

Stability in Online Coalition Formation

Martin Bullinger

*Department of Computer Science
University of Oxford
7 Parks Road, Oxford OX1 3QD, United Kingdom*

MARTIN.BULLINGER@CS.OX.AC.UK

René Romen

*School of Computation, Information and Technology
Technical University of Munich
Boltzmannstr 3, 85748 Munich, Germany*

RENE.ROMEN@TUM.DE

Abstract

Coalition formation is concerned with the question of how to partition a set of agents into disjoint coalitions according to their preferences. Deviating from most of the previous work, we consider an online variant of the problem, where agents arrive in sequence. Whenever an agent arrives, they must be assigned to a coalition immediately and irrevocably. The scarce existing literature on online coalition formation has focused on maximizing social welfare, a demanding requirement, even in the offline setting. Instead, we seek to achieve *stable* coalition structures online and treat the most common stability concepts based on deviations by single agents and groups of agents. We present a comprehensive picture in additively separable hedonic games, leading to dichotomies, where positive results are obtained by deterministic algorithms and negative results even hold for randomized algorithms.

1. Introduction

The formation of stable coalition structures is an important concern in multi-agent systems. The critical question is how to partition a set of agents into reasonable coalitions. A standard framework for this is the consideration of hedonic games (Drèze & Greenberg, 1980). In these games, a set of agents expresses their preferences over subsets of agents containing themselves, i.e., their potential coalitions. The output is a coalition structure where all agents are assigned to a unique coalition. In our work, we consider additively separable hedonic games, one of the most prominent classes of hedonic games, where cardinal utilities for single agents encode the preferences, and a sum-based aggregation defines the utility for a coalition (Bogomolnaia & Jackson, 2002).

Hedonic games have been used to model various aspects of group interaction, such as the formation of research teams (Alcalde & Revilla, 2004) or the detection of communities (Aziz et al., 2019). A commonality of most research on hedonic games is that the focus is on fully specified games, for which a desirable outcome is searched. However, this misses an important feature of many real-life scenarios: Agents might arrive over time and have to be assigned to existing coalitions. For instance, in a company, most workers are already assigned to a department or team, and new hires join existing teams.

Based on such considerations, Flammini et al. (2021) introduced an online variant of hedonic games that adds the arrival of agents over time. Their goal is to achieve high social

welfare, and they find that a greedy algorithm performs best in a competitive analysis. However, this algorithm has an unbounded competitive ratio if the number of agents or the utility range is unbounded. In subsequent work, Bullinger and Romen (2023) show that it is possible to get rid of utility dependencies of the competitive ratio under certain model assumptions, e.g., a uniformly random arrival of agents. They achieve a competitive ratio of $\Theta(n)$, which is essentially the best *offline* approximation guarantee that we can hope for by efficient algorithms. Indeed, for every $\epsilon > 0$, it is NP-complete to approximate social welfare by a factor of $n^{1-\epsilon}$ (Flammini et al., 2022, Theorem 17).¹

By contrast, the scarce existing literature on online coalition formation omits other common objectives in coalition formation. Stability is perhaps the most studied solution concept for hedonic games in general and in additively separable hedonic games in particular (see, e.g., Bogomolnaia & Jackson, 2002; Sung & Dimitrov, 2010; Aziz et al., 2013; Woeginger, 2013; Gairing & Savani, 2019; Brandt et al., 2023, 2024). In our work, we close this research gap and consider the question of whether notions of stability can be achieved in an online manner.

We cover a broad range of the most common stability concepts for hedonic games based on deviations by single agents and groups of agents. Specifically, we consider the single-deviation concepts of Nash stability, individual stability, contractual Nash stability, and contractual individual stability, as well as the group-deviation concepts of the core and strict core. In addition, we study Pareto optimality, which is particularly interesting because it is a natural weakening of the demanding objective of maximizing social welfare while it can still be interpreted as a notion of stability (Morrill, 2010). We present a comprehensive picture of the capabilities and impossibilities of online algorithms aiming to compute stable partitions. Note that for some of the solution concepts, e.g., Nash stability, no partition satisfies the stability notion in some instances. This also implies trivial impossibilities for the online setting. Therefore, we consider natural utility restrictions that are known to allow for stable outcomes. These are symmetric utilities and restricted utilities based on the distinction between friends and enemies. The latter is a natural approach that has been thoroughly explored in coalition formation settings (Dimitrov et al., 2006; Ota et al., 2017; Flammini et al., 2022; Kerkmann et al., 2022; Brandt et al., 2024). It yields classes of games with just two utility values: a positive utility for all friends and a negative utility for all enemies. Within the framework of additively separable hedonic games, prominent such subclasses encompass appreciation-of-friends games (AFGs) and aversion-to-enemies games (AEGs), as first considered by Dimitrov et al. (2006).

Our findings are summarized in Table 1. A large part of our results is negative and proves the nonexistence of randomized algorithms capable of computing stable coalition structures under strong utility restrictions. This is even the case for solution concepts like Pareto optimality or contractual individual stability, for which solutions are guaranteed to exist in any hedonic game, i.e., including those that are not additively separable (Aziz & Savani, 2016). In fact, all our negative results only use games in which coalition structures satisfying the considered solution concept are guaranteed to exist. By contrast, we obtain deterministic online algorithms capable of computing contractually Nash-stable and Pareto-

1. This result even holds in restricted classes of games that we will introduce and investigate later. The proof by Flammini et al. (2022) holds for aversion-to-enemies games, and Bullinger et al. (2025) extend the result to, among others, friends-and-enemies games.

Table 1: Computability of stable partitions by online algorithms; for definitions of stability concepts and utility restrictions, see Section 3. A checkmark (✓) means that a deterministic online algorithm can compute the desired partition. A cross (✗) means that no randomized online algorithm exists that outputs the desired partition with probability bounded away from 0. All negative results hold even for the case of symmetric games. Of the positive results (highlighted in gray), only the results for contractual Nash stability need symmetry.

Utility restriction of ASHG	Strict	FENG	FEG	AFG	AEG
Allowed utility values	$\mathbb{Q} \setminus \{0\}$	$\{1, 0, -1\}$	$\{1, -1\}$	$\{n, -1\}$	$\{1, -n\}$
Nash stability	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)
Individual stability	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)	✗(Th. 4.10)
Contractual Nash stability	✗(Th. 4.4)	✗(Th. 4.9)	✓(Th. 4.2)	✗(Th. 4.4)	✓(Th. 4.2)
Contractual individual stability	✓(Cor. 4.6)	✗(Th. 4.9)	✓(Cor. 4.6)	✓(Cor. 4.6)	✓(Cor. 4.6)
Strict core stability	✗(Th. 4.11)	✗(Th. 4.11)	✗(Th. 4.11)	✗(Th. 4.12)	✗(Th. 4.11)
Core stability	✗(Th. 4.11)	✗(Th. 4.11)	✗(Th. 4.11)	✗(Th. 4.12)	✗(Th. 4.11)
Pareto optimality	✓(Th. 4.5)	✗(Th. 4.9)	✓(Th. 4.5)	✓(Th. 4.5)	✓(Th. 4.5)

optimal coalition structures in restricted classes of games. While such positive results seem rare, they entail very strong stability guarantees. The associated algorithms do not only output a final stable coalition structure, but they *maintain* stability throughout the entire arrival process of agents. Otherwise, they would fail their promised guarantee on a partial instance. Hence, these algorithms suit every application with an indefinite time horizon, where new agents can arrive continuously.

2. Related Work

Our work contributes to two streams of work: the consideration of coalition formation models in economic theory and, more recently, the AI literature, as well as the investigation of online algorithms in related settings, mostly in theoretical computer science. Here, we give an account of both.

2.1 Coalition Formation

Coalition formation in the framework of hedonic games was first studied by Drèze and Greenberg (1980) and popularized two decades later (Banerjee et al., 2001; Ceclárová & Romero-Medina, 2001; Bogomolnaia & Jackson, 2002). Bogomolnaia and Jackson (2002) introduced additively separable hedonic games (ASHGs), which have since been an ongoing object of study. Aziz and Savani (2016) and Bullinger et al. (2024) present introductory texts on hedonic games.

The majority of the research on ASHG considers the offline setting. It focuses on the computational complexity of stability concepts. In Table 2, we provide an overview of the complementary offline results for our results as presented in Table 1. Since strict utilities

Table 2: Existence and computational complexity of stable partitions in offline additively separable hedonic games; for definitions of stability concepts and utility restrictions, see Section 3. We indicate whether stable partitions are guaranteed to exist (\exists), may not exist (\nexists), exist and can be computed in polynomial time (FP, i.e., Function-P), or whether deciding if there exists a stable partition is hard. ^a: Aziz and Brandl (2012), ^b: Aziz et al. (2013), ^c: Brandt et al. (2024) ^d: Dimitrov et al. (2006), ^e: Olsen (2009), ^f: Sung and Dimitrov (2010), ^g: Peters (2017), ^h: Woeginger (2013)

Utility restriction of ASHG	None	FEG	AFG	AEG
Allowed utility values	\mathbb{Q}	$\{1, -1\}$	$\{n, -1\}$	$\{1, -n\}$
Nash stability	NP-complete ^e	NP-complete ^c	NP-complete ^c	NP-complete ^c
Individual stability	NP-complete ^f	FP ^c	FP ^d	FP ^a
Contractual Nash stability	NP-complete ^c	FP ^c	FP ^c	FP ^c
Contractual individual stability	FP ^b	FP ^b	FP ^b	FP ^b
Strict core stability	Σ_2^p -complete ^g	\nexists	FP ^d	\nexists^d
Core stability	Σ_2^p -complete ^h	?	FP ^d	\exists , FNP-hard ^d
Pareto optimality	\exists	FP ^b	FP ^b	FP ^b

and friend-enemies-and-neutrals games (FENGs) were previously not considered in detail, we omit them from the table.

The existence problem for Nash stability was first shown to be NP-complete by Olsen (2009). His reduction was from the weakly NP-complete partition problem. Subsequently, a strong NP-completeness result was achieved by Sung and Dimitrov (2010). This was then further strengthened by Brandt et al. (2024), who show that the result holds under utility restrictions as in FEGs, AFGs, and AEGs.

Without utility restrictions, individual stability and contractual Nash stability also have an NP-complete decision problem (Sung & Dimitrov, 2010; Brandt et al., 2024). Moreover, the computational complexity for the core and strict core is even higher (Woeginger, 2013; Peters, 2017). Still, utility restrictions lead to better existence results. For single-deviation concepts, Brandt et al. (2024) show that deviation dynamics converge to stable partitions in polynomial time when utilities are restricted as in FEGs, AFGs, and AEGs.² Furthermore, partitions in the strict core of AFGs and in the core of AEGs always exist (Dimitrov et al., 2006). However, while the former can be elicited in polynomial time, finding the latter is FNP-hard (i.e., hard as a search problem). Notably, except for the core in AFGs, these existence guarantees for group stability break down once agents with a utility of 0 are allowed (Ota et al., 2017). To the best of our knowledge, the core and strict core have not been considered when utilities are restricted to $\{-1, 1\}$, but it is easy to construct a game where the strict core is empty under these utility restrictions.³

2. The existence and efficient computability of individually stable partitions in AFGs and AEGs can also be inferred from results by Dimitrov et al. (2006) and Aziz and Brandl (2012), respectively.

3. Consider a symmetric game with 5 agents, where the positive utilities form a cycle.

Pareto-optimal and contractually individually stable partitions are guaranteed to exist; in fact, Pareto optimality implies contractual individual stability. Aziz et al. (2013) provide a polynomial-time algorithm for the former when utilities cannot assume a value of 0, and for the latter without utility restrictions. Arguably, a utility of 0 feels quite natural because it expresses indifference towards other players. Hence, towards computing Pareto-optimal partitions in more generality, Bullinger (2020) provides an algorithm that works under a weak condition for 0-utilities. In particular, it works if these utilities are mutual. It is, however, an open problem to compute Pareto-optimal partitions in unrestricted ASHG. In the context of Pareto optimality in hedonic games, Elkind et al. (2020) introduce the price of Pareto optimality, which measures the welfare loss of Pareto-optimal outcomes.

Apart from restricting utility values, some other conditions lead to the existence of stable outcomes. Notably, Nash-stable (and, therefore, stable outcomes under all single-deviation concepts) exist if utilities are symmetric (Bogomolnaia & Jackson, 2002). In fact, Nash-stable partitions are found when running deviation dynamics, the canonical associated local search algorithm. Nonetheless, this is not an efficient procedure and leads to PLS-completeness, i.e., hardness for local search problems, even when considering individual stability (Gairing & Savani, 2019). Additionally, Bullinger and Kraicz (2024) prove that individually stable and contractually Nash-stable partitions exist and can be computed efficiently with high probability in random (nonsymmetric) additively separable hedonic games. By contrast, they show that Nash-stable outcomes do not exist with high probability.

As we mentioned in the introduction, online ASHG have been introduced by Flammini et al. (2021) and subsequently been studied by Bullinger and Romen (2023). Moreover, Pavone et al. (2022) study online hypergraph matching. Their model can be interpreted as coalition formation with bounded coalition sizes. In contrast to Flammini et al. (2021) and Bullinger and Romen (2023), the work by Pavone et al. (2022) does not assume additively separable utilities, and agents do not have to be matched immediately at arrival. Instead, they depart unmatched after a fixed time. All three works solely consider the maximization of social welfare or the minimization of total cost.

In addition, a recent series of work considers deviation dynamics for hedonic games, which constitute another time-dependent model of coalition formation (see, e.g., Bilò et al., 2018; Brandt et al., 2023; Carosi et al., 2019). In particular, ASHG and close variants are studied in depth (Bilò et al., 2022; Brandt et al., 2024; Bullinger & Suksompong, 2024; Boehmer et al., 2025).

2.2 Online Algorithms

From the literature on online algorithms, online matching is the most related to our setting, which originates from the seminal paper by Karp et al. (1990). Huang and Tröbst (2023) survey this line of work. Matchings can be seen as a variant of hedonic games, where coalitions are restricted to size at most 2. Unlike our work, Karp et al. (1990) consider a setting where the input instances are bipartite, and only one side of the agents appears online. The objective in online matching is usually to find a matching of maximum cardinality or weight. Karp et al. (1990) introduce the famous ranking algorithm, which achieves a competitive ratio of $1 - 1/e$. Subsequent work considers related models with edge weights, all agents arriving online, or nonbipartite matching (Feldman et al., 2009; Wang & Wong,

2015; Huang et al., 2018; Ezra et al., 2022). While it is possible to achieve the competitive ratio of $1 - 1/e$ in the weighted setting (Feldman et al., 2009), this is usually impossible if all agents arrive online (Wang & Wong, 2015; Huang et al., 2018). Closest to our setting is the work by Ezra et al. (2022). They show an optimal bound of $5/12$ for maximum weight matching in general graphs in the online random arrival setting and provide an algorithm that matches this bound asymptotically.

Additionally, stability has been considered for online matching to some extent. Doval (2022) extends stability according to Gale and Shapley (1962) to an online setting and shows that her extension can always be satisfied. Still, her model has several conceptual differences from the online models discussed thus far. Most notably, agents do not have to be matched immediately (but suffer from a discount in utility if matched later) and may arrive in batches. Moreover, Gajulapalli et al. (2020) study a two-stage process for school choice with the goal of preserving stability. While this is generally impossible, they present efficient algorithms that maximize the number of additionally matched agents or minimize the number of reallocations compared to the matching of the first stage. More loosely related, Benadè and Sahoo (2023) touch upon an online model for recommender systems, where they—mostly experimentally—investigate stability.

3. Preliminaries

In this section, we present preliminaries. For an integer $i \in \mathbb{N}$, we define $[i] := \{1, \dots, i\}$ and we define $[0] := \emptyset$.

3.1 Additively Separable Hedonic Games

Let N be a finite set of n agents. Any subset of N is called a *coalition*. We denote the set of all possible coalitions containing agent $i \in N$ by $\mathcal{N}_i := \{C \subseteq N : i \in C\}$. A *coalition structure* (or *partition*) is a partition of the agents. Given an agent $i \in N$ and a partition π , let $\pi(i)$ denote the coalition of i , i.e., the unique coalition $C \in \pi$ with $i \in C$.

A *hedonic game* is a pair (N, \succsim) consisting of a set N of agents and a preference profile $\succsim = (\succsim_i)_{i \in N}$ where \succsim_i is a weak order over \mathcal{N}_i which represents the preferences of agent i . The semantics is that agent i strictly prefers coalition C to coalition C' if $C \succ_i C'$ and is indifferent between coalitions C and C' if $C \sim_i C'$.

Following Bogomolnaia and Jackson (2002), an *additively separable hedonic game* (ASHG) consists of a set N of agents and a tuple $u = (u_i)_{i \in N}$ of *utility functions*⁴ $u_i : N \setminus \{i\} \rightarrow \mathbb{Q}$, such that, for every pair of coalitions $C, C' \in \mathcal{N}_i$, it holds that

$$C \succsim_i C' \text{ if and only if } \sum_{j \in C \setminus \{i\}} u_i(j) \geq \sum_{j \in C' \setminus \{i\}} u_i(j).$$

We usually represent an ASHG by the pair (N, u) . Clearly, an ASHG is a hedonic game. We abuse notation and extend the definition of u to coalitions $C \in \mathcal{N}_i$ and partitions π by $u_i(C) := \sum_{j \in C \setminus \{i\}} u_i(j)$ and $u_i(\pi) := u_i(\pi(i))$, respectively.

4. We assume utility functions to be rational-valued as our standard computational framework. However, all of our results also hold for real-valued utilities and computations.

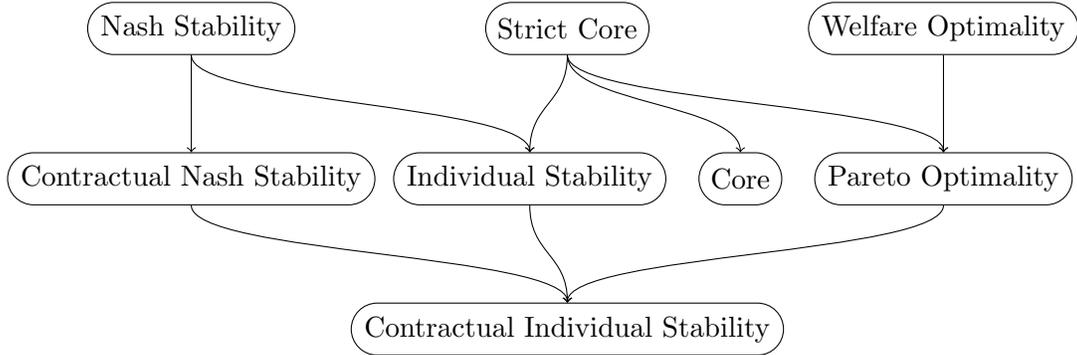


Figure 1: Logical relationships between our solution concepts (see, e.g., Aziz & Savani, 2016). An arrow from concept α to concept β indicates that if a partition satisfies α , then it also satisfies β .

Also, an ASHG can be represented equivalently by a complete directed graph $G = (N, E)$ with weight $u_i(j)$ on arc (i, j) . An ASHG is said to be *symmetric* if, for every pair of agents $i, j \in N$ with $i \neq j$, it holds that $u_i(j) = u_j(i)$. In this case, we also write $u(i, j) := u_i(j)$. A complete undirected graph can naturally represent a symmetric ASHG. An ASHG is *strict* if, for every pair of agents $i, j \in N$ with $i \neq j$, it holds that $u_i(j) \neq 0$ (Aziz et al., 2013).

There are various important subclasses of ASHGs with restricted utility values. Given a subset $U \subseteq \mathbb{Q}$, an ASHG is called a U -ASHG if, for every pair of agents $i, j \in N$ with $i \neq j$, it holds that $u_i(j) \in U$. In particular, some ASHGs allow a natural interpretation in terms of friends and enemies. Following Dimitrov et al. (2006) and Brandt et al. (2024), a U -ASHG is called an

- *appreciation-of-friends game* (AFG) if $U = \{n, -1\}$,
- *aversion-to-enemies game* (AEG) if $U = \{1, -n\}$,
- *friends-and-enemies game* (FEG) if $U = \{1, -1\}$, and
- *friends-enemies-and-neutrals game* (FENG) if $U = \{1, 0, -1\}$.

In all of these games, the utility for a coalition depends on the distinction of friends and enemies, i.e., players that yield positive and negative utility, respectively. In FEGs and FENGs, friends and enemies have equal importance, whereas in AFGs, a single friend outweighs an arbitrary number of enemies, and in AEGs, a single enemy annihilates any number of friends.

3.2 Solution Concepts

In this section, we define the solution concepts considered in this paper. Figure 1 provides an overview of their logical relationships. We assume that we are given a fixed ASHG (N, u) .

Notions of stability capture agents' incentives to perform deviations (Bogomolnaia & Jackson, 2002; Dimitrov & Sung, 2007). A *single-agent deviation* performed by agent i

transforms a partition π into a partition π' where $\pi(i) \neq \pi'(i)$ and, for all agents $j \neq i$, it holds that $\pi(j) \setminus \{i\} = \pi'(j) \setminus \{i\}$. The basic idea of deviations is that the deviating agent should immediately benefit from a deviation. A *Nash deviation* is a single-agent deviation performed by agent i such that $u_i(\pi') > u_i(\pi)$. Any partition in which no Nash deviation is possible is said to be *Nash-stable* (NS).

The drawback of Nash stability is that only the preferences of the deviating agent are considered, which might seem too demanding in the context of cooperation. Therefore, various refinements have been proposed, which additionally require the consent of other involved agents. An *individual deviation* is a Nash deviation by agent i transforming π into π' such that, for all agents $j \in \pi'(i) \setminus \{i\}$, it holds that $u_j(\pi') \geq u_j(\pi)$. Similarly, a *contractual deviation* is a Nash deviation by agent i transforming π into π' such that, for all agents $j \in \pi(i) \setminus \{i\}$, it holds that $u_j(\pi') \geq u_j(\pi)$. Hence, an individual deviation requires the unanimous consent of the agents within the welcoming coalition, while a contractual deviation analogously pertains to the abandoned coalition. The latter is appropriate in settings where the membership in a coalition is based on a contract that would only be dissolved if this is beneficial for all involved parties. A partition is then said to be *individually stable* (IS) or *contractually Nash-stable* (CNS) if it allows for no individual or contractual deviation, respectively. A single-agent deviation is called a *contractual individual deviation* if it is both a contractual deviation and an individual deviation. A partition is said to be *contractually individually stable* (CIS) if it allows for no contractual individual deviation.

Next, we introduce stability based on group deviations. Consider a partition π and a coalition $B \subseteq N$. Then, B is called a *blocking coalition* for π if, for all agents $i \in B$, it holds that $u_i(B) > u_i(\pi)$. Moreover, B is called a *weakly blocking coalition* for π if, for all agents $i \in B$, it holds that $u_i(B) \geq u_i(\pi)$, and there exists an agent $j \in B$ with $u_j(B) > u_j(\pi)$. A partition is said to be in the *core* (CR) if it admits no blocking coalition, and in the *strict core* (SCR) if it admits no weakly blocking coalition. Note that every blocking coalition is also weakly blocking. Hence, the strict core prevents a larger set of possible group deviations and, therefore, is the stronger solution concept. Some authors refer to the strict core as strong core (Bogomolnaia & Jackson, 2002). When we speak of performing a group deviation, we mean that agents abandon their respective coalition and jointly form a new coalition.

For a more concise notation, we refer to deviations with respect to stability concept $\alpha \in \{\text{NS, IS, CNS, CIS, CR, SCR}\}$ as α *deviations*, e.g., IS deviations for $\alpha = \text{IS}$. Similarly, we refer to a partition satisfying stability concept α as an α partition.

Finally, we also consider Pareto optimality, which can be seen as a stability guarantee, where the whole set of agents cannot beneficially rearrange their coalitions. A partition π' is said to *Pareto-dominate* a partition π if, for every agent $i \in N$, it holds that $u_i(\pi') \geq u_i(\pi)$, and there exists an agent $j \in N$ with $u_j(\pi') > u_j(\pi)$. A partition π is said to be *Pareto-optimal* (PO) if it is not Pareto-dominated by another partition. Pareto optimality is a classical concept of economic efficiency and was already a primal objective during the birth of hedonic games (Drèze & Greenberg, 1980).

At first glance, Pareto dominance feels like a global variant of weakly blocking coalitions and, therefore, seemingly leads to a more demanding solution concept. However, the exact opposite is true. While group deviations based on (weakly) blocking coalitions lead to partitions that can be inferior for other agents, every Pareto dominance gives rise to a

weakly blocking coalition (the ones containing agents that are strictly better off). Hence, partitions in the strict core are Pareto-optimal.

The *social welfare* of a partition π is defined as $\mathcal{SW}(\pi) := \sum_{i \in N} u_i(\pi)$, i.e., it is the sum of the utilities of all agents. A partition is said to be *welfare-optimal* if it maximizes social welfare among all partitions. Note that Pareto optimality is a weakening of welfare optimality, which was the objective in the literature on online ASGHs thus far (Flammini et al., 2021; Bullinger & Romen, 2023).

Notably, apart from the relationships indicated in Figure 1 by arrows (and their transitive closure), there are no other logical relationships between the solution concepts that we introduced. For example, there exists an additively separable hedonic game containing a partition that is contractually Nash-stable but neither individually stable, nor in the core, nor Pareto-optimal (and, therefore, neither Nash-stable, nor in the strict core, nor welfare-optimal), see Example 4.3.

3.3 Online Coalition Formation

In this section, we introduce our model of online coalition formation, following the notation of Bullinger and Romen (2023). The online setting is not restricted to ASGHs, so we define it for a general hedonic game $G = (N, \succsim)$. Let $\Sigma(G) := \{\sigma: [|N|] \rightarrow N \text{ bijective}\}$ be the set of all *orders* of the agent set N . Given a subset of agents $M \subseteq N$, let $G[M]$ be the hedonic game restricted to agent set M , i.e., the hedonic game (M, \succsim^M) where $C \succsim_i^M D$ if and only if $C \succsim_i D$. Moreover given a partition π of a set N and a subset of agents $M \subseteq N$, we define $\pi[M]$ as the *partition restricted to M* as $\pi[M] := \{C \cap M: C \in \pi, C \cap M \neq \emptyset\}$. Specifically, if M consists of all agents except a single agent $i \in N$, then we write $\pi - i := \pi[N \setminus \{i\}]$.

An instance of an *online coalition formation* problem is a pair (G, σ) , where $G = (N, \succsim)$ is a hedonic game and $\sigma \in \Sigma(G)$. An *online coalition formation algorithm* for instance (G, σ) gets as input the sequence G_1, \dots, G_n , where, for every $i \in [n]$, $G_i := G[\{\sigma(j): 1 \leq j \leq i\}]$. Then, for every $i \in [n]$, the algorithm has to produce a partition π_i of $\{\sigma(j): 1 \leq j \leq i\}$ such that

- the algorithm has only access to G_i and π_{i-1} , and
- the newly formed partition π_i satisfies $\pi_i - \sigma(i) = \pi_{i-1}$, i.e., extends the existing partition.

The output of the algorithm is the partition π_n . Given an online coalition formation algorithm ALG , let $ALG(G, \sigma)$ be its output, for instance (G, σ) . If σ is clear from the context, we omit it from this notation and simply write $ALG(G)$.

More informally, the algorithm iteratively creates a partition such that it only has access to the utilities of the currently present agents when irrevocably adding a new agent to an existing or new coalition. In addition to deterministic algorithms, we also consider randomized algorithms. This means that the decisions as to which coalition an agent is added to can be random.

Unlike welfare optimality, stability concepts do not naturally yield a quantitative maximization objective, and we cannot directly perform the usual competitive analysis. Instead, we have qualitative objectives that are either satisfied or not by an output. Therefore,

we desire algorithms that output stable partitions with a high probability if agents arrive online, which once again is a quantitative objective.

Consider a solution concept $\alpha \in \{\text{NS, IS, CNS, CIS, PO, CR, SCR}\}$ and an algorithm ALG . We define the α *guarantee* of ALG as

$$W_\alpha(ALG) := \inf_G \min_{\sigma \in \Sigma(G)} \mathbb{P}(ALG(G, \sigma) \text{ is an } \alpha \text{ partition}).$$

Here, the probability is taken according to the randomized decisions of ALG . Hence, $W_\alpha(ALG)$ is the worst-case probability that ALG outputs an α partition.

Note that the α guarantee of deterministic algorithms is either one—if the algorithm always outputs an α partition—or zero—if the algorithm does not output an α partition for some input instance.

4. Results

In this section, we present our results. Recall that only instances that allow for a stable partition are relevant for the consideration of stability in online coalition formation. Otherwise, no algorithm, and, therefore, especially no online algorithm, can make any guarantee.⁵

In the literature on stability in ASHG, two restrictions have turned out to be vital for the existence of stable partitions, namely symmetry and severe utility restrictions (Bogomolnaia & Jackson, 2002; Dimitrov et al., 2006; Brandt et al., 2024). First, in symmetric ASHG, Nash-stable partitions (and, therefore, partitions satisfying all weaker single-deviation stability notions) are guaranteed to exist (Bogomolnaia & Jackson, 2002). Moreover, in (possibly asymmetric) FEGs, AEGs, and AFGs, it is guaranteed that individually stable and contractually Nash-stable partitions exist (Brandt et al., 2024). Finally, utility restrictions may also lead to group stability (Dimitrov et al., 2006). In particular, AFGs contain partitions in the strict core. While this is not the case for AEGs, these at least contain partitions in the core. In this section, we will see (the conjunction of) which of these assumptions are sufficient to allow for the computation of stable outcomes online, and which of the results in the offline setting cause problems online.

4.1 Contractual Nash Stability

We start with the consideration of contractual Nash stability, where every agent in the abandoned coalition can veto a single-agent deviation. As a warm-up, we begin with a simple proposition that gives a first hint as to why computing stable partitions in an online manner is a nontrivial task. Even the conjunction of symmetry and utility restrictions is not sufficient for computing CNS partitions.

Proposition 4.1. *There exists no deterministic online algorithm, which always outputs a CNS partition for symmetric AFGs.*

Proof. Assume for contradiction that ALG is a deterministic algorithm that always outputs a CNS partition for symmetric AFGs. Consider the following two AFGs (N, u) and

5. We can also take the viewpoint of comparing the capabilities of online algorithms with offline possibilities: if no stable partition exists, then any online algorithm is as good as an optimal offline algorithm.

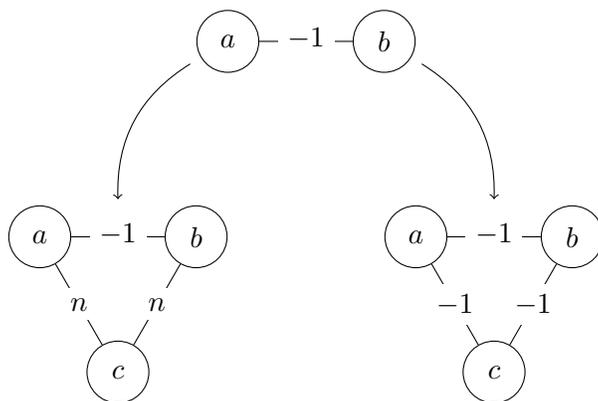


Figure 2: Adversarial AFGs for computing CNS partitions. Every deterministic algorithm fails for one of the two possible instances.

(N, \hat{u}) with identical agent set $N = \{a, b, c\}$ and symmetric utilities $u(a, b) = \hat{u}(a, b) = -1$, $u(a, c) = u(b, c) = n$ (in this case, $n = 3$), and $\hat{u}(a, c) = \hat{u}(b, c) = -1$. Consider the arrival order (a, b, c) . Before the arrival of c , ALG cannot distinguish, whether the game will be (N, u) or (N, \hat{u}) . The situation is depicted in Figure 2.

If ALG creates $\{a, b\}$ at the arrival of b , then it fails to compute a CNS partition for (N, \hat{u}) . Hence, ALG has to create a new coalition $\{b\}$. When c arrives, ALG cannot form a new singleton coalition as otherwise a (or b) has a CNS deviation to join c . Assume without loss of generality that ALG forms $\{a, c\}$. Then, b has a CNS deviation to join them. Hence, ALG fails to compute a CNS partition for (N, u) . \square

However, the previous result seems to rely on small negative utilities. By contrast, we now present an online algorithm that computes CNS partitions for other restricted classes of ASHG. The basic idea of our algorithm is to establish that agents in a coalition of size at least 2 are not allowed to leave their coalition because some other agent would veto this. Moreover, we use our assumption on the utility values to show that agents in singleton coalitions can never gain positive utility by joining a constructed coalition.

Theorem 4.2. *Let $y \geq x > 0$. Then, there exists a deterministic online algorithm, which always outputs a CNS partition for symmetric $\{-y, x\}$ -ASHGs.*

Proof. Let $y \geq x > 0$ and consider a symmetric $\{-y, x\}$ -ASHG with agent set $N = \{a_i : 1 \leq i \leq n\}$ and arrival order (a_1, \dots, a_n) . We apply Algorithm 1 to compute a partition π . This algorithm proceeds as follows. Whenever a new agent arrives, we compute the set of present agents with positive utility for the new agent. Note that, by symmetry, this implies a mutual positive utility. Assume there is at least one. Then, we check if at least one of them is in a singleton coalition. If such an agent exists, we let the new agent join this singleton coalition. Otherwise, we add a_i to any coalition of an agent with positive utility. If no such agent exists, we form a new singleton coalition.

Algorithm 1 Contractually Nash-stable partition of online symmetric $\{-y, x\}$ -ASHGs for $y \geq x > 0$.

Input: Symmetric $\{-y, x\}$ -ASHG

Output: Contractually Nash-stable partition π

```

 $\pi \leftarrow \emptyset$ 
for  $i = 1, \dots, n$  do
   $N_i \leftarrow \{j \in [i-1] : u(a_i, a_j) > 0\}$ 
  if  $\exists j \in N_i$  with  $|\pi(a_j)| = 1$  then
     $\pi \leftarrow \pi \setminus \{\{a_j\}\} \cup \{\{a_i, a_j\}\}$ 
  else if  $\exists j \in N_i$  with  $|\pi(a_j)| > 1$  then
     $\pi \leftarrow \pi \setminus \{\pi(a_j)\} \cup \{\pi(a_j) \cup \{a_i\}\}$ 
  else
     $\pi \leftarrow \pi \cup \{\{a_i\}\}$ 
return  $\pi$ 

```

We claim that the partition π is contractually Nash-stable. We show the claim by induction. Recall that, for every $i \in [n]$, π_i is the partition created by the algorithm after agent a_i has been assigned to a coalition.

Claim 1. *For every $i \in [n]$, the following statements are true:*

1. *The partition π_i is CNS.*
2. *For every coalition $C \in \pi_i$ with $|C| \geq 2$ and $a_k \in C$, there exists an agent $a_\ell \in C$ with $u(a_k, a_\ell) = x$.*
3. *For every $k \in [i]$ with $|\pi_i(a_k)| = 1$, it holds that, if $\ell \in [i]$ with $u(a_k, a_\ell) = x$, then $|\pi_i(a_\ell)| \geq 2$ and $u(a_k, b) = -y$ for all $b \in \pi_i(a_\ell) \setminus \{a_\ell\}$.*

Proof. All three statements are true for $i = 1$. Assume now that the statements are true for some $1 \leq i < n$. Let $N_{i+1} := \{j \in [i] : u(a_{i+1}, a_j) > 0\}$. We start by proving the second and third assertions by a case distinction according to the different cases in the algorithm.

Assume first that there exists $j \in N_{i+1}$ with $|\pi_i(a_j)| = 1$ and that we have $\pi_{i+1} = \pi_i \setminus \{\{a_j\}\} \cup \{\{a_j, a_{i+1}\}\}$. Clearly, the second assertion for $i + 1$ follows by induction and since $u(a_{i+1}, a_j) = x$. For the third assertion, let $k \in [i+1]$ with $|\pi_{i+1}(a_k)| = 1$ and consider $\ell \in [i+1]$ with $u(a_k, a_\ell) = x$. Since $|\pi_{i+1}(a_{i+1})| = 2$, it holds that $k \neq i+1$. By the induction hypothesis for the third assertion, the third assertion is true unless $\ell \in \{i+1, j\}$. Moreover, again by the induction hypothesis for the third assertion, $u(a_k, a_j) = -y$, and, therefore, the assertion is true if $\ell = i+1$.

Next, assume that there exists no $j \in N_{i+1}$ with $|\pi_i(a_j)| = 1$, but $N_{i+1} \neq \emptyset$ and that $\pi_{i+1} = \pi_i \setminus \{\pi_i(a_j)\} \cup \{\pi_i(a_j) \cup \{a_{i+1}\}\}$ for some $j \in N_{i+1}$. Note that the second assertion is true for agent a_{i+1} because $a_j \in \pi_{i+1}(a_{i+1})$. For all other agents, the second assertion follows by induction. As there exists no $j \in N_{i+1}$ with $|\pi_i(a_j)| = 1$, the third assertion follows by induction.

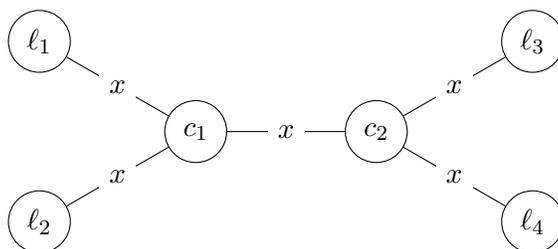


Figure 3: Instance defined in Example 4.3. Missing edges indicate a symmetric utility of $-y$. Algorithm 1 returns the grand coalition consisting of all agents, which is neither individually stable, nor in the core, nor Pareto-optimal.

Finally, if $N_{i+1} = \emptyset$, then $\pi_{i+1} = \pi_i \cup \{\{a_{i+1}\}\}$. Hence, the second assertion follows by induction, and the third assertion follows by induction for all agents except for a_{i+1} . For a_{i+1} , it is true because $N_{i+1} = \emptyset$.

It remains to prove the first assertion. We show how it follows from the second and third assertions. By the second assertion, every agent in a coalition of size at least 2 is denied to leave their coalition and can, therefore, not perform a CNS deviation. On the other hand, by the third assertion, no agent in a singleton coalition can improve their utility by joining any other coalition. Hence, π_{i+1} is a CNS partition. This completes the proof of the claim. \triangleleft

The assertion of Theorem 4.2 follows from Claim 1 for the case $i = n$. \square

In particular, Theorem 4.2 applies to symmetric FEGs and AEGs. For the latter, we must deal with variable utility values that depend on the number of players. However, Theorem 4.2 applies to individual games, and each AEG satisfies the conditions of the theorem. By contrast, Theorem 4.2 breaks down if we additionally allow for the utility value of 0, even if the positive and negative utilities are restricted further. We defer this result to Section 4.2, where we get it as a byproduct during the consideration of CIS.

Example 4.3. We provide an example to show that the partitions produced by Algorithm 1 do not satisfy any of our solution concepts other than CNS and the weaker notion of CIS. Let $y \geq x > 0$. We define a symmetric $\{-y, x\}$ -ASHG (N, u) with agent set is $N = \{c_1, c_2, l_1, l_2, l_3, l_4\}$ where we display the agents according to their arrival order. The symmetric utilities are as illustrated in Figure 3, i.e., $u(a, b) = x$ for $(a, b) \in \{(c_1, c_2), (c_1, l_1), (c_1, l_2), (c_2, l_3), (c_2, l_4)\}$, and $u(a, b) = -y$, otherwise. Hence, the utility structure resembles a double-centered star with two centers c_1 and c_2 and four leaves l_1 through l_4 .

Clearly, Algorithm 1 forms the grand coalition consisting of all agents as, whenever an agent arrives, there exists a positive utility towards some agent that is already present. The resulting partition is contractually Nash-stable as every agent is denied to abandon this coalition. However, for each $i \in [4]$, l_i can perform an individual deviation to form a singleton coalition. Hence, the partition is not individually stable and, therefore, not

Nash-stable. Moreover, forming a singleton coalition is also a (weakly) blocking coalition and thus the partition is not in the (strict) core.

Additionally, the partition $\pi' = \{\{c_1, \ell_1, \ell_2\}, \{c_2, \ell_3, \ell_4\}\}$ Pareto-dominates the grand coalition. Indeed, for $i \in [2]$ and $j \in [4]$, it holds that $u_{c_i}(\pi') = 2x > 3x - 2y$ and $u_{\ell_j}(\pi') = x - y > x - 4y$. There, we use that $y \geq x > 0$.

Finally, consider the partition $\pi'' = \{\{c_1, c_2\}\} \cup \{\{\ell_j\} : j \in [4]\}$. We observe that the social welfare of the grand coalition is $10x - 20y < 0$, whereas $\mathcal{SW}(\pi'') = 2x > 0$. Hence, the partition produced by Algorithm 1 can have a negative social welfare in instances that admit a partition with positive social welfare. Thus, Algorithm 1 does not yield any approximation guarantee for welfare optimality. \triangleleft

To conclude this section, we show how to strengthen Proposition 4.1 for randomized algorithms. The idea is to construct a random instance where every deterministic algorithm performs poorly. We can then apply Yao's principle (Yao, 1977) to bound the performance of any randomized algorithm. For this, we create a random version of the game in Proposition 4.1, where every deterministic algorithm succeeds in computing a CNS partition with probability at most $1/2$. Modifying this instance by concatenating k copies of it implies that every deterministic algorithm succeeds with probability at most 2^{-k} .

Theorem 4.4. *Let ALG be any randomized online algorithm for symmetric AFGs. Then, it holds that $W_{\text{CNS}}(ALG) = 0$.*

Proof. Let $k \in \mathbb{N}$ be a positive integer. We define the random AFG $G = (N, u)$ where $N = \bigcup_{i \in [k]} N_i$ for $N_i = \{a_i, b_i, c_i\}$. The random utilities are given by $u(a_i, b_i) = -1$ and, with probability $1/2$ each, it holds that $u(a_i, c_i) = u(b_i, c_i) = 3k$ or $u(a_i, c_i) = u(b_i, c_i) = -1$. Note that $3k$ is the number of agents. All other utilities are set to -1 . The randomizations for the utilities within N_i and N_j for $1 \leq i < j \leq k$ are performed independently. Hence, we have defined a random AFG.

The agents arrive in the order (N_1, \dots, N_k) and within a set N_i in the order (a_i, b_i, c_i) . In other words, G is a uniformly random choice from a set of 2^k instances, each of which is a composition of k gadgets drawn independently from the same distribution. Each gadget is one of the two games considered in Proposition 4.1 with equal probability.

Now, let ALG be an arbitrary deterministic algorithm for AFGs and define $\pi := ALG(G)$. By the proof of Proposition 4.1, for every $i \in [k]$, ALG fails with probability at least $1/2$ on $G[N_i]$.

Moreover, by design of the random instance, if ALG computes a CNS partition on G , then, for all $i \in [k]$, $\pi[N_i]$ is CNS for $G[N_i]$. Indeed, if $u(a_i, c_i) = u(b_i, c_i) = -1$, then π is CNS only if all agents in N_i are in singleton coalitions, and hence $\pi[N_i]$ is CNS for $G[N_i]$. If $u(a_i, c_i) = u(b_i, c_i) = 3k$, then $a_i \in \pi(c_i)$ and $b_i \in \pi(c_i)$ as these agents would perform a CNS deviation to join c_i , otherwise. Hence, $\pi[N_i] = \{N_i\}$, which is CNS for $G[N_i]$.

Now, observe that, by the arrival sequence of the agents, the performance of ALG on N_i is at most as good as the performance of the best algorithm for $G[N_i]$. Therefore, using the independence of the random selection of the utilities, the probability that ALG computes a CNS partition on G is bounded by the product of the probabilities that $\pi[N_i]$ is CNS for $G[N_i]$. Hence, π is a CNS partition with probability at most 2^{-k} .

By Yao’s principle, no randomized algorithm can compute a CNS partition with probability more than 2^{-k} for every (deterministic) symmetric AFG. Since k was chosen arbitrarily, this proves the assertion. \square

4.2 Contractual Individual Stability and Pareto Optimality

Next, we consider contractual individual stability, which is a weaker solution concept than contractual Nash stability. It relies on a more constrained deviation concept, where an agent in the welcoming coalition can also veto a single-agent deviation. Algorithm 1 can be shown to compute CIS partitions, even if we allow for strict and symmetric ASHG as input. However, our following result achieves even more and shows the existence of an online algorithm for computing PO partitions in strict (and possibly nonsymmetric) ASHG. Recall that PO is a stronger notion than CIS. The presented algorithm is an online adaptation of serial dictatorships. This algorithmic approach is known to be successful in achieving Pareto optimality for offline ASHG (Aziz et al., 2013; Bullinger, 2020) and online fair division (Aleksandrov & Walsh, 2019).⁶ The idea is to assign a dictator to every created coalition, and these are asked in the order of their arrival whether they want newly arriving agents to be part of their coalition.

Theorem 4.5. *There exists a deterministic online algorithm, which always outputs a PO partition for strict ASHG.*

Proof. Consider a strict ASHG with agent set $N = \{a_i: 1 \leq i \leq n\}$ and arrival order (a_1, \dots, a_n) . Apply Algorithm 2 to compute a partition π . This algorithm proceeds as follows: Whenever a new agent arrives, we ask for each existing coalition whether the first agent in that coalition has a positive utility for the new agent. If such a coalition exists, the new agent joins the coalition among those that was created first. Otherwise, the algorithm starts a new coalition with the new agent.

For the proof, we use the notation from the algorithm and assume that $\pi = \{C_i: 1 \leq i \leq m\}$ for some $m > 0$. Assume further that these coalitions were formed in the order C_1, \dots, C_m and that agent f_j was the first agent in coalition C_j for all $j \in [m]$. The algorithm fulfills the property that, for all $j \in [m]$, and agents $x \in C_j \setminus \{f_j\}$, it holds that $u_{f_j}(x) > 0$. We refer to this property as observation (*).

We are ready to prove that π is Pareto-optimal. Assume that π' is any partition such that, for all agents $x \in N$, it holds that $u_x(\pi') \geq u_x(\pi)$. We claim that $\pi' = \pi$.

By observation (*) and the design of the algorithm, agent f_1 is in their best coalition in π and, by the strictness of the utilities, their best coalition is unique, so $\pi'(f_1) = \pi(f_1) = C_1$. We call this fact observation (**). We now prove our claim that $\pi' = \pi$ by induction over m , i.e., the number of coalitions in the partition π .

First, consider the case $m = 1$. Then, by observation (**), it holds that $\pi' = \{\pi'(f_1)\} = \{\pi(f_1)\} = \pi$.

Now, assume that $m > 1$. By observation (**), it suffices to show that $\pi' \setminus \{C_1\} = \pi \setminus \{C_1\}$. Consider the ASHG restricted to the agent set $N' = \bigcup_{2 \leq j \leq m} C_j$. Then, $\pi \setminus \{C_1\}$ is the output of Algorithm 2 of the restricted ASHG if the arrival order is the subsequence

6. Similar to the online fair division literature, our online serial dictatorship algorithm can be shown to have the additional desirable property of strategyproofness.

Algorithm 2 Pareto-optimal partition of online strict ASHG.

Input: Strict ASHG

Output: Pareto-optimal partition π

```

 $\pi \leftarrow \emptyset$ 
for  $i = 1, \dots, n$  do
  if  $\{j \in [|\pi|] : u_{f_j}(a_i) > 0\} \neq \emptyset$ 
    /* check if  $a_i$  is a friend of the first agent  $f_j$  in the  $j$ th existing
    coalition */
    then
       $j^* \leftarrow \min_{j \in [|\pi|]} \{u_{f_j}(a_i) > 0\}$ 
       $\pi \leftarrow \pi \setminus \{C_{j^*}\} \cup \{C_{j^*} \cup \{a_i\}\}$ 
    else
       $C_{|\pi|+1} \leftarrow \{a_i\}$ 
       $f_{|\pi|+1} \leftarrow a_i$ 
       $\pi \leftarrow \pi \cup \{C_{|\pi|+1}\}$ 
return  $\pi$ 

```

of the original arrival order. Moreover, by assumption, for all $x \in N \setminus C_1$, it holds that $u_x(\pi' \setminus \{C_1\}) \geq u_x(\pi \setminus \{C_1\})$. Hence, by induction, it holds that $\pi' \setminus \{C_1\} = \pi \setminus \{C_1\}$, as desired. This shows that $\pi' = \pi$.

Consequently, no partition exists that Pareto-dominates π , and, therefore, π is Pareto-optimal. \square

Since every Pareto-optimal partition is also a CIS partition, we obtain the following corollary.

Corollary 4.6. *There exists a deterministic online algorithm, which always outputs a CIS partition for strict ASHGs.*

However, no other of our solution concepts is satisfied by the partitions produced by Algorithm 2.

Example 4.7. We provide an example to show that the partitions produced by Algorithm 2 do not satisfy any of our solution concepts other than PO and CIS. Our game only uses any positive utility x and any negative utility $-y$, where $x, y > 0$, and the agent set depends on a positive integer $k \in \mathbb{N}$. Specifically, we define a symmetric $\{-y, x\}$ -ASHG (N, u) with agent set $N = \{a, c\} \cup \{b_i : i \in [k]\}$ and arrival order (a, b_1, \dots, b_k, c) . The symmetric utilities are as illustrated in Figure 3, i.e., $u(d, b_i) = x$ for $d \in \{a, c\}$ and $i \in [k]$. All other utilities are set to $-y$.

Clearly, Algorithm 2 produces the coalition structure $\pi = \{\{a, b_1, \dots, b_k\}, \{c\}\}$. For all $i \in [k]$, it holds that $u_{b_i}(\pi) = x - (k-1)y$. Thus, for large enough k , agent b_i would prefer to be in a coalition of their own. Hence, this constitutes an IS deviation as well as a (weakly) blocking coalition, and, therefore, π is no IS (and NS) partition, and not in the (strict) core.

Moreover, it holds that $u_c(\pi(a) \cup \{c\}) = u_c(N) = kx - y$. Thus, for large enough k , agent c can join $\pi(a)$ by a CNS deviation, and, therefore, π is no CNS partition either.

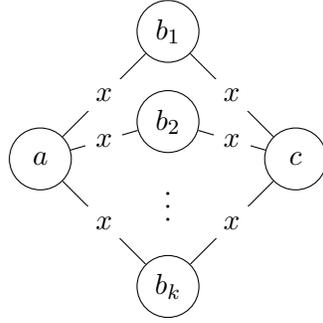


Figure 4: Instance defined in Example 4.7. Missing edges indicate a symmetric utility of $-y$. For arrival order (a, b_1, \dots, b_k, c) , Algorithm 2 returns $\pi = \{\{a, b_1, \dots, b_k\}, \{c\}\}$. For large enough k , this is neither individually stable, nor contractually Nash-stable, nor in the core.

Finally, we have that

$$SW(\pi) = 4kx - \left(2\binom{k}{2} + 2\right)y = 4kx - (k(k-1) + 2)y.$$

Hence, for large enough k , we have that $SW(\pi) < 0$. However, the welfare-optimal partition has positive welfare. Thus, Algorithm 2 does not yield any approximation guarantee for welfare optimality. \triangleleft

Of course, Theorem 4.5 and Corollary 4.6 work for subclasses of ASHG like FEGs, AFGs, and AEGs. Interestingly, however, the situation changes once we allow for a utility of 0. It becomes impossible to compute CIS partitions, and thus PO partitions, online, and this already holds for symmetric FENGs. This is a clear contrast to offline hedonic games, where PO (and CIS) partitions are guaranteed to exist without any restriction on the game.

Proposition 4.8. *There exists no deterministic online algorithm, which always outputs a CIS partition for symmetric FENGs.*

Proof. Assume for contradiction that *ALG* always outputs a CIS partition for symmetric FENGs. Consider the following two AFGs (N, u) and (N, \hat{u}) with identical agent set $N = \{a, b, c\}$ and symmetric utilities $u(a, b) = \hat{u}(a, b) = 0$, $u(a, c) = -1$, $u(b, c) = 1$, and $\hat{u}(a, c) = \hat{u}(b, c) = 1$. Consider the arrival order (a, b, c) . Before the arrival of c , *ALG* cannot distinguish, whether the game will be (N, u) or (N, \hat{u}) . Figure 5 depicts the situation.

If, at the arrival of b , *ALG* forms $\{a, b\}$, assume that we are in game (N, u) . Then, $\{\{a, b\}, \{c\}\}$ is not a CIS partition, because b has a CIS deviation to join agent c . However, forming $\{a, b, c\}$ does not lead to a CIS partition either, because then agent a has a CIS deviation to form a singleton coalition.

If, however, *ALG* forms two singleton coalitions for a and b , then consider (N, \hat{u}) . If *ALG* forms $\{a, c\}$ or $\{c\}$, then b has a CIS deviation to join this coalition. Finally, if *ALG* forms $\{b, c\}$, then a has a CIS deviation to join. \square

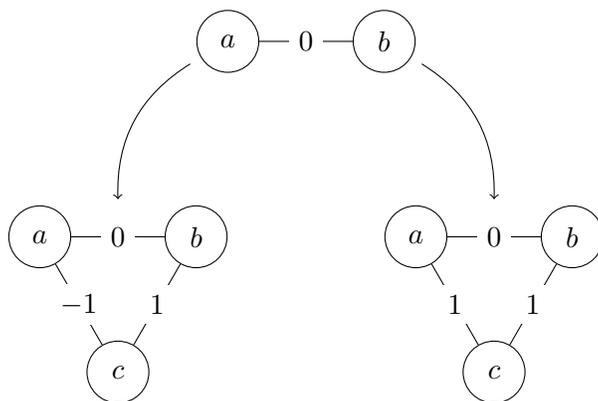


Figure 5: Adversarial FENGs for computing CIS partitions. Every deterministic algorithm fails for one of the two possible instances.

Similar to the previous section, we can extend this result to randomized algorithms.

Theorem 4.9. *Let ALG be any randomized online algorithm for symmetric FENGs. Then, it holds that $W_{CIS}(ALG) = 0$.*

Proof. Let $k \in \mathbb{N}$ be a positive integer. We define the random FENG $G = (N, u)$ where $N = \bigcup_{i \in [k]} N_i$ for $N_i = \{a_i, b_i, c_i\}$. The random utilities are given by $u(a_i, b_i) = 0$ and $u(a_i, c_i) = 1$, and we randomize uniformly between $u(b_i, c_i) = 1$ and $u(b_i, c_i) = -1$. All other utilities are set to 0. The randomization for the utilities within N_i and N_j for $1 \leq i < j \leq k$ are performed independently. The agents arrive in the order (N_1, \dots, N_k) and within N_i in the order (a_i, b_i, c_i) . Hence, for $i \in [k]$, $G[N_i]$ is one of the two FENGs from Proposition 4.8 selected by an unbiased coin flip.

Now, let ALG be an arbitrary deterministic algorithm for symmetric FENGs and define $\pi := ALG(G)$. By the proof of Proposition 4.8, for every $i \in [k]$, ALG fails with probability $1/2$ on $G[N_i]$.

Moreover, by design of the random instance, if ALG computes a CIS partition on G , then, for all $i \in [k]$, $\pi[N_i]$ is CIS for $G[N_i]$. Indeed, assume that π is CIS but there exists $i \in [k]$ such that $\pi[N_i]$ is not CIS for $G[N_i]$. Then, some agent $x \in N_i$ has a CIS deviation in $G^k[N_i]$ with respect to $\pi[N_i]$. However, since the utilities of x to all agents in $N \setminus N_i$ are 0, they permit x to leave or join their respective coalitions and do not influence the utility change of x . Therefore, π is not CIS, a contradiction.

Now, observe that, by the arrival sequence of the agents, the performance of ALG on N_i is at most as good as the performance of the best algorithm for $G[N_i]$. Therefore, using the independence of the random selection of the utilities, the probability that ALG computes a CIS partition on G is bounded by the product of the probabilities that $\pi[N_i]$ is CIS for $G[N_i]$. Hence, π is a CIS partition with probability at most 2^{-k} .

By Yao's principle, no randomized algorithm can compute a CIS partition with probability more than 2^{-k} for every (deterministic) symmetric FENG. Since k was chosen arbitrarily, this proves the assertion. \square

4.3 Individual Stability

As a last single-deviation stability concept, we consider individual stability, which entails contractual individual stability and is the complementary (but logically incomparable) notion to contractual Nash stability. In contrast to CNS, each agent in the welcoming (instead of abandoned) partition has the power to veto a single-agent deviation. We find that, even for the combination of symmetry and severe utility restrictions, online algorithms fail to be able to compute IS partitions. Note that reasonable classes of games contain at least one negative and one positive utility value. Otherwise, the partition consisting of the grand coalition containing all agents or the partition consisting of singleton coalitions is stable. By contrast, we show next that computing IS partitions becomes difficult when any positive and any negative utility is present.

Compared to the proofs of Theorems 4.4 and 4.9, simply concatenating identical games by negative utilities can be problematic for some utility values. For instance, if the positive utility is sufficiently large compared to the negative utility (e.g., in AFGs), then the grand coalition is IS if each agent has a positive utility for some other agent. Instead, we prove the statement by considering one large random adversarial instance for deterministic algorithms and apply Yao’s principle once again.

Theorem 4.10. *Let $x, y > 0$ and let ALG be any randomized online algorithm for symmetric $\{-y, x\}$ -ASHGs, symmetric AFGs, or symmetric AEGs. Then, $W_{IS}(ALG) = 0$.*

Proof. Let $x, y > 0$. We will define a random adversarial symmetric $\{-y, x\}$ -ASHG based on an integer $k \geq 2$ with $n = k^2 + 2$ agents. We will then show that any online algorithm ALG computes an IS partition in this game with probability at most $\frac{1}{k}$. This result holds independent of the specific values for x and y and we can, therefore, assume that these values depend on k (and, therefore, n). Hence, the construction works in particular for AFGs and AEGs.

We define the game $G = (N, u)$, which is illustrated in Figure 6. Let $A = \{a_i : 1 \leq i \leq k^2\}$ and $N = A \cup \{b, c\}$. Agents arrive in the order $(a_1, \dots, a_{k^2}, b, c)$. Independent of randomizations, it always holds that $u(b, c) = x$ and $u(a_i, a_j) = -y$ for all $1 \leq i < j \leq k^2$. The remaining utilities are selected at random as follows. First, we uniformly draw a random subset $B \subseteq A$ of k agents. Second, one agent is selected from B uniformly at random and labeled d . We set $u(c, d) = x$, $u(c, a) = -y$ for all $a \in A \setminus \{d\}$, $u(b, a) = x$ for all $a \in B$, and $u(b, a) = -y$ for all $a \in A \setminus B$. Therefore, G is a uniformly random choice from a set of $\binom{k^2}{k}k$ instances.

Before we bound the probability of a deterministic algorithm for forming an IS partition, we determine the IS partitions in the obtained instance dependent on x and y .

Claim 2. *Let $\emptyset \neq S \subseteq B \setminus \{d\}$. The following are individually stable partitions in G :*

- $\{\{b, c, d\}\} \cup \{\{a_i\} : a_i \in A \setminus \{d\}\}$ for all $x, y > 0$,
- $\{\{c, d\}, \{b\} \cup S\} \cup \{\{a_i\} : a_i \in A \setminus (\{d\} \cup S)\}$ if $2 \leq |S| \leq \frac{x}{y} + 1$, and
- $\{\{b, c, d\} \cup S\} \cup \{\{a_i\} : a_i \in A \setminus (\{d\} \cup S)\}$ if $|S| \leq \frac{x}{y} - 1$

Moreover, no other partition is individually stable.

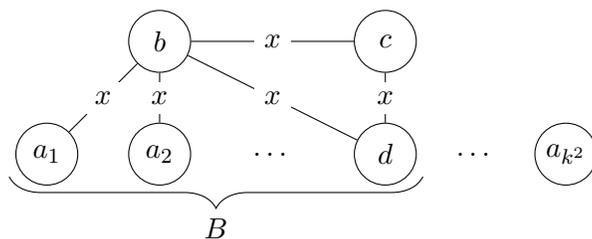


Figure 6: Adversarial instance for achieving individual stability in $\{-y, x\}$ -ASHGs for $x, y > 0$. We only depict the positive utilities of x . All remaining utilities are $-y$.

Proof. Let π be an IS partition for G . Independently of x and y , it has to hold that all agents in $A \setminus B$ are in singleton coalitions, and all agents in B are either in a coalition with b , in a coalition with c , or in a singleton coalition as well. We disregard the agents in $A \setminus B$ from consideration for the rest of the proof. Another important observation is that there exists only one pair of agents with positive utility that does not include player b . We base the proof on a case distinction depending on $\pi(b)$.

First, it holds that $\pi(b) \neq \{b\}$, as otherwise every agent in $B \setminus \{d\}$ has an IS deviation to join $\{b\}$, a contradiction. Second, it holds that $\pi(b) \neq \{b, c\}$ and $\pi(b) \neq \{b, d\}$, as otherwise agent d or c respectively have an IS deviation to join $\pi(b)$.

Now, assume that there exists $a \in B \setminus \{d\}$ such that $\pi(b) = \{a, b\}$. Then, $\{c, d\} \notin \pi$, as otherwise agent b has an IS deviation to join $\{c, d\}$. Hence, c and d must be in singleton coalitions as they have IS deviations to leave any other coalition to form a singleton coalition. However, then, they both have the IS deviation to join each other, which is a contradiction. Hence, $\pi(b) \neq \{a, b\}$.

Next, assume that there exists $\emptyset \neq S \subseteq B \setminus \{d\}$ such that $\pi(b) = S \cup \{b, c\}$ or $\pi(b) = S \cup \{b, d\}$. Then, the agent among $\{c, d\}$ not in $\pi(b)$ has to be in a singleton coalition as they could otherwise form a singleton coalition by an IS deviation. But then the agent among $\{c, d\}$ contained in $\pi(b)$ can perform an IS deviation to join this singleton coalition. Hence, this case is also excluded.

So far, we have excluded several cases in which no IS partition is possible. We conclude that b must be in a coalition of size at least 3 and that the coalition of b must contain either both c and d or none of them. In the remaining cases, we find some IS partitions.

First, consider the case where $\pi(b) = \{b, c, d\}$. Then $\pi(a) = \{a\}$ for all $a \in A$, and the resulting partition is an IS partition. This proves that the first partition of the claim is an IS partition.

Next, assume that there exists $\emptyset \neq S \subseteq B \setminus \{d\}$ such that $\pi(b) = S \cup \{b\}$. Then, all agents $a \in A \setminus (S \cup \{d\})$ must be in singleton coalitions, as otherwise they have an IS deviation to form a singleton coalition. This only leaves agents c and d , and we conclude that $\{c, d\} \in \pi$, as otherwise, they have an IS deviation to join each other. Finally, the partition is only IS if $|S| \leq \frac{x}{y} + 1$. Otherwise, all agents $a \in S$ have a utility of $x - (|S| - 1)y < 0$ and thus can perform an IS deviation to form a singleton coalition. In addition, we need $|S| \geq 2$ as otherwise b performs a deviation to join $\{c, d\}$.

Moreover, for $2 \leq |S| \leq \frac{x}{y} + 1$, the partition $\{\{c, d\}, \{b\} \cup S\} \cup \{\{a\} : a \in A \setminus (\{d\} \cup S)\}$ can be shown to be individually stable: Clearly, none of the agents in A in singleton coalitions can enter $\pi(b)$ because this would be blocked by agents in S . Similarly, they are also blocked to enter other singleton coalitions. Agents b and c are also blocked to join any other coalition. Next, agents in S cannot improve by performing any deviation. Finally, because $|S| \geq 2$, b cannot improve by joining any other coalition.

Together, for $\emptyset \neq S \subseteq B \setminus \{d\}$, the partition $\{\{c, d\}, \{b\} \cup S\} \cup \{\{a\} : a \in A \setminus (\{d\} \cup S)\}$ is an IS partition if and only if $2 \leq |S| \leq \frac{x}{y} + 1$.

Finally, assume that there exists $\emptyset \neq S \subseteq B \setminus \{d\}$ such that $\pi(b) = S \cup \{b, c, d\}$. Then, as before, all agents $a \in A \setminus (S \cup \{d\})$ must be in singleton coalitions. Moreover, the partition is individually stable only if $|S| \leq \frac{x}{y} - 1$. Otherwise, all agents $a \in S$ have a utility of $x - (|S| + 1) \cdot y < 0$ and can thus perform an IS deviation to form a singleton coalition. Agents b, c , and d all have two friends and thus no IS deviation when $|S| \leq \frac{x}{y} - 1$. This proves that for $\emptyset \neq S \subseteq B \setminus \{d\}$, the partition $\{\{b, c, d\} \cup S\} \cup \{\{a\} : a \in A \setminus (\{d\} \cup S)\}$ is individually stable if and only if $|S| \leq \frac{x}{y} - 1$.

As this case distinction covers all possible partitions, we have found all IS partitions as stated in the assertion. \triangleleft

Now, let ALG be any deterministic online algorithm and define $\pi := ALG(G)$. To conclude the proof, we show that π is an IS partition with probability at most $\frac{1}{k}$.

Note that $G[A]$ is identical independent of the randomization for the instance. We perform a case distinction based on $\pi[A]$. Intuitively, ALG can either attempt to reach the IS partition $\{\{b, c, d\}\} \cup \{\{a\} : a \in A \setminus \{d\}\}$ by forming all singleton coalitions, or it forms a single coalition of size greater than one to reach any of the other IS partitions. In all other cases, the algorithm can no longer reach an IS partition.

Assume first that $\pi[A]$ contains exactly one coalition of size strictly larger than one. Note that then $\pi \neq \{\{b, c, d\}\} \cup \{\{a\} : a \in A \setminus \{d\}\}$ and it can only create an IS partition for the latter two cases of Claim 2. Let $S \subseteq A$ be the coalition of size strictly larger than 1 in $\pi[A]$. Then, π can only be individually stable if $S \subseteq B$ for the random set B , particularly $|S| \leq k$. As B is chosen uniformly at random from A , the probability that $S \subseteq B$ is $\frac{\binom{k}{|S|}}{\binom{k^2}{|S|}}$.

This is true because there are $\binom{k^2}{|S|}$ choices to select S and $\binom{k}{|S|}$ of these choices result in $S \subseteq B$. We compute

$$\begin{aligned} \frac{\binom{k}{|S|}}{\binom{k^2}{|S|}} &= \frac{k!|S|!(k^2 - |S|)!}{k^2!|S|!(k - |S|)!} = \frac{k(k-1) \cdots (k - |S| + 1)}{k^2(k^2 - 1) \cdots (k^2 - |S| + 1)} \\ &= \frac{1}{k} \frac{k-1}{k^2 - 1} \cdots \frac{k - |S| + 1}{k^2 - |S| + 1} < \frac{1}{k}. \end{aligned}$$

Therefore, ALG successfully forms an IS partition with probability at most $\frac{1}{k}$ in this case.

Next, assume that $\pi[A]$ contains only singleton coalitions. By Claim 2, ALG only outputs an IS partition if $\pi = \{\{b, c, d\}\} \cup \{\{a\} : a \in A \setminus \{d\}\}$. Therefore, when b arrives, the set B is revealed to the algorithm, and it needs to match b to d , as otherwise $\{b, c, d\}$

cannot be formed. The probability for this event is precisely $\frac{1}{k}$ since each element in B is d with equal probability. Hence, also in this case, π is IS with probability at most $\frac{1}{k}$.

Together, π is an IS partition with probability at most $\frac{1}{k}$. By Yao's principle, no randomized algorithm can compute an IS partition with probability more than $\frac{1}{k}$ for every (deterministic) symmetric $\{-y, x\}$ -ASHGs. Since our choice of k is arbitrary, the assertion follows. \square

4.4 Group Stability

Finally, we consider group stability, that is, computing partitions in the core and strict core. Since partitions in the strict core are individually stable, Theorem 4.10 already suggests computational difficulties for achieving partitions in the strict core. There is, however, a caveat. For some parameters of x and y , the instances considered in Theorem 4.10 have an empty strict core. Hence, an online algorithm for these instances is not worse than the best offline algorithm (see also Footnote 5). In this section, we, therefore, restrict attention to instances containing partitions in the strict core. Note that this assumption is, for instance, trivially true for AFGs, in which partitions in the strict core are guaranteed to exist (Dimitrov et al., 2006).

We now prove a negative result encompassing both the core and the strict core. The proof is similar to the proof of Theorem 4.10, but a suitable set of agents arrives instead of the single agent c . In addition, we can simplify the instances from Theorem 4.10 because we can omit agents in $A \setminus B$.

Theorem 4.11. *Let $x, y > 0$ and ALG be any randomized online algorithm for symmetric $\{-y, x\}$ -ASHGs (or symmetric AEGs) that contain partitions in the strict core. Then, it holds that $W_{\text{CR}}(\text{ALG}) = 0$ and $W_{\text{SCR}}(\text{ALG}) = 0$.*

Proof. Let $x, y > 0$ and define $q := \left\lceil \frac{x}{y} \right\rceil + 1$. Note that q is a fixed parameter of our construction that only depends on x and y .

We will define a random adversarial symmetric $\{-y, x\}$ -ASHG $G = (N, u)$ based on q as well as an integer $k \geq 2$ with $n = k + q + 1$ agents. Figure 7 illustrates the game. Let $A = \{a_i : 1 \leq i \leq k\}$, $C = \{c_i : 1 \leq i \leq q\}$ and $N = A \cup C \cup \{b\}$, where b is an additional agent. Agents arrive in the order $(a_1, \dots, a_k, b, c_1, \dots, c_q)$. We define

- for all $i \in [q]$, $u(b, c_i) = x$,
- for all $i \in [k]$, $u(b, a_i) = x$,
- for all $i, j \in [q]$ with $i \neq j$, $u(c_i, c_j) = x$, and
- for all $i, j \in [k]$ with $i \neq j$, $u(a_i, a_j) = -y$.

The remaining utilities are selected at random as follows. We uniformly draw an index $j^* \in [k]$. Then, for all $i \in [q]$ and $j \in [k]$ with $j \neq j^*$, we set $u(c_i, a_{j^*}) = x$ and $u(c_i, a_j) = -y$, i.e., agent a_{j^*} has positive utility for all agents in C while all other agents in A have negative utilities. Therefore, G is a uniformly random choice from a set of k instances.

First, we ensure that G always contains partitions in the strict core. Define the partition $\pi^* = \{C \cup \{a_{j^*}, b\}\} \cup \{\{a_j\} : 1 \leq j \leq k, j \neq j^*\}$.

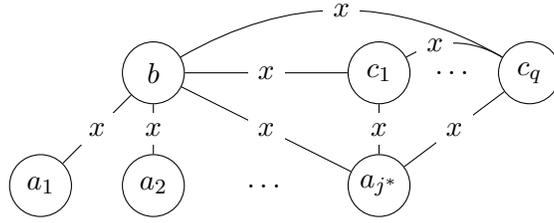


Figure 7: Adversarial instance for achieving partitions in the (strict) core in $\{-y, x\}$ -ASHGs for $x, y > 0$. All such instances have a nonempty strict core. We only depict the positive utilities of x . All remaining utilities are $-y$.

Claim 3. *The partition π^* is in the strict core.*

Proof. First, note that an agent that is part of their unique best coalition can never be part of a weakly blocking coalition. Hence, since this is the case for agents in $C \cup \{a_{j^*}\}$, we only have to exclude weakly blocking coalitions containing b and agents in $A \setminus \{a_{j^*}\}$. Let $D \subseteq (A \setminus \{a_{j^*}\}) \cup \{b\}$ with $b \in D$ and assume that $u_b(D) \geq u_b(\pi^*)$. Then, $|D \cap (A \setminus \{a_{j^*}\})| \geq q + 1$, otherwise coalition D lowers agent b 's utility compared to $\pi^*(b)$. Hence, for $d \in D \setminus \{b\}$, it holds that $u_d(D) = x - (|D| - 2)y < x - qy < 0$, where we use that $q > \frac{x}{y}$. It follows that D is not a weakly blocking coalition, and therefore, b is not part of a weakly blocking coalition. However, the agents in $A \setminus \{a_{j^*}\}$ all have a negative utility for each other, and therefore, they cannot form a weakly blocking coalition. Hence, there is no weakly blocking coalition, and π^* is in the strict core. \triangleleft

The previous claim shows that our considered instances have a suitable form, i.e., they contain elements in the strict core. To continue the proof, we show that π^* is the only partition in the strict core and even in the core.

Claim 4. *The partition π^* is the unique partition in the core.*

Proof. By Claim 3, we already know that π^* is in the strict core and, therefore, in the core. It remains to show that the core does not contain any other partition. Let π be a partition in the core. We will show that $\pi = \pi^*$.

First, we will prove that there exists a coalition $D \in \pi$ with $C \cup \{a_{j^*}\} \subseteq D$. To prove this, observe that the coalition $C \cup \{a_{j^*}\}$ yields a utility of qx to all its members. Hence, because π does not admit a blocking coalition, there exists an agent $d \in C \cup \{a_{j^*}\}$ with $u_d(\pi) \geq qx$. This can only happen if the coalition of d contains at least q agents for which they receive a positive utility. Hence if $b \notin \pi(d)$, then $C \cup \{a_{j^*}\} \subseteq \pi(d)$ and our assertion is true. Moreover, if there exists an agent $a \in A \setminus \{a_{j^*}\}$ with $a \in \pi(d)$, then $u_d(\pi) \geq qx$ is only possible if all $q + 1$ agents for which d achieves a positive utility are in π . Then, once again $C \cup \{a_{j^*}\} \subseteq \pi(d)$.

Together, it remains to consider the case where $b \in \pi(d)$ and $\pi(d) \subseteq C \cup \{a_{j^*}, b\}$. We are done if $\pi(d) = C \cup \{a_{j^*}, b\}$. Otherwise, because $|\pi(d)| \geq q + 1$, there is a unique agent $d' \in (C \cup \{a_{j^*}\}) \setminus \pi(d)$. However, forming $C \cup \{a_{j^*}, b\}$ is preferred by all agents in $\pi(d)$ as

well as by d' . Hence, this is a blocking coalition, contradicting the fact that π is in the core. Thus, it must hold that $\pi(d) = C \cup \{a_{j^*}, b\}$.

Now, consider the coalition $D \in \pi$ with $C \cup \{a_{j^*}\} \subseteq D$. Then, for all $1 \leq j \leq k$ with $j \neq j^*$, it holds that $a_j \notin D$. Otherwise, $u_{a_j}(\pi) \leq x - (q+1)y < 0$ and $\{a_j\}$ would be a blocking coalition.

Next, assume for contradiction that $b \notin D$. Note that all members in D prefer $D \cup \{b\}$. Hence, for this not to be a blocking coalition, it must hold that $u_b(\pi) \geq x(q+1)$. Therefore, $|\pi(b) \cap (A \setminus \{a_{j^*}\})| \geq q+1$. But then, for $a \in \pi(b) \setminus \{b\}$, it holds that $u_a(\pi) \leq x - qy < 0$. This is a contradiction. Hence, it must hold that $b \in D$. Taken together, we conclude that $D = C \cup \{a_{j^*}, b\}$.

For the remaining agents, i.e., for agents in $A \setminus \{a_{j^*}\}$, every nonempty coalition among themselves yields a negative utility. Hence, these agents have to form singleton coalitions in π . We conclude that $\pi^* = \pi$. \triangleleft

Now, let ALG be any deterministic online algorithm and define $\pi := ALG(G)$. To conclude the proof, we show that π is in the core (and, therefore, in the strict core) with probability at most $\frac{1}{k}$.

Recall that the first k agents to arrive are the agents in A . Since π^* is the only coalition in the core, we can restrict attention to the case where ALG assigns all agents in A to singleton coalitions. If ALG forms a singleton coalition when b arrives, then $\pi \neq \pi^*$ and π is not in the core. Assume that ALG forms a coalition of b with an agent from A . Then, $\pi = \pi^*$ is only possible if b forms a coalition with a_{j^*} . The probability for this event is exactly $\frac{1}{k}$ since $j^* = j$ for all $j \in [k]$ with equal probability. Hence, π is in the core with probability at most $\frac{1}{k}$.

By Yao's principle, no randomized algorithm can compute a partition in the core (and, therefore, in the strict core) with probability more than $\frac{1}{k}$ for every (deterministic) symmetric $\{-y, x\}$ -ASHGs containing a partition in the strict core. Since our choice of k is arbitrary, the assertion follows.

Finally, we observe that our construction also works for AEGs. These are $\{-y, x\}$ -ASHGs, where $x = 1$ and y depends on the number of agents. We can then simply set $q = 1$. Since all our games contain at least 2 agents, we then have that $q > \frac{1}{n}$, i.e., our construction is valid for $x = 1$ and $y = n$. \square

Note that the construction in the previous proof does not work for AFGs because of the dependence of q on x . Indeed, we cannot simultaneously satisfy $q > \frac{n}{1}$ and $n = k + q + 1$ when $k \geq 2$. Instead, we provide a different construction to obtain a result for AFGs analogous to Theorem 4.11.

Theorem 4.12. *Let ALG be any randomized online algorithm for symmetric AFGs. Then, it holds that $W_{CR}(ALG) = 0$ and $W_{SCR}(ALG) = 0$.*

Proof. Let $k \geq 3$. We will define a random adversarial symmetric AFG $G = (N, u)$ with $n = k + 1$ agents. Figure 8 illustrates the game. Let $A = \{a_i : 1 \leq i \leq k\}$ and define $N = A \cup \{b\}$, where b is an additional agent. Agents arrive in the order (a_1, \dots, a_k, b) . For all $i, j \in [k]$ with $i \neq j$, we define $u(a_i, a_j) = -1$.

The remaining utilities are selected at random as follows. We draw an index set $J \subseteq [k]$ with $|J| = 3$ uniformly at random, say $J = \{j_1, j_2, j_3\}$. We set $u(b, a_j) = n$ if $j \in J$ and

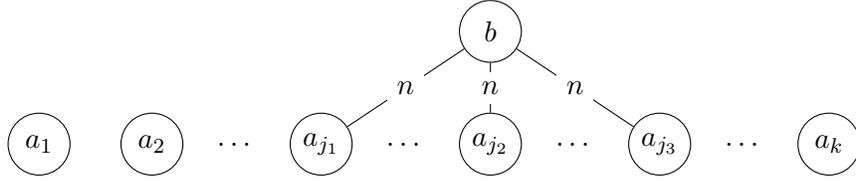


Figure 8: Adversarial instance for achieving partitions in the (strict) core in AFGs. We only depict the positive utilities of n . All remaining utilities are -1 .

$u(b, a_j) = -1$ if $j \in [k] \setminus J$. Hence, there are exactly three positive utilities from b to a random subset of agents in A and all other utilities are negative. Therefore, G is a uniformly random choice from a set of $\binom{k}{3}$ instances.

We start by determining the partitions in the core in the strict core.

Claim 5. *The following statements are true.*

1. *The partition $\{\{b\} \cup \{a_j : j \in J\}\} \cup \{\{a_j\} : j \in [k] \setminus J\}$ is in the strict core.*
2. *For $J' \subseteq J$ with $|J'| = 2$, the partition $\{\{b\} \cup \{a_j : j \in J'\}\} \cup \{\{a_j\} : j \in [k] \setminus J'\}$ is in the core.⁷*
3. *No other partition is in the core.*

Proof. Let π be a partition in the core. First note that for $j \in [k] \setminus J$, all coalitions except a singleton coalition yield a negative utility for a_j . Hence, $\{a_j\} \in \pi$.

Now, let $C = \pi(b)$ and define $F := \{a_j : j \in J\}$. We already know that $C \subseteq \{b\} \cup F$. If $C = \{b\}$, then $\{b\} \cup F$ is a blocking coalition. Next, assume that $|C \cap F| = 1$, say $C = \{b, a_{j_1}\}$. Then, $\{b, a_{j_2}, a_{j_3}\}$ is a blocking coalition. Hence π must be of the form described in the first or second case of the claim, which proves the third assertion.

If $C = \{b\} \cup F$, then b is in their unique most preferred coalition, and, therefore, not part of a weakly blocking coalition. However, no coalition consisting only of agents in A can be a weakly blocking coalition. Hence, π is in the strict core, proving the first assertion.

Finally, if $|C \cap F| = 2$, then the only coalition that is better for b is $\{b\} \cup F$, which is worse for the other agents in C . Similar as before, no coalition consisting only of agents in A is a (weakly) blocking coalition. Hence, π is in the core, proving the second assertion. \triangleleft

Now, let ALG be any deterministic online algorithm for symmetric AFGs and define $\pi := ALG(G)$. We claim that π is in the core (and, therefore, in the strict core) with probability at most $\frac{6}{k(k-1)}$.

Recall that the first k agents to arrive are the agents in A . By Claim 5, π can only be in the core if ALG forms a coalition C of size 2 or size 3 among the agents in A . Even more, π can only be in the core if $C \subseteq F$. If $|C| = 2$, then $C \subseteq F$ occurs with probability $\frac{\binom{3}{2}}{\binom{k}{2}} = \frac{6}{k(k-1)}$. If $|C| = 3$, then $C \subseteq F$ occurs with probability $\frac{1}{\binom{k}{3}} = \frac{6}{k(k-1)(k-2)} \leq \frac{6}{k(k-1)}$. Hence, π is in the core (and, therefore, in the strict core) with probability at most $\frac{6}{k(k-1)}$.

7. We remark that these partitions are not in the strict core as $B = \{b, a_j, a_\ell\}$ for $j \in J'$ and $\ell \in J \setminus J'$ is a weakly blocking coalition.

By Yao’s principle, no randomized algorithm can compute a partition in the core (and, therefore, in the strict core) with probability more than $\frac{6}{k(k-1)}$ for every (deterministic) symmetric AFG. Since our choice of k is arbitrary, the assertion follows. \square

5. Conclusion

In this paper, we have studied stability in online coalition formation. We have considered stability notions based on single-agent deviations and group deviations as well as Pareto optimality. Table 1 in Section 1 provides an overview of our results.

Our positive results follow from two deterministic algorithms. The first one outputs CNS partitions for symmetric games with utility restrictions that include FEGs and AEGs. The second one applies to strict ASHG and outputs PO partitions. The latter is interesting because Pareto optimality has the flavor of both stability and optimality. However, except for additionally achieving the weaker solution concept of contractual individual stability, the algorithms do not guarantee to satisfy any other of our solution concepts. Moreover, both may output partitions of negative social welfare for instances in which the maximum social welfare is positive, and, therefore, yield no approximation guarantee for maximizing social welfare, which is known to be a challenging objective (Flammini et al., 2021; Bullinger & Romén, 2023).

Beyond our algorithmic possibilities, we obtain negative results in the sense that no randomized algorithm can guarantee any fixed probability to output a stable partition. Surprisingly, our negative results even encompass concepts like contractual individual stability and Pareto optimality, for which solutions exist in every hedonic game. Hence, the online capabilities of algorithms can be severely weaker than strong offline possibilities. Moreover, negative results naturally extend to stronger solution concepts. For instance, a consequence of our negative results for IS and CNS is that the NS guarantee is 0 for all considered game restrictions.

We believe that moving beyond the mere consideration of welfare maximality in online coalition formation is an important step. There is plenty of room for future research in this direction. Possible directions include considering other solution concepts, such as envy-freeness or popularity (Aziz et al., 2013; Brandt & Bullinger, 2022). Another avenue for future research would be to consider game classes different from ASHG, for example, fractional hedonic games (for which only welfare maximization has been studied by Flammini et al. (2021)) or modified fractional hedonic games (Aziz et al., 2019; Olsen, 2012). Finally, especially since stability concepts are ordinal concepts (i.e., whether a deviation is possible only depends on the comparison of coalitions rather than exact utilities), it might be interesting to consider ordinal classes of hedonic games, such as anonymous hedonic games or hedonic diversity games (Bogomolnaia & Jackson, 2002; Bredereck et al., 2019).

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/11-2 and BR 2312/12-1, and by the AI Programme of The Alan Turing Institute. A preliminary version of this article appeared in the Proceedings of the 38th AAI Conference.

ence on Artificial Intelligence (February 2024). We thank Saar Cohen and the anonymous reviewers from AAAI and JAIR for their helpful comments.

References

- Alcalde, J., & Revilla, P. (2004). Researching with whom? Stability and manipulation. *Journal of Mathematical Economics*, 40(8), 869–887.
- Aleksandrov, M., & Walsh, T. (2019). Strategy-proofness, envy-freeness and Pareto efficiency in online fair division with additive utilities. In *Proceedings of the 16th Pacific Rim International Conference on Artificial Intelligence (PRICAI)*, pp. 527–541. Springer.
- Aziz, H., & Brandl, F. (2012). Existence of stability in hedonic coalition formation games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 763–770.
- Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., & Peters, D. (2019). Fractional hedonic games. *ACM Transactions on Economics and Computation*, 7(2), 1–29.
- Aziz, H., Brandt, F., & Seedig, H. G. (2013). Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, 195, 316–334.
- Aziz, H., & Savani, R. (2016). Hedonic games. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. D. (Eds.), *Handbook of Computational Social Choice*, chap. 15. Cambridge University Press.
- Banerjee, S., Konishi, H., & Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice and Welfare*, 18, 135–153.
- Benadè, G., & Sahoo, N. (2023). Stability, fairness and the pursuit of happiness in recommender systems. Tech. rep. 4241170, SSRN.
- Bilò, V., Fanelli, A., Flammini, M., Monaco, G., & Moscardelli, L. (2018). Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. *Journal of Artificial Intelligence Research*, 62, 315–371.
- Bilò, V., Monaco, G., & Moscardelli, L. (2022). Hedonic games with fixed-size coalitions. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 9287–9295.
- Boehmer, N., Bullinger, M., & Kerkmann, A. M. (2025). Causes of stability in dynamic coalition formation. *ACM Transactions on Economics and Computation*. Forthcoming.
- Bogomolnaia, A., & Jackson, M. O. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2), 201–230.
- Brandt, F., & Bullinger, M. (2022). Finding and recognizing popular coalition structures. *Journal of Artificial Intelligence Research*, 74, 569–626.
- Brandt, F., Bullinger, M., & Tappe, L. (2024). Stability based on single-agent deviations in additively separable hedonic games. *Artificial Intelligence*, 334, 104160.

- Brandt, F., Bullinger, M., & Wilczynski, A. (2023). Reaching individually stable coalition structures. *ACM Transactions on Economics and Computation*, 11(1–2), 4:1–65.
- Bredereck, R., Elkind, E., & Igarashi, A. (2019). Hedonic diversity games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 565–573.
- Bullinger, M. (2020). Pareto-optimality in cardinal hedonic games. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 213–221.
- Bullinger, M., Chatziafratis, V., & Shahkar, P. (2025). Welfare approximation in additively separable hedonic games. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Forthcoming.
- Bullinger, M., Elkind, E., & Rothe, J. (2024). Cooperative game theory. In Rothe, J. (Ed.), *Economics and Computation: An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, chap. 3, pp. 139–229. Springer.
- Bullinger, M., & Kraicz, S. (2024). Stability in random hedonic games. In *Proceedings of the 25th ACM Conference on Economics and Computation (ACM-EC)*, p. 212.
- Bullinger, M., & Romen, R. (2023). Online coalition formation under random arrival or coalition dissolution. In *Proceedings of the 31st European Symposium on Algorithms (ESA)*, pp. 27:1–27:18.
- Bullinger, M., & Suksompong, W. (2024). Topological distance games. *Theoretical Computer Science*, 981, 114238.
- Carosi, R., Monaco, G., & Moscardelli, L. (2019). Local core stability in simple symmetric fractional hedonic games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 574–582.
- Cechlárová, K., & Romero-Medina, A. (2001). Stability in coalition formation games. *International Journal of Game Theory*, 29, 487–494.
- Dimitrov, D., Borm, P., Hendrickx, R., & Sung, S. C. (2006). Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, 26(2), 421–433.
- Dimitrov, D., & Sung, S. C. (2007). On top responsiveness and strict core stability. *Journal of Mathematical Economics*, 43(2), 130–134.
- Doval, L. (2022). Dynamically stable matching. *Theoretical Economics*, 17(2), 687–724.
- Drèze, J. H., & Greenberg, J. (1980). Hedonic coalitions: Optimality and stability. *Econometrica*, 48(4), 987–1003.
- Elkind, E., Fanelli, A., & Flammini, M. (2020). Price of pareto optimality in hedonic games. *Artificial Intelligence*, 288, 103357.
- Ezra, T., Feldman, M., Gravin, N., & Tang, Z. G. (2022). General graphs are easier than bipartite graphs: Tight bounds for secretary matching. In *Proceedings of the 22nd ACM Conference on Economics and Computation (ACM-EC)*, pp. 1148 – 1177.
- Feldman, J., Korula, N., Mirrokni, V., Muthukrishnan, S., & Pál, M. (2009). Online ad assignment with free disposal. In *Proceedings of the 5th International Conference on Web and Internet Economics (WINE)*, pp. 374–385.

- Flammini, M., Kodric, B., & Varricchio, G. (2022). Strategyproof mechanisms for friends and enemies games. *Artificial Intelligence*, *302*, 103610.
- Flammini, M., Monaco, G., Moscardelli, L., Shalom, M., & Zaks, S. (2021). On the online coalition structure generation problem. *Journal of Artificial Intelligence Research*, *72*, 1215–1250.
- Gairing, M., & Savani, R. (2019). Computing stable outcomes in symmetric additively separable hedonic games. *Mathematics of Operations Research*, *44*(3), 1101–1121.
- Gajulapalli, K., Liu, J., Mai, T., & Vazirani, V. V. (2020). Stability-preserving, time-efficient mechanisms for school choice in two rounds. In *Proceedings of the 40th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*.
- Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, *69*(1), 9–15.
- Huang, Z., Kang, N., Tang, Z. G., Wu, X., Zhang, Y., & Zhu, X. (2018). How to match when all vertices arrive online. In *Proceedings of the 50th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 17–29.
- Huang, Z., & Tröbst, T. (2023). Online matching. In Echenique, F., Immorlica, N., & Vazirani, V. V. (Eds.), *Online and Matching-Based Market Design*. Cambridge University Press.
- Karp, R. M., Vazirani, U. V., & Vazirani, V. V. (1990). An optimal algorithm for on-line bipartite matching. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (STOC)*, pp. 352–358.
- Kerkmann, A. M., Nguyen, N.-T., Rey, A., Rey, L., Rothe, J., Schend, L., & Wiechers, A. (2022). Altruistic hedonic games. *Journal of Artificial Intelligence Research*, *75*, 129–169.
- Morrill, T. (2010). The roommates problem revisited. *Journal of Economic Theory*, *145*(5), 1739–1756.
- Olsen, M. (2009). Nash stability in additively separable hedonic games and community structures. *Theory of Computing Systems*, *45*, 917–925.
- Olsen, M. (2012). On defining and computing communities. In *Proceedings of the 18th Computing: The Australasian Theory Symposium (CATS)*, Vol. 128 of *Conferences in Research and Practice in Information Technology (CRPIT)*, pp. 97–102.
- Ota, K., Barrot, N., Ismaili, A., Sakurai, Y., & Yokoo, M. (2017). Core stability in hedonic games among friends and enemies: Impact of neutrals. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 359–365.
- Pavone, M., Saberi, A., Schiffe, M., & Tsao, M. W. (2022). Online hypergraph matching with delay. *Operations Research*, *70*(4), 2194 – 2212.
- Peters, D. (2017). Precise complexity of the core in dichotomous and additive hedonic games. In *Proceedings of the 5th International Conference on Algorithmic Decision Theory (ADT)*, Lecture Notes in Computer Science (LNCS), pp. 214–227. Springer-Verlag.

- Sung, S. C., & Dimitrov, D. (2010). Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3), 635–639.
- Wang, Y., & Wong, S. C. (2015). Two-sided online bipartite matching and vertex cover: Beating the greedy algorithm. In *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming (ICALP)*, pp. 1070–1081.
- Woeginger, G. J. (2013). A hardness result for core stability in additive hedonic games. *Mathematical Social Sciences*, 65(2), 101–104.
- Yao, A. C.-C. (1977). Probabilistic computations: Toward a unified measure of complexity. In *Proceedings of the 18th Symposium on Foundations of Computer Science (FOCS)*, pp. 222–227. IEEE Computer Society Press.