of agents whose friendship relation is complete and transitive, together with one additional agent each that gives a temptation for the agent of the transitive substructures with the most friends.

An important reason for the non-existence 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. ◀

In the previous proof, it is particularly useful to establish disjoint coalitions of groups of agents who dislike each other. On the other hand, 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 MOS. This condition essentially means completeness of the friendship relation.[3] Note that this proposition is not true for other stability concepts such as NS or even IS.

▶ **Proposition 6.** *The grand coalition is MOS in every FEG with complete friendship relation.*

[3] Technically, the friendship relation may not be reflexive, but we can set $v_i(i) = 1$ for all $i \in N$ in an FEG to formally achieve completeness.

Proof. Let (N, v) be an FEG with complete friendship relation, and let π be the grand coalition. We claim that π is MOS. 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 non-existence of MOS partitions in FEGs shown in Proposition 5 can be leveraged to prove an intractability result. Interestingly, in contrast to the proofs of Theorem 1 and Theorem 4, 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 7.** *Deciding whether an FEG contains an MOS partition is NP-complete.*

In our next result, we construct an FEG without an MIS partition. Despite a lot of structure, the game is quite large encompassing 183 agents.

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

Proof sketch. We illustrate the example with the aid of Figure 5 and briefly discuss some key features. Again, 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 $\{a_i^0, \dots, a_i^9\}$ 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 possible to balance all diametric temptations, the instance does not admit an MIS partition. ◀

Similar to Proposition 6, it is easy to see that the singleton partition is MIS 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, MIS can also prevent typical run-and-chase games which do not admit NS partitions. We are ready to prove hardness of deciding on the existence of MIS partitions in FEGs.

▶ **Theorem 9.** *Deciding whether an FEG contains an MIS partition is NP-complete.*

6 Discussion and Conclusion

We have investigated single-agent stability in additively separable hedonic games. Our main results determine strong boundaries to the efficient computability of stable partitions. Table 1 provides a complete picture of the computational complexity of all considered stability notions and subclasses of ASHGs, where our results close all remaining open problems. First, we resolve the computational complexity of computing CNS partitions, which considers the last open unanimity-based stability notion in unrestricted ASHGs. The derived hardness result stands in contrast to positive results when considering appropriate subclasses such as FEGs,

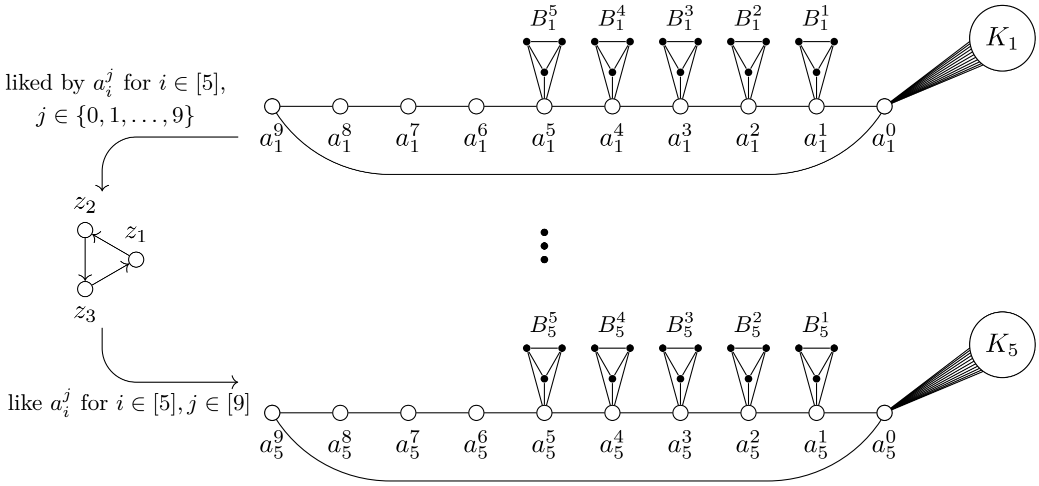

Figure 5 FEG without an MIS partition. The depicted edges represent friends. Undirected edges represent mutual friendship. For $i \in [5]$, some of the edges of agents in A_i are omitted. In fact, these agents form cliques. Also, each K_i represents a clique of 11 agents.

AEGs, or AFGs [10]. Second, our intractability concerning AFGs stands in contrast to known 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. 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 used as a black box.

Our results complete the 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. Their thorough investigation in other classes of hedonic games might lead to intriguing discoveries.

Table 1 Overview of the computational complexity of single-agent stability concepts in different classes of ASHGs. The NP-completeness results concern deciding on the existence of a stable partition. Membership in Function-P means that the search problem of constructing a stable partition can be solved in polynomial time.

ASHG	Unrestricted	Friends-and-enemies games	Appreciation-of-friends games
NS	NP-complete [29]	NP-complete [10]	NP-complete [10]
IS	NP-complete [29]	Function-P [10]	Function-P [16]
CNS	NP-complete (Th. 1)	Function-P [10]	Function-P [10]
MIS	NP-complete [10]	NP-complete (Th. 9)	Function-P [10]
MOS	NP-complete [10]	NP-complete (Th. 7)	NP-complete (Th. 4)

References

1 H. Aziz and F. Brandl. Existence of stability in hedonic coalition formation games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 763–770, 2012.

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

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

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

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

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

7 V. Bilò, A. Fanelli, M. Flammini, G. Monaco, and L. Moscardelli. Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. *Journal of Artificial Intelligence Research*, 62:315–371, 2018.

8 A. Bogomolnaia and M. O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

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

10 F. Brandt, M. Bullinger, and L. Tappe. Single-agent dynamics in additively separable hedonic games. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, 2022. Forthcoming.

11 F. Brandt, M. Bullinger, and A. Wilczynski. Reaching individually stable coalition structures in hedonic games. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5211–5218, 2021.

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

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

14 R. Carosi, G. Monaco, and L. Moscardelli. Local core stability in simple symmetric fractional hedonic games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 574–582, 2019.

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

16 D. Dimitrov, P. Borm, R. Hendrickx, and S. C. Sung. Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, 26(2):421–433, 2006.

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

18 J. H. Drèze and J. Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica*, 48(4):987–1003, 1980.

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

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

21 M. Gairing and R. Savani. Computing stable outcomes in symmetric additively separable hedonic games. *Mathematics of Operations Research*, 44(3):1101–1121, 2019.

22 D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

23 M. Hoefer, D. Vaz, and L. Wagner. Dynamics in matching and coalition formation games with structural constraints. *Artificial Intelligence*, 262:222–247, 2018.
24 A. B. Kahn. Topological sorting of large networks. *Communications of the ACM*, 5(11):558–562, 1962.
25 R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.
26 M. Olsen. On defining and computing communities. In *Proceedings of the 18th Computing: Australasian Theory Symposium (CATS)*, volume 128 of *Conferences in Research and Practice in Information Technology (CRPIT)*, pages 97–102, 2012.
27 W. Suksompong. Individual and group stability in neutral restrictions of hedonic games. *Mathematical Social Sciences*, 78:1–5, 2015.
28 S. C. Sung and D. Dimitrov. On myopic stability concepts for hedonic games. *Theory and Decision*, 62(1):31–45, 2007.
29 S. C. Sung and D. Dimitrov. Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3):635–639, 2010.

A Omitted Proofs

In the appendix we provide missing proofs and full proofs whenever we only provided a proof sketch.

A.1 Appreciation-of-Friends Games

▶ **Theorem 4.** *Deciding whether an AFG contains an MOS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C. We define an ASHG (N, v) as follows. Let $N = N_R \cup N_S$ where $N_R = \cup_{r\in R} N^r$ and $N_S = \cup_{s\in S} N_s$ with $N^r = \{a_i^r, b_i^r, c_i^r : i \in [5]\} \cup \{z^r\}$ for $r \in R$ and $N_s = \{s_r : r \in s\} \cup \{s_0\}$ for $s \in S$. 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 $[5]$.

We define utilities v as follows:

- For all $s \in S$, $r \in s$: $v_{s_r}(s_0) = n$.
- For all $s \in S$, $r, r' \in s$, $r \neq r'$: $v_{s_r}(s_{r'}) = n$.
- For all $s \in S$, $r \in s$: $v_{s_r}(a_1^r) = n$.
- For all $r \in R$, $i \in [5]$, and $x \in \{a, b, c\}$: $v^r_{x_i}(x^r_{i+1}) = n$.
- For all $r \in R$, $x \in \{a, b, c\}$: $v_{x_1^r}(z) = n$.
- All other valuations are -1.

An illustration of the reduction is provided in Figure 6. Intuitively, 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 2. The sets in S are represented by S-gadgets consisting of a triple of agents representing its elements in R which are linked to the respective R-gadgets. Furthermore, there is one special agent without any friends attracting the other agents in the S-gadget.

We claim that (R, S) is a Yes-instance if and only if the reduced AFG contains an MOS partition. Suppose first that $S' \subseteq S$ 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 $s \in S', r \in s$, form $\{s_r, a_1^r\}$.
- For $s \in S'$, form $\{s_0\}$.
- For $s \in S \setminus S'$, form N_s.

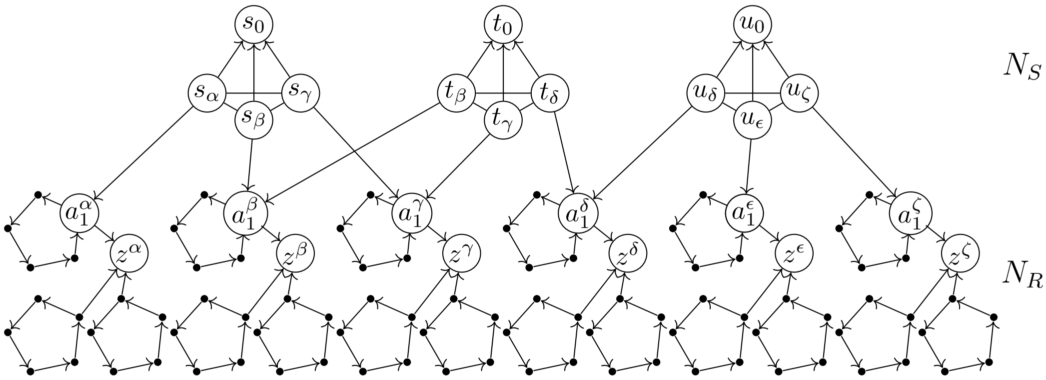

Figure 6 Schematic of the reduction from the proof of Theorem 4. We depict the reduced instance for the instance (R, S) of E3C where $R = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$ and $S = \{s, t, u\}$ with $s = \{\alpha, \beta, \gamma\}$, $t = \{\beta, \gamma, \delta\}$, and $u = \{\delta, \epsilon, \zeta\}$. Directed edges indicate a 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 2.

We prove that π is MOS 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 $s \in S$ and $r \in s$, the agent s_r 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 $s \in S'$, the agent s_0 is in her most preferred coalition and has no incentive to perform a deviation. Finally, for $s \in S \setminus S'$, the agent s_0 is not allowed to perform an MOS deviation.

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 10. For all $s \in S$, it holds that $\pi(s_0) = \{s_0\}$ or $N_s \subseteq \pi(s_0)$.

Proof. Let $s \in S$, say $s = \{u, w, x\}$, and define $C = \pi(s_0)$ and $D = \{s_u, s_w, s_x\}$. Assume that $C \supsetneq \{s_0\}$. Then, since s_0 has no friends, she would prefer to stay in a singleton coalition. Hence, $C \cap D \neq \emptyset$, say $s_u \in C$.

Assume for contradiction that $D \setminus C \neq \emptyset$, say $s_w \notin C$. Then, $s_x \in \pi(s_w)$. Indeed, if $s_x \notin \pi(s_w)$, then s_w has at most one friend in her coalition, and no agent would prevent her from performing an MOS deviation to join C. Hence, $s_x \in \pi(s_w)$. Then, $C = \{s_0, s_u\}$, as s_0 could leave her coalition to form a singleton coalition if any other agent was part of it. But then, s_u has an incentive to join $\pi(s_w)$, and could perform a valid MOS deviation to do so. This is a contradiction and therefore $D \subseteq C$. ◁

In the next claim, we improve upon Claim 10 and show that there are in fact only two possible coalitions for s_0.

▷ Claim 11. For all $s \in S$, it holds that $\pi(s_0) = \{s_0\}$ or $\pi(s_0) = N_s$.

Proof. Let $s \in S$ and define $C = \pi(s_0)$. Assume that $C \supsetneq \{s_0\}$. By Claim 10, it holds that $N_s \subseteq C$ and since s_0 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 $t \in S$ with $t \neq s$ and $u, v \in t$ with $C = N_s \cup \{t_u, t_v\}$. Then, by Claim 10, $\{t_0\} \in \pi$, implying that t_u has an MOS deviation to join t_0. This is a contradiction and we can therefore conclude that $\pi(s_0) = N_s$. ◁

Next, we consider the coalitions of other agents in gadgets related to sets in S.

▷ Claim 12. For all $s \in S$ and $r \in R$, it holds that $\pi(s_r) = \{s_r, a_1^r\}$ or $N_s \setminus \{s_0\} \subseteq \pi(s_r)$.

Proof. Let $s \in S$, say $s = \{r, u, w\}$, and define $C = \pi(s_r)$. If $s_0 \in C$, then $C = N_s$ by Claim 11 and the assertion is true. Suppose therefore that $s_0 \notin C$. Assume now that there is $x \in s$ with $s_x \notin C$, say $s_u \notin C$. If $s_w \notin C$, then no agent in C has s_r as a friend and could therefore vote against a deviation. Moreover, since the deviation of s_r to join s_0 is not an MOS deviation, it must be the case that $v_{s_r}(\pi) = n$, which can, under the given assumptions, only be the case if $\pi(s_r) = \{s_r, a_1^r\}$.

It remains to consider the case that $s_w \in C$. But then, s_u is in a coalition with at most one friend (note that it is excluded that $s_0 \in \pi(s_u)$ by Claim 11) and no agent in her coalition has her as a friend. Hence, s_u has an MOS deviation to join C, a contradiction. Together, we have shown that if there is $x \in s$ with $s_x \notin C$, then $\pi(s_r) = \{s_r, a_1^r\}$, which proves this claim. ◁

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

▷ Claim 13. For all $s \in S$, $r \in s$, and $u \in R$, it holds that if $\pi(s_r) \cap N^u \neq \emptyset$, then $r = u$ and $\pi(s_r) = \{s_r, a_1^u\}$.

Proof. Let $s \in S$, $r \in s$, and $u \in R$. The assertion is true if $\pi(s_r) = \{s_r, a_1^r\}$. Hence, by Claim 12, we may assume that $N_s \setminus \{s_0\} \subseteq C$. we will show that $\pi(s_r) \cap N^u = \emptyset$. 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 $N_s \setminus \{s_0\}$ 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 $t \in S$ with $u \in t$ and $t_u \in C$. But then, Claim 12 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, a contradiction. Therefore, we have shown that $\pi(s_r) \cap N^u = \emptyset$. ◁

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

▷ Claim 14. 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 11 and Claim 13, it holds that $C \cap N_s = \emptyset$ for all $s \in S$. 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 13. 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. ◁

Finally, we can conclude the proof by showing that there exists $S' \subseteq S$ partitioning R. Therefore, let $S' = \{s \in S : \pi(s_r) = \{s_r, a_1^r\}$ for some $r \in s\}$. We show that 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 2, 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 11, Claim 13, and Claim 14, this can only be the case if there exists $s \in S$ with $r \in s$ and $\pi(s_r) = \{s_r, a_1^r\}$. Consequently, S' covers R.

For the second part, assume for contradiction that some element in R is covered at least twice by sets in S'. Then, there exists $s \in S'$ with $r \in s$ and $\{s_r, a_1^r\} \notin \pi$. By Claim 12, $N_s \setminus \{s_0\} \subseteq \pi(s_r)$. But then, according to the definition of S', it follows that $s \notin S'$, a contradiction. Hence, the elements of S' are disjoint sets. This completes the proof. ◀

A.2 Majority-Out Stability in Friends-and-Enemies Games

In this section, we provide missing proofs about majority-based stability concepts in FEGs.

▶ **Proposition 5.** *There exists an FEG without an MOS partition.*

Proof. Recall that the game is illustrated in Figure 4. Formally, 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 15 describes possible coalitions of agents x_0 where $x \in \{a, b, c, d\}$. Second, Claim 16 establishes that coalitions from agents of different sets of N_x, $x \in \{a, b, c, d\}$, are disjoint. Then, Claim 17 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 15. 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. ◁

▷ Claim 16. 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 15, $x_0 \notin \Gamma$ 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

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

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.

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

We consider now the agents in N_c. By Claim 15, $(*)$, 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 15, it follows that $\pi(c_1) \subseteq C \setminus \{c_0\}$.

Let $\Phi \subseteq C \setminus \{c_0\}$ be a coalition of 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.

Case 2: 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 $z_{i^*} \notin \Gamma$, 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. ◁

▷ Claim 17. 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 15 and Claim 16, 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 15, $a_0 \notin \Gamma$. We claim that $|\Gamma \cap N_a| \leq 3$. For the contrary, assume that $|\Gamma \cap N_a| = 4$. Then, Claim 15 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 16, 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. ◁

We are ready to obtain a final contradiction. By Claim 17, there exist $i, j \in [3]$ with $z_i \notin \pi(z_j)$. Without loss of generality, we may assume that $i = 2$ and $j = 1$.

Case 1: It holds that $z_3 \in \pi(z_2)$.

By Claim 15, $v_{z_k}(x) = 1$ 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.

Case 2: The agents in N_z are in pairwise disjoint coalitions.

Let $p_k = |\pi(z_k)|$ for $k \in [3]$ and $k^* = \arg\max_{k \in [3]} p_i$. Without loss of generality, $k^* = 1$. As in the previous case, it follows from Claim 15 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. ◀

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 18.** *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 forth 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 proceed with proving the hardness result.

▶ **Theorem 7.** *Deciding whether an FEG contains an MOS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R,S) be an instance of E3C. We define a reduced FEG (N,v) as follows. By Proposition 5, 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 $N = N_R \cup N_S$ where $N_R = \cup_{r\in R} N^r$ with $N^r = \{y^r \colon y \in N'\}$ for $r \in R$ and $N_S = \cup_{s\in S} N_s$ with $N_s = \{s_0\} \cup \{s_r \colon r \in s\}$ for $s \in S$. Specifically, we denote the agent corresponding to the special agent $x \in N'$ by x^r. Agents of the type s^r 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 $s \in S$, $y,z \in N_s$: $v_y(z) = 1$.
- For all $s \in S$, $r \in s$: $v_{s_r}(x^r) = 1$.
- For all $r \in R$ and $y,z \in N'$: $v_{y^r}(z^r) = v'_y(z)$, 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.

We claim that (R,S) is a Yes-instance if and only if the reduced FEG contains an MOS partition. Suppose first that $S' \subseteq S$ partitions R. We define a partition π as follows.

- For $s \in S \setminus S'$: $N_s \in \pi$ and for $s \in S'$: $\{s_0\} \in \pi$.
- For $s \in S'$, $r \in s$: $\{s_r, x^r\} \in \pi$.
- For $r \in R$ and $z \in N' \setminus \{x\}$: $\pi(z^r) = \{y^r \in y \in \pi'(x)\}$.

We claim that the partition π is MOS.

- Let $r \in R$ and consider an agent $y \in N' \setminus \{x\}$. Then, y^r cannot perform an MOS deviation to join $\pi(z^r)$ for any $z \in N' \setminus \{x\}$, because π' restricted to $N' \setminus \{x\}$ is an MOS partition. Moreover, joining $\pi(z)$ for any $z \in N \setminus N^r$ yields utility at most 0 (in fact, the only such coalition that y^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 $s \in S'$. Then $v_{s_0}(\pi) = 0$ and joining any other coalition yields utility at most 0. In particular, $v_{s_0}(\pi(s_r) \cup \{s_0\}) = 0$ for all $r \in s$. Moreover, for $r \in s$, $v_{s_r}(\pi) = 1$ and joining any other coalition yields utility at most 1. In particular, $v_{s_r}(\pi(s_0) \cup \{s_r\}) = 1$.
- Let $s \in S \setminus S'$. Then, $\pi(s_0)$ is s_0's most preferred coalition and she has no incentive to perform an MOS deviation. Moreover, for $r \in s$, $v_{s_r}(\pi) = 3$ and joining any other coalition yields a utility of at most 0.

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

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 s_0.

$\triangleright$ Claim 19. Let $s \in S$. Then, $\pi(s_0) = \{s_0\}$ or $\pi(x) \subseteq N_s$ for all $x \in N_s$.

Proof. Let $s \in S$ and define $C := \pi(s_0)$. A close inspection of the utilities in the definition of the reduced instance lets us apply Lemma 18 multiple times to conclude that
- for all $s' \in S \setminus \{s\}$, $C \cap N_{s'} = \emptyset$,
- 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, $F_{\text{in}}(C, x^r) \subseteq \{s_r\}$ and $s_0 \in F_{\text{out}}(C, x^r)$. Hence, x^r could perform an MOS deviation to form a singleton coalition. We can therefore conclude that $C \subseteq N_s$.

Assume that $C \supsetneq \{s_0\}$. If $|C| = 3$, then there exists a unique $r \in s$ with $s_r \notin C$. Since s_r has only one friend outside C, this would imply that $v_{s_r}(\pi) \leq 1$ whereas $v_{s_r}(C \cup \{s_r\}) = 3$ and $F_{\text{in}}(\pi(s_r), s_r) = \emptyset$. Hence, s_r 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 $C = \{s_0, s_{r_1}\}$. Then, it holds that $s_{r_3} \in \pi(s_{r_2})$, as otherwise $v_{s_{r_2}}(\pi) \leq 1$ whereas $v_{s_r}(C \cup \{s_{r_2}\}) = 3$ and $F_{\text{in}}(\pi(s_{r_2}), s_{r_2}) = \emptyset$. But then, $\pi(s_{r_2}) = \{s_{r_2}, s_{r_3}\}$. Any other agent would only have enemies in $\pi(s_{r_2})$ and is allowed to leave by a weak majority. This concludes the proof of the claim. $\triangleleft$

Our next claim investigates elements $s \in S$ for which $\{s_0\} \in \pi$.

$\triangleright$ Claim 20. Let $s \in S$ such that $\{s_0\} \in \pi$. Then, for every $r \in s$, it holds that $\pi(s_r) = \{s_r, x^r\}$.

Proof. Let $s \in S$ with $\{s_0\} \in \pi$ and consider any $r \in s$. Define $C := \pi(s_r)$ and assume for contradiction that there exists $r' \in s$ with $r' \neq r$ and $s_{r'} \in C$. We can combine the following observations:
- Claim 19 shows that $s'_0 \notin C$ for every $s' \in S \setminus \{s\}$.
- Let $\hat{r} \in R$. We can apply Lemma 18 for s_r (or $s_{r'}$) 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 $s' \in S$ and $\hat{r} \in s'$. We can apply Lemma 18 for s_r (or $s_{r'}$) and $s'_{\hat{r}}$ to show that $s'_{\hat{r}} \notin C$ if $\hat{r} \neq r$ (or $\hat{r} = r$).

Together, the observations show that $C \subseteq N_s$. But then, $v_{s_0}(C \cup \{s_0\}) \geq 2$, and s_0 could perform an MOS deviation to join C. This is a contradiction and we can conclude that $C \cap N_s = \{r_s\}$.

This means in particular, that $F_{\text{in}}(C, s_r) = \emptyset$. Since $v_{s_r}(\{s_0, s_r\}) = 1$, it must hold that $v_{s_r}(\pi) = 1$. Since the unique friend of s_r outside N_s is x^r, we can conclude that $\pi(s_r) = \{s_r, x^r\}$. $\triangleleft$

We are ready to finish the proof. Therefore, let $S' := \{s \in S \colon \{s_0\} \in \pi\}$. We show that S' partitions R in two steps. First, the sets in S' are disjoint. Indeed, if $s, s' \in S'$ with $s \neq s'$ and $r \in s \cap s'$, then Claim 20 implies that $\{s_r, x^r\} \in \pi$ and $\{s'_r, x^r\} \in \pi$, contradicting the fact that π is a partition.

It remains to show that all elements of R are covered by a set in S'. Therefore, consider an arbitrary $r \in R$ and let $y \in N'$. By Lemma 18, $\pi(y^r) \cap N^{r'} = \emptyset$ for all $r' \in R$ with $r' \neq r$. Moreover, Claim 19 and Claim 20 imply that $\pi(y^r) \cap N_s = \emptyset$ for all $s \in S$ with $r \notin s$. Assume for contradiction that there exists no $s \in S'$ with $r \in s$. Then, Claim 19 implies that $\pi(y^r) \cap N_s = \emptyset$ for all $s \in S$ with $r \in s$. Together, $\pi(y^r) \subseteq N^r$. 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 S'. ◀

A.3 Majority-In Stability in Friends-and-Enemies Games

We start this section by providing the full proof investigating the FEG without MIS partition. First, 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 5).

▶ **Lemma 21.** *Consider an FEG (N, v) with an MIS partition π. Let $W \subseteq N$ such that the following conditions hold:*

1. *For all $i, j \in W$, $k \in N$: $v_i(j) = 1$.*
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 MIS, it holds that $v_i(\pi) \geq 0$ and therefore, because the game is an FEG,

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

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. ◀

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

Proof. We define an FEG for which we prove that it does not contain an MIS partition. As discussed before, the game is rather large (the number of agents is 183), but it has a lot of structure and an illustration was already provided in Figure 5. 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 $A_i = \{a_i^j : j = \{0, 1, \dots, 9\}$ 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 $(x, y) \in \{(z_1, z_2), (z_2, z_3), (z_3, z_1)\}$,
- $v_{z_i}(a_j^l) = 1$ if $i \in [3]$, $j \in [5]$, and $l \in [9]$,
- $v_{a_i^j}(a_i^l) = 1$ if $i \in [5]$, $j, l \in \{0, 1, \dots, 9\}$,
- $v_{a_i^j}(z_l) = 1$ if $i \in [5]$, $j \in \{0, 1, \dots, 9\}$, and $l \in [3]$,
- $v_{a_i^0}(k_i^j) = v_{k_i^j}(a_i^0) = 1$ if $i \in [5]$, $j \in [11]$,
- $v_{a_i^j}(b_i^{j,l}) = 1$ if $i, j \in [5]$, $l \in [3]$,
- $v_{b_i^{j,l}}(b_i^{j,l'}) = 1$ if $i, j \in [5]$, $l, l' \in [3]$,
- $v_{k_i^j}(k_i^l) = 1$ 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:

Every agent receives non-negative utility in π, i.e.,
$v_i(\pi) \geq 0$ for all $i \in N$. $(*)$

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 21 for the agents in sets K_i for $i \in [5]$.

▷ Claim 22. For all $i \in [5]$, there exists $C \in \pi$ with $K_i \subseteq C$.

The next claim improves upon the previous claim.

▷ Claim 23. If $i \in [5]$, then $K_i \in \pi$ or $K_i \cup \{a_i^0\} \in \pi$.

Proof. Let $i \in [5]$ and assume for contradiction that there exists $C \in \pi$ with $K_i \subseteq C$ and $C \setminus (K_i \cup \{a_i^0\}) \neq \emptyset$. By $(*)$, $v_{k_i^1}(\pi) \geq 0$ and therefore $|C \setminus (K_i \cup \{a_i^0\})| \leq |K_i \cup \{a_i^0\}| - 1 = 11$. Let $x \in C \setminus (K_i \cup \{a_i^0\})$. Then, $a_i^0 \in C$, $|C \setminus (K_i \cup \{a_i^0\})| = 11$, and $v_x(y) = 1$ for all $y \in C \setminus (K_i \cup \{a_i^0\})$. Otherwise, x has at most 10 friends leading to $v_x(\pi) \leq 10 - |K_i| < 0$, contradicting $(*)$. Consequently, the agents $C \setminus (K_i \cup \{a_i^0\})$ form a set of 11 mutual friends which all have a_i^0 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_i^j. First, we can apply Lemma 21 again for a statement analogous to Claim 22.

▷ Claim 24. For all $i, j \in [5]$, there exists $C \in \pi$ with $B_i^j \subseteq C$.

We also refine this claim.

▷ Claim 25. 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 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 25 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 the 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 26. 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 23, Claim 25, 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 23, $K_i \in \pi$ 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 27. 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 23 and Claim 25, $C \subseteq Z \cup \bigcup_{i \in [5]} A_i$. Define $I = \{i \in [5] \colon A_i \cap C \neq \emptyset\}$ and let

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

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

Case 1: $|I| = 5$.

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

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

Case 2: $|I| = 4$.

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 26, 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.

Case 3: $|I| = 3$.

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.

Case 4: $|I| = 2$.

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 25, $B^j_{i^*} \in \pi$ 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 25, $B^j_{i'} \in \pi$ and since π is MIS, $u_{a^j_{i'}}(\pi) \geq 3$. 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 25, 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. ◁

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

▷ Claim 28. 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 $Z \nsubseteq C$.

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 23, $K_i \in \pi$ and a^0_i would perform an MIS deviation to join this coalition. By this observation and using Claim 23 and Claim 25, 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 26, $A_1 \setminus \{a^0_1\} \in \pi$ 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 27 and Claim 28 directly leads to a contradiction, we have shown that the constructed FEG has no MIS partition. ◀

Towards turning this counterexample into an intractability result for FEGs, 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 29.** *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 $v_i(k) + v_j(k) \leq 0$. 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 $v_j(\pi) < 0$, a contradiction. ◀

▶ **Theorem 9.** *Deciding whether an FEG contains an MIS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C. We define an FEG (N, v) as follows. Let $N = N_R \cup N_S$ where $N_R = \cup_{r \in R} N^r$ and $N_S = \cup_{s \in S} N_s$ with $N_s = V_s \cup \bigcup_{r \in s} V_s^r$ for $s \in S$. There, we define, for $s \in S$, $V_s = \{c_{s,i} \colon i \in [10]\}$, and for $s \in S$ and $r \in s$, $V_s^r = \{c_{s,i}^r \colon i \in [10]\}$. 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_s^r by means of a positive utility correspondence. We define utilities v as follows:

- For all $s \in S$, $x, y \in N_s$: $v_x(y) = 1$.
- For all $s \in S$, $r \in s$, and $x \in V_s^r$: $v_x(z_1^r) = v_{z_1^r}(x) = 1$.
- 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, S) is a Yes-instance if and only if the reduced FEG contains an MIS partition. Suppose first that $S' \subseteq S$ 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 $i, j \in [5]$, $B_i^j \in \pi'$.
- For $i \in \{2,3,4,5\}$, $A_i \setminus \{a_i^0\} \in \pi'$ and $K_i \cup \{a_i^0\} \in \pi'$.

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

- For $s \in S \setminus S'$: $N_s \in \pi$ and for $s \in S'$: $V_s \in \pi$.
- For $s \in S'$, $r \in s$: $V_s^r \cup \{z_1^r\} \in \pi$.
- For $r \in R$ and $x \in N' \setminus \{z_1\}$: $\pi(x^r) = \{y^r : y \in \pi'(x)\}$.

Showing that π is MIS 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 $s \in S \setminus S'$ and $x \in N_s$, $v_x(\pi) = 39$, and joining any other coalition does not give agent x positive utility.
- For $s \in S'$ and $x \in V_s$, $v_x(\pi) = 9$, and joining any other coalition does not give her a better utility. In particular, joining $V_s^r \cup \{z_1^r\}$ for $r \in s$ would also give her a utility of 9.
- For $s \in S'$, $r \in s$, and $x \in V_s^r$, $v_x(\pi) = 10$, 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).

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

Note that the assumptions of Lemma 29 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 29 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 $s, s' \in S$, $s \neq s'$, $x \in V_s$, and $y \in N_{s'}$, it holds that $y \notin \pi(x)$.
3. For every $s \in S$, $r \in R \setminus s$, $x \in N_s$, and $y \in N^r$, it holds that $y \notin \pi(x)$.
4. For every $s \in S$, $r \in s$, and $x \in V_s$, it holds that $\pi(x) \cap N^r \subseteq \{z_1^r\}$.

Next, we can apply Lemma 21 to obtain the next two facts.

5. For every $s \in S$, there exists a coalition $C \in \pi$ with $V_s \subseteq C$.
6. For every $s \in S$, $r \in S$, there exists a coalition $C \in \pi$ with $V_s^r \subseteq C$.

Moreover, combining Lemma 29 with Fact 6 allows us to further refine Fact 4 yielding the fact

7. For every $s \in S$, $r \in s$, and $x \in V_s$, it holds that $V_s^r \subseteq \pi(x)$ whenever $z_1^r \in \pi(x)$.

We are ready to restrict the coalitions of agents in sets V_s to two possibilities.

$\triangleright$ Claim 30. For all $s \in S$, it holds that $V_s \in \pi$ or $N_s \in \pi$.

Proof. Let $s \in S$ 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 $V_s^r \subseteq C$. 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 $C \supsetneq V_s \cup V_s^r \cup \{z_1^r\}$. This would contradict that π is an MIS partition and therefore $C = V_s \cup V_s^r \cup \{z_1^r\}$. 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 $y' \in V_s^{r'}$ and $y'' \in V_s^{r''}$. If there exists an agent $q \in N \setminus (V_s^{r'} \cup V_s^{r''})$, 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 $D = V_s^{r'} \cup V_s^{r''} \cup \{z_1^{r'}, z_1^{r''}\}$. But then, $v_{z_1^{r'}}(\pi) = -1$, a contradiction. Hence, $D = V_s^{r'} \cup V_s^{r''}$. 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 $V_s^r \subseteq C$, 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 $y \in V_s^{r'}$. Note that $v_y(C \cup \{y\}) \geq 20$, and therefore, it must hold that $v_y(\pi) \geq 20$. Hence, $V_s^{r'} \cup V_s^{r''} \cup \{z_1^{r'}\} \subseteq \pi(y)$. Therefore, even $z_1^{r''} \in \pi(y)$ because otherwise, an agent in $V_s^{r''}$ 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. $\triangleleft$

Our next goal is to pinpoint the coalitions of agents in sets of the type V_s^r.

▷ Claim 31. For all $s \in S$ and $r \in s$, it holds that $V_s^r \cup \{z_1^r\} \in \pi$ or $N_s \in \pi$.

Proof. For $s \in S$ and $r \in s$ consider an agent $x \in V_s^r$ and define $C := \pi(x)$. Assume that $C \neq N_s$. We have to show that $C = V_s^r \cup \{z_1^r\}$. By Claim 30, we know then that $V_s \in \pi$. By Fact 3, we know that $C \subseteq N_S \cup \bigcup_{t\in s} N^t$. Assume that $s = \{r, r', r''\}$.

Assume for contradiction that there exists an agent $y \in (V_s^{r'} \cup V_s^{r''}) \cap C$. 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 29 for q and one of x and y. A similar argument shows that $N_S \cap C \subseteq N_s$. Hence, $C \subseteq N_s \cup \bigcup_{t\in s}\{z_1^t\}$.

By Fact 6 and our assumptions, we know that in addition $V_s^r \cup V_s^t \subseteq C$ for $t \in s$ with $y \in V_s^t$. 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 $C \subseteq N_s \cup (N_S \setminus N_s) \cup \bigcup_{t\in s} N^t$, this is only possible if $C = V_s^r \cup \{z_1^r\}$. ◁

We are ready to prove that (R, S) is a Yes-instance. Define $S' = \{s \in S \colon N_s \notin \pi\}$. First, note that the sets in S' are disjoint. Indeed, let $s \in S'$ and consider $r \in s$. By Claim 31, $V_s^r \cup \{z_1^r\} \in \pi$. Hence, for every $s' \in S \setminus \{s\}$ with $r \in s'$, it cannot be the case that $V_{s'}^r \cup \{z_1^r\} \in \pi$. Hence, another application of Claim 31 yields $N_{s'} \in \pi$, and therefore $s' \notin S'$.

It remains to show that S' covers all elements in R. Therefore, let $r \in R$. By Fact 1, Claim 30, and Claim 31, 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 $\pi(z_1^r) = V_s^r \cup \{z_1^r\}$ for some $s \in S$. 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., $\pi(z_1^r) = V_s^r \cup \{z_1^r\}$ for some $s \in S$. Then, $s \in S'$, and therefore r is covered by an element in S'. ◀