Control in Hedonic Games

Jiehua Chen
TU Wien
Vienna, Austria
jiehua.chen@ac.tuwien.ac.at

Jakob Guttmann
TU Wien
Vienna, Austria
e11810289@student.tuwien.ac.at

Merisa Mustajbašić
TU Wien
Vienna, Austria
e12450330@student.tuwien.ac.at

Sofia Simola
TU Wien
Vienna, Austria
ssimola@ac.tuwien.ac.at

ABSTRACT

We initiate the study of control in hedonic games, where an external actor influences coalition formation by adding or deleting agents. We consider three basic control goals (1) enforcing that an agent is *not alone* (NA); (2) enforcing that a *pair* of agents is in the same coalition (PA); (3) enforcing that all agents are in the *same grand* coalition (GR), combined with two control actions: adding agents (AddAg) or deleting agents (DelAg). We analyze these problems for friend-oriented and additive preferences under individual rationality, individual stability, Nash stability, and core stability. We provide a complete computational complexity classification for control in hedonic games.

KEYWORDS

Hedonic Games, Control; Computational Social Choice; Additive Preferences; Friend-Oriented Preferences

ACM Reference Format:
Jiehua Chen, Jakob Guttmann, Merisa Mustajbašić, and Sofia Simola. 2026. Control in Hedonic Games. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. https://doi.org/10.65109/WVYI7850

1 INTRODUCTION

Hedonic games [19] are coalition formation games where agents form coalitions based on their preferences over which coalition to join. The task is to partition agents into disjoint coalitions satisfying certain stability criteria. Typical stability criteria include *individual rationality* (no agent prefers being alone to his current coalition), *individual stability* (no agent prefers to join an existing coalition that would accept them), *Nash stability* (no agent can improve by unilaterally moving to another existing coalition), and *core stability* (no group of agents can all improve by forming a new coalition together). Since their introduction, hedonic games have become an important research topic in algorithmic game theory [8, 10] and computational social choice [2].

Most research has focused on analyzing fixed hedonic games–proposing new stability concepts, investigating existential questions (does a stable partition exist?), and computational questions (verifying stability, finding stable partitions when they exist). Far less attention has been paid to how external influence–such as adding or deleting agents–can shape outcomes in hedonic games. Such external influence is formalized as *control* in voting [4], where an external actor manipulates elections by adding or deleting voters or candidates. Control naturally arises in coalition formation settings where external actors shape outcomes: department chairs forming research collaborations, conference organizers balancing working groups, or managers assigning employees to project teams.

In this paper, we introduce and perform a systematic study of control in hedonic games, where an external party influences outcomes by adding or deleting agents. We formalize control problems combining two actions: adding agents (AddAg) or deleting agents (DelAg), with three goals: ensuring a specific agent is not alone (NA), ensuring a specific pair is together (PA), or forming a grand coalition (GR), all while reaching a stable partition. These goals capture common control scenarios:

- NA: A company forms project teams. A new employee joins who has not built relationships yet. Without intervention, they might remain isolated and unassigned. The manager can hire another employee with complementary skills, making them willing to form a team together or more attractive to existing teams.
- PA: The department chair wants two senior researchers to collaborate. A third, polarizing colleague, cannot work with one of them, which prevents a stable joint assignment. Reassigning this colleague to another project enables a mutually acceptable collaboration between the two senior researchers.
- GR: Multiple small research labs work on similar problems but compete rather than collaborate. A funding agency wants to form a single large consortium for a major grant. The agency can offer funding for postdocs who would only join if everyone collaborates, forcing consolidation into a grand coalition.

We analyze these problems for friend-oriented (FriHG) and additive (AddHG) preferences under four stability concepts: individual rationality (IR), individual stability (IS), Nash stability (NS), and core stability (CS).

Our contributions. Besides introducing the control model, we provide a comprehensive complexity picture of control in hedonic games. We summarize the key findings below:

- We provide a few polynomial-time algorithms for achieving NA and PA by adding agents in the friend-oriented preference setting. The algorithms exploit specific graph structures: For IR (and IS), we reduce to finding minimum-weight paths and cycles in arc-weighted directed graphs; For CS, we reduce to finding

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/WVYI7850

minimum-weight subgraphs in Steiner networks. These positive results show that targeted control goals (ensuring individuals are not isolated or specific pairs collaborate) are computationally tractable via agent addition.
– We discover that for most stability concepts, achieving NA or PA by deleting agents is impossible–the problems are immune; see Definition 4. This asymmetry between addition and deletion is somewhat surprising: You can add agents to force desired coalitions, but removing agents rarely helps.
– For all intractable (NP-hard, coNP-hard, or Σ_2^P-hard) control problems, the base problem–verifying whether the goal already holds without any control–is already intractable. This provides a natural barrier against control: If determining whether control is necessary is already hard, executing control attacks is not easier.
– For the control goal GR, the problem is mostly polynomial-time solvable for constant number k of control actions (i.e., in XP wrt. k). Interestingly, the complexity picture inverts between preference types: GR is more resistant to control than NA and PA in FriHG, but less resistant in AddHG. This suggests grand coalition formation has different structural properties than targeted pair or individual goals.

Our main results regarding the computational complexity are summarized in Table 1.

Related work. Hedonic games were first conceptualized by Drèze and Greenberg [19] and formally reintroduced by Banerjee et al. [3] and Bogomolnaia and Jackson [6]. They independently defined hedonic coalition formation games and analyzed fundamental stability concepts such as the core and Nash stability for additive preferences. Dimitrov et al. [17] introduced the friends-and-enemies model. Woeginger [30] surveyed different preference models for core stability. Brandt et al. [7] demonstrated that most stability-related problems remain intractable even under fairly restrictive preference assumptions. For a comprehensive overview of complexity results regarding verifying whether a partition is stable and determining the existence of stable partitions, we refer to a recent survey by Chen et al. [13].

To our knowledge, no prior work has studied control in hedonic games. Bartholdi III et al. [4] introduced electoral control (adding or deleting voters or candidates to change the election winner) into voting theory. Hemaspaandra et al. [25] explored both constructive and destructive control in elections under various voting rules, showing that many forms of control can be computationally difficult (a desirable property for election security), and providing a template for defining control actions in a precise algorithmic way; also see the book chapter by Faliszewski and Rothe [20] for more references on voting control.

Boehmer et al. [5] first studied control in matching markets, systematically exploring external control in Stable Marriage by defining a range of manipulative actions including adding or deleting agents. Chen and Schlotter [15] recently provided a comprehensive complexity overview of adding or removing agents in both Stable Marriage and Stable Roommates settings to achieve certain outcomes (such as guaranteeing the existence of a stable matching or ensuring a particular pair is matched). For control problems in other domains, we refer to a recent survey by Chen et al. [14].

Outline of the paper. In Section 2, we define hedonic games, the stability concepts we consider, and our control problems. In Section 3, we provide some structural results for hedonic games and our control problems. In Sections 4 and 5, we investigate the complexity for the friend-oriented preference and additive preference settings, respectively. For both sections, we first consider the control goal of NA, then PA, and finally GR. We conclude with a discussion on potential areas for future research in Section 6. Due to space constraints, proofs and statements marked with (⋆) are deferred to the full version of the paper [12]

2 PRELIMINARIES

Given an integer t, let $[t]$ denote the set $\{1, 2, \dots, t\}$. Given a directed graph G and a vertex $v \in V(G)$, let $N^+_G(v)$ and $N^-_G(v)$ denote the out- and in-neighborhood of v, respectively. Given a directed graph $G = (V, A)$ and a subset $V' \subseteq V$ of vertices, a subgraph *induced* by V', written as $G[V']$, is a subgraph (V', A') of G where $A' = \{a' \in A \mid a' \subseteq V'\}$.

Let $\mathcal{V}$ be a finite set of n agents. A *coalition* is a non-empty subset of $\mathcal{V}$. We call the entire agent set $\mathcal{V}$ the *grand coalition*. The input of a hedonic game is a tuple $(\mathcal{V}, (\succeq_i)_{i\in\mathcal{V}})$, where $\mathcal{V}$ is the agent set, and each agent $i \in \mathcal{V}$ has a preference order $\succeq_i$ over all coalitions that contain i. Each preference order is a weak order (i.e., complete, reflexive, and transitive). For two coalitions S and T containing i, we say that agent i *weakly prefers* S to T if $S \succeq_i T$; i (*strictly*) *prefers* S to T (written as $S \succ_i T$) if $S \succeq_i T$, but $T \not\succeq_i S$; i is *indifferent between* S and T (written as $S \sim_i T$) if $S \succeq_i T$ and $T \succeq_i S$.

A *partition* Π of $\mathcal{V}$ is a division of $\mathcal{V}$ into disjoint coalitions, i.e., the coalitions in Π are pairwise disjoint and $\bigcup_{A\in\Pi} A = \mathcal{V}$. Given a partition Π of $\mathcal{V}$ and an agent $i \in \mathcal{V}$, let $\Pi(i)$ denote the coalition which contains i. We also call the partition where every agent is in the grand coalition the *grand coalition partition*.

Compact preference representations of hedonic games. Since the number of possible coalitions that contain an agent is exponential in the number of agents, there is a need for compact representation of the preference orders of the agents. In this paper, we focus on two compact presentation models that can be encoded polynomially in the number of agents. The first setting is called hedonic games with *additively separable preferences (AddHG)*. In AddHG, every agent only needs to express a cardinal utility to every other agent. The second setting is simple restriction of the first setting and is called hedonic games with *friend-oriented preferences (FriHG)*. In FriHG, every agent regards every other agent either as a *friend* or an *enemy* such that he prefers coalitions with more friends to those with less friends.

Definition 1 (AddHG [3]). *Let $\mathcal{V}$ be a set of agents. The input of AddHG is a tuple $(\mathcal{V}, (\mu_i)_{i\in\mathcal{V}})$, where every agent $i \in \mathcal{V}$ has a* cardinal utility function *$\mu_i \colon \mathcal{V} \to \mathbb{R}$ such that for each two coalitions S and T containing i, agent i weakly prefers S to T if $\sum_{j\in S} \mu_i(j) \geq \sum_{j\in T} \mu_i(j)$. We assume $\mu_i(i) = 0$. The utility functions can also be compactly represented by a* preference graph. *It is a tuple $(\mathcal{G}, \omega)$, where $\mathcal{G}$ is a directed graph on the agent set and ω is an arc-weighting function. For each two agents i and j there is an arc (i, j) if and only if the utility of i to j is non-zero. The weight of this arc is equal to the non-zero utility $\omega(i, j) = \mu_i(j)$.*

Table 1: Complexity results of control in hedonic games, by either adding (AddAg) or deleting agents (DelAg), with the goals of ensuring that either a given agent is not alone (NA), or a pair of agents is in the same coalition (PA), or all agents are in the grand coalition (GR). We study two compact preference representations–additive preferences (AddHG) or friend-oriented preferences in the friends-and-enemies model (FriHG)–and four stability concepts–individual rationality (IR), individual stability (IS), Nash stability (NS), and core stability (CS). All hardness results hold even if no control action is allowed (i.e., $k = 0$). Problems with tag "DAG" (resp. "SYM") means that the corresponding hardness results hold even if the preference graph is a DAG (resp. symmetric). Entries labeled "P" denote polynomial-time solvability, "NPc" NP-completeness, and "$\Sigma_2^P c$" Σ_2^P-completeness. "imm" means immune and "imm★" means that it is immune while deciding yes-instances with $k = 0$ remains NP-hard. W[2]h and XP are with respect to the budget k. "never" means that the given instance is always a no instance. Results in bold are our contributions.

HG	NA-AddAg	NA-DelAg	PA-AddAg	PA-DelAg	GR-AddAg	GR-DelAg
FriHG-{IR,IS,CS}^{SYM}	**P** [T1,T2]	**imm** [P1]	**P** [T1,T2]	**imm** [P1]	**W[2]h**, **XP** [T3]	**P** [P4]
FriHG-NS	NPc [7],[O5]	NPc [7],[O5]	NPc [7],[O5]	NPc [7],[O5]	**W[2]h**, **XP** [T3]	**P** [P4]
SYM	**P** [P3]	**imm** [P1]	**P** [P3]	**imm** [P1]	**W[2]h**, **XP** [T3]	**P** [P4]
FriHG-{IR,IS,NS,CS}^{DAG}	**never** [O8]	**never** [O8]	**never** [O8]	**never** [O8]	**never** [O8]	**never** [O8]
AddHG-IR	**NPc** [T4]	**imm★** [P2]	**NPc** [T7,T8]	**imm★** [P2]	**W[2]h**, **XP** [T9,T10]	**W[2]h**, **XP** [T9,T10]
DAG/SYM	**P** [P5]	**imm** [P2]	**NPc** [T7,T8]	**imm★** [P2]	**W[2]h**, **XP** [T9,T10]	**W[2]h**, **XP** [T9,T10]
AddHG-{IS,NS}	NPc [29],[O5]	NPc [29],[O5]	NPc [29],[O5]	NPc [29],[O5]	**W[2]h**, **XP** [T9,T10]	**W[2]h**, **XP** [T9,T10]
DAG/SYM	**NPc** [T5,T6]	**NPc** [T5,T6]	**NPc** [T5,T6]	**NPc** [T5,T6]	**W[2]h**, **XP** [T9,T10]	**W[2]h**, **XP** [T9,T10]
AddHG-CS^{SYM}	$\Sigma_2^P c$ [28, 31],[O5]	$\Sigma_2^P c$ [28, 31],[O5]	$\Sigma_2^P c$ [28, 31],[O5]	$\Sigma_2^P c$ [28, 31],[O5]	**coNPc** [T11]	**coNPc** [T11]
DAG	**P** [P5]	**imm** [P2]	**NPc** [T7]	**imm★** [P2]	**W[2]h**, **XP** [T9]	**W[2]h**, **XP** [T9]

Definition 2 (FriHG [17]). *The input of FriHG is a directed graph $\mathcal{F}$ where every vertex corresponds to an agent such that every agent i considers another agent j as a friend if and only if there is an arc from i to j; otherwise i considers j as an enemy. Graph $\mathcal{F}$ is also called a friendship graph. For each agent $i \in \mathcal{V}$ and each two coalitions S and T containing i, agent i weakly prefers S to T if*
(i) either $|N^+_{\mathcal{F}}(i) \cap S| > |N^+_{\mathcal{F}}(i) \cap T|$, or
(ii) $|N^+_{\mathcal{F}}(i) \cap S| = |N^+_{\mathcal{F}}(i) \cap T|$ and $|S \setminus N^+_{\mathcal{F}}(i)| \leq |T \setminus N^+_{\mathcal{F}}(i)|$.

Remark 1. *Note that FriHG is a simple restriction of AddHG: Set $\mu_x(y) = n$ if (x, y) is an arc in $\mathcal{F}$, and $\mu_x(y) = -1$ otherwise.*

We say that an instance of AddHG (resp. FriHG) is a DAG if the preference graph (resp. friendship graph) is acyclic. Correspondingly, we say that it has f *feedback arcs* if the graph can be turned acyclic by deleting at most f arcs. Further, we say that AddHG is *symmetric* if for every pair of agents i, j it holds that $\mu_i(j) = \mu_j(i)$. Similarly, we say that an instance of FriHG is *symmetric* if the friendship graph $\mathcal{F}$ is symmetric. In these cases we can assume that the preference graph (resp. friendship graph) is undirected.

Relevant stability concepts. In this paper, we study four relevant stability concepts.

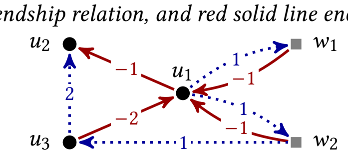
Definition 3. *Let Π be a partition of $\mathcal{V}$. A coalition B is blocking a partition Π if every agent $i \in B$ prefers B to $\Pi(i)$. An agent i and a coalition B form a blocking tuple if i prefers $B \cup \{i\}$ to $\Pi(i)$ and each agent $j \in B$ weakly prefers $B \cup \{i\}$ to $\Pi(j)$.*
- *Π is individually rational (IR) if no agent i prefers $\{i\}$ to $\Pi(i)$. If i prefers $\{i\}$ to $\Pi(i)$, we say he wishes to deviate from Π.*
- *Π is Nash stable (NS) if no agent i and coalition $B \in \Pi \cup \{\emptyset\}$ exist such that i strictly prefers $B \cup \{i\}$ to his coalition $\Pi(i)$.*
- *Π is individually stable (IS) if no agent i and coalition $B \in \Pi \cup \{\emptyset\}$ can form a blocking tuple.*
- *Π is core stable (CS) if no coalition is blocking Π.*

By definition, the stability concepts satisfy the following:

Observation 1. *An NS partition is IS. An IS partition is IR. A CS partition is IR.*

Our control problems and their complexity upper bounds. In this paper, we study three different control goals (1) enforcing that an agent is *not alone* (NA); (2) enforcing that a *pair* of agents is in the same coalition (PA); (3) enforcing that all agents are in the *same grand* coalition (GR). Moreover, we study two possible control actions we can use to obtain the control goals: adding agents (AddAg) or deleting agents (DelAg).

Example 1. *Let $U = \{x = u_1, u_2, u_3\}$ and $W = \{w_1, w_2\}$, where U is a set of consisting of three original agents while W consisting of two additional agents, respectively. The preference graph of the agents $U \cup W$ is depicted below. Throughout, we use blue dotted line to indicate friendship relation, and red solid line enemy relation.*



Our special agent is $x = u_1$. In the original instance, consisting of only U, agent u_1 must be alone in an IR partition since he dislikes u_2 but u_3 dislikes him. One can verify that it is impossible to make u_1 not alone by deleting agents since otherwise the new IR partition augmented with the deleted agents in singletons would yield an IR partition for the original instance.

We can add agent w_2 (not w_1) to the original instance to obtain an IR partition, where every agent is in the same grand coalition.

Now we are ready to formally define our control problems (as decision problems) HG-S-$\mathcal{G}$-$\mathcal{A}$, where S ∈ {IR, IS, NS, CS} denotes one of the four stability concepts, $\mathcal{G}$ ∈ {NA, PA, GR} one of the three control goals, and $\mathcal{A} \in \{\text{AddAg}, \text{DelAg}\}$ one of the two control actions, respectively. The problem of control by adding agents is defined as follows:

HG-S-NA-AddAg (*resp.* HG-S-PA-AddAg)

Input: A hedonic game instance $(\mathcal{V} = \mathcal{U} \cup \mathcal{W}, (\succeq_i)_{i\in\mathcal{V}})$, a selected agent $x \in \mathcal{U}$ (resp. selected agent pair $\{x, y\} \subseteq \mathcal{U}$), and a non-negative value $k \in \mathbb{N} \cup \{0\}$.

Question: Is there a subset $\mathcal{W}' \subseteq \mathcal{W}$ of size $|\mathcal{W}'| \le k$ such that $\mathcal{U} \cup \mathcal{W}'$ admits an S-partition Π where $\Pi(x) \neq \{x\}$ (resp. $\Pi(x) = \Pi(y)$)?

For the control goal GR, the input only consists of the hedonic game instance and a budget:

HG-S-GR-AddAg

Input: A hedonic game instance $(\mathcal{V} = \mathcal{U} \cup \mathcal{W}, (\succeq_i)_{i\in\mathcal{V}})$ and a non-negative value $k \in \mathbb{N} \cup \{0\}$.

Question: Is there a subset $\mathcal{W}' \subseteq \mathcal{W}$ of size $|\mathcal{W}'| \le k$ such that the partition consisting of the grand coalition $\mathcal{U} \cup \mathcal{W}'$ is S?

In the above, we call the set $\mathcal{U}$ the *original* agents and the set $\mathcal{W}$ the *additional* agents.

Similarly, we define the problems for the setting where we can delete agents.

HG-S-NA-DelAg (*resp.* HG-S-PA-DelAg)

Input: A hedonic game instance $(\mathcal{U}, (\succeq_i)_{i\in\mathcal{U}})$, a selected agent $x \in \mathcal{U}$ (resp. selected agent pair $\{x, y\} \subseteq \mathcal{U}$), and a non-negative value $k \in \mathbb{N} \cup \{0\}$.

Question: Is there a subset $\mathcal{U}' \subseteq \mathcal{U}$ of size $|\mathcal{U}'| \le k$ such that $\mathcal{U} \setminus \mathcal{U}'$ admits an S-partition Π where $\Pi(x) \neq \{x\}$ (resp. $\Pi(x) = \Pi(y)$)?

HG-S-GR-DelAg

Input: A hedonic game instance $(\mathcal{U}, (\succeq_i)_{i\in\mathcal{U}})$ and a non-negative value $k \in \mathbb{N} \cup \{0\}$.

Question: Is there a subset $\mathcal{U}' \subseteq \mathcal{U}$ of size $|\mathcal{U}'| \le k$ such that the partition consisting of the grand coalition $\mathcal{U} \setminus \mathcal{U}'$ is S?

If the hedonic game is restricted to be AddHG or FriHG, we replace the prefix HG with AddHG or FriHG.

To classify cases where control is impossible, we define the following.

Definition 4 (immune and never). *We say a hedonic game control problem is immune if for every No-instance with $k = 0$, it remains a No-instance even if we set $k = \infty$. We say that the problem is never if all instances containing at least two agents are No-instances.*

Clearly, an instance with the control goal NA with one agent overall is always a No-instance, and an instance with the control goal PA contains at least two agents by definition. On the other hand, an instance with the control goal GR that contains one agent is trivially a Yes-instance.

It is known that for both preference settings, one can check in polynomial time whether a given partition is IR, IS, or NS. However, the verification problem for CS is coNP-complete [11]. This immediately yields the following complexity upper bounds for our control problems.

Observation 2. *For preference model M $\in$ {AddHG, FriHG}, control goal $\mathcal{G} \in$ {NA, PA, GR}, and control action $\mathcal{A} \in$ {AddAg, DelAg}, the first three sets of problems M-IR-$\mathcal{G}$-$\mathcal{A}$, M-IS-$\mathcal{G}$-$\mathcal{A}$, and M-NS-$\mathcal{G}$-$\mathcal{A}$ are in NP, while the problems M-CS-NA-$\mathcal{A}$ and M-CS-PA-$\mathcal{A}$ are in Σ_2^P, and the problem M-CS-GR-$\mathcal{A}$ is in coNP.*

W[2] and XP. In this paper, we prove that several problems are W[2]-hard (resp. in XP) with respect to a parameter p. W[2]-hardness is shown via a *parameterized reduction* from a W[2]-hard problem Q (parameter q) running in time $f(q) \cdot |I_Q|^{O(1)}$ and producing an instance with parameter $p \le g(q)$, for computable f, g. Thus, it is unlikely that the problems are fixed-parameter tractable (FPT), i.e., solvable in time $f'(p) \cdot |I|^{O(1)}$ for computable f', unless FPT = W[2]. Membership in XP wrt. p means solvability in time $|I|^{h(p)}$ for some computable h (equivalently, polynomial time for every constant p). If a problem remains NP-hard for some fixed constant value of p, then it is not in XP wrt. p, unless P = NP. See [16, 27] for details.

3 STRUCTURAL OBSERVATIONS

In this section, we collect some useful structural properties that may be of independent interest.

Relations among the stability concepts. The first two observations describe two cases when a more stringent stability concept is equivalent to IR. This is useful in searching for both algorithms and hardness results.

The first observation follows directly from definition.

Observation 3. *For every S $\in$ {NS, IS}, the grand coalition partition is S if and only if it is IR.*

Next, we observe that if the preference graph is a DAG, the two stability concepts IR and CS coincide.

Observation 4. *For every instance of AddHG with an acyclic preference graph, a partition is CS if and only if it is IR.*

Proof. By Observation 1, it suffices to show that IR implies CS. Suppose, for the sake of contradiction, that Π is IR but not CS. Let B be a blocking coalition with at least two agents. Since the preference graph is acyclic, one agent $t \in B$ is a sink in the preference graph induced by B. Such a sink t has zero utility towards B, but has utility at least zero towards $\Pi(t)$, since Π is IR. Thus, agent t has no incentive to deviate to B, a contradiction. □

The next lemma shows that under the friend-oriented model and for the control goal PA, the existence questions for IR and IS are essentially the same.

Lemma 1 (★). *Let $I = (\mathcal{V}, \mathcal{F})$ be a FriHG-instance, and x and y two agents in $\mathcal{V}$. From each IR partition Π with $\Pi(x) = \Pi(y)$, one can construct in polynomial time an IS partition Π' with $\Pi'(x) = \Pi'(y)$.*

Proof sketch. The idea is to start with an IR partition where x and y are in the same coalition, say C, and merge it with all strongly connected components in $\mathcal{F}$. We then repeatedly add agents to C whenever they can reach some agent in C. All remaining agents will be in their own singleton coalitions. □

Influence of the control goals on the complexity. The next result implies that enforcing the control goal NA or PA will not help in lowering the complexity if the underlying problem of determining the existence of a stable partition is hard.

Observation 5 (★). *For every* M ∈ {AddHG, FriHG} *and* S ∈ {IR, NS, IS, CS} *if it is* Σ_2^P*-hard (resp.* NP*-hard) to determine the existence of an* S*-partition under preference representation* M*, then so is* M-S-$\mathcal{G}$-$\mathcal{A}$ *for every control goal* $\mathcal{G}$ ∈ {NA, PA}*, control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*, even when* $k = 0$*. The same implication holds even when the preferences are symmetric.*

Proof sketch. Add two additional agents $\{x, y\}$ who have positive utility towards each other and negative to all other agents. They will not affect the stability of the rest of the structure. □

The next result allows us to focus on the control goal PA when searching for efficient algorithms for the goal NA.

Observation 6 (★). *For every* M ∈ {AddHG, FriHG}*,* S ∈ {IR, NS, IS, CS}*, and* $\mathcal{A}$ ∈ {AddAg, DelAg}*, if* M-S-PA-$\mathcal{A}$ *is polynomial-time solvable, then so is* M-S-NA-$\mathcal{A}$.

Proof sketch. We guess the other agent $y \in (\mathcal{U} \cup \mathcal{W})$ with whom x should be paired with. If there is no y such that M-S-PA-$\mathcal{A}$ is a Yes-instance, x will be always alone. □

Next, we state a well-known tight relation between the coalitions in a CS partition and the strongly connected components of the friendship graph, which is very useful for the PA goal and CS stability.

Observation 7 ([17, 30],★). *For every FriHG-instance and two agents x and y, there is a core stable partition Π such that x and y belong to the same coalition if and only if they belong to the same strongly connected component.*

Immune cases. We close this section by summarizing cases when control is immune.

Proposition 1. *For each stability concept* S ∈ {IR, IS, CS} *and each control goal* $\mathcal{G}$ ∈ {NA, PA}, FriHG-S-$\mathcal{G}$-DelAg *is immune. Under symmetric preferences* FriHG-NS-NA-DelAg *and* FriHG-NS-PA-DelAg *are immune.*

Proof. First observe that if it is possible to delete some agents to have an IR-partition such that x is not alone (resp. x and y are in the same coalition), then adding back the deleted agents each as a singleton coalition would be an IR-partition with the same goal as well. By Lemma 1, this result extends to IS as well.

By Observation 7, if it is possible to delete some agents to have a CS-partition such that x is not alone (resp. x and y are in the same coalition), then x must be part of a connected component of size at least two (resp. x and y are part of the same strongly connected component) in the friendship graph of the remaining agents. If the agents are connected, they are also connected after adding back the agents. By Observation 7, there is then a CS partition of the original agent set such that x is not alone (resp. x and y are together).

If it is possible to delete some agents to have a NS partition such that x is not alone (resp. x and y are in the same coalition), then x must have a friend (resp. x and y both have a friend). Let us now add back all the removed agents. Consider the partition in which all agents with degree at least one are placed in the same big coalition, while all remaining agents form singleton coalitions. This partition is NS. Since preferences are symmetric, no agent in the big coalition is friends with the singletons, and since every agent in the big coalition has a friend within that coalition, he prefers it to being alone. Since x has a friend (resp. x and y have a friend) he (resp. they) must be in the big coalition. □

DAGs in FriHG are very restrictive. The IR partition is unique and no two agents are together in it.

Observation 8 (★). *For FriHG-instances with acyclic friendship graph, the only partition that may be stable for* S ∈ {IR, IS, NS, CS} *consists of every agent being in a singleton coalition.*

Proposition 2. *For control goal* $\mathcal{G}$ ∈ {NA, PA}, AddHG-IR-$\mathcal{G}$-DelAg *is immune.*

Proof. If it would be possible to delete some agents to have an IR-partition such that x is not alone (resp. x and y are in the same coalition), then adding back the deleted agents each as a singleton coalition would be an IR-partition with the same goal as well. □

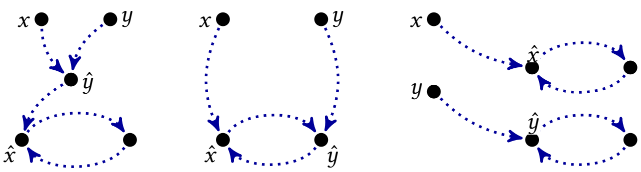


Figure 1: The three possible substructures for an IR coalition containing x and y. The arcs indicate paths of varied lengths. Structure on left corresponds to line 6, middle to line 8, and right to line 9.

4 FRIEND-ORIENTED PREFERENCES

In this section, we discuss the results relating to FriHG. For the control goals NA and PA, we discover that IR, IS, and CS are all polynomial-time solvable for the control action AddAg: These stability concepts require certain friendship structures and finding a minimum number of agents to add to realize those can be done in polynomial time. In Proposition 1 we saw that these stability concepts were also immune to the control action DelAg.

For NS, on the other hand, FriHG-NS-$\mathcal{G}$-$\mathcal{A}$ is NP-hard for every $\mathcal{G}$ ∈ {NA, PA}, $\mathcal{A}$ ∈ {AddAg, DelAg} even when $k = 0$ by a result from [7]. We however discover that symmetric preferences make the problems tractable.

Finally, we look into enforcing that the grand coalition partition is stable (GR) and see that regardless of the stability concept, obtaining grand coalition through agent deletion is computationally easier than through agent addition.

We start the section by presenting an algorithm that solves FriHG-IR-PA-AddAg.

Theorem 1 (★). *For each stability concept* S ∈ {IR, IS} *and each control goal* $\mathcal{G}$ ∈ {NA, PA}, FriHG-S-$\mathcal{G}$-AddAg *is polynomial-time solvable.*

Proof Sketch. The idea is that a coalition containing both x and y, which is part of an IR partition, must contain one of three possible structures, which we illustrate in Figure 1. Any agent who is part of an IR partition must be part of a "ρ-shaped" subgraph [23, Chapter 14], that is, a graph that consists of a directed cycle of

Algorithm 1: Algorithm for FriHG-IR-PA-AddAg.

Input: $(\mathcal{V} = \mathcal{U} \cup \mathcal{W}, \mathcal{F} = (\mathcal{U} \cup \mathcal{W}, \mathcal{A}), x, y, k)$

foreach $(i, j) \in \mathcal{A}$ **do** $\omega((i, j)) \leftarrow \begin{cases} 1, & \text{if } i \in \mathcal{W} \\ 0, & \text{otherwise.} \end{cases}$;

foreach $i, j \in \mathcal{V}$ **do**

  $\omega^P(i, j) \leftarrow$ weight of a min-weight path from i to j.

foreach $i \in \mathcal{V}$ **do** $\omega^C(i) \leftarrow$ weight of a min-weight non-trivial cycle containing i.;

foreach $\hat{x}, \hat{y} \in \mathcal{V}$ **do**

  if $\omega^P(x, \hat{y}) + \omega^P(y, \hat{y}) + \omega^P(\hat{y}, \hat{x}) + \omega^C(\hat{x}) \leq k$ **then**

    return yes

  if $\omega^P(x, \hat{x}) + \omega^P(y, \hat{y}) + \omega^P(\hat{x}, \hat{y}) + \omega^P(\hat{y}, \hat{x}) \leq k$ *and* $\hat{x} \neq \hat{y}$ **then return yes**;

  if $\omega^P(x, \hat{x}) + \omega^P(y, \hat{y}) + \omega^C(\hat{x}) + \omega^C(\hat{y}) \leq k$ **then**

    return yes

return no

length at least two and a path (possibly of length zero) that reaches a vertex in it. This is a consequence of every agent obtaining a friend in an IR coalition of size at least two: If we start from an agent and follow a friendship path, we must eventually encounter an agent we have already seen. Since x and y are both a part of an IR coalition, they both must be a part of a "ρ-shaped" subgraph. There are three ways in which they can intersect, as shown in Figure 1. We show that these structures can be found in polynomial time by combining with finding minimum-weight paths. The algorithm is in Algorithm 1.

In the algorithm we construct an arc-weight function ω for $\mathcal{F}$ based on whether an arc starts with an agent in $\mathcal{W}$ or not. We compute the all-pair minimum-weight paths on $(\mathcal{F}, \omega)$ in polynomial time using Floyd-Warshall [22] or some other algorithm. While running the algorithm, we also store, for each agent, an extra entry that records the minimum-weight path from the agent to himself of length at least one, i.e., the minimum-weight non-trivial cycle containing the agent. Then, we use those paths and cycles to search for the structures shown in Figure 1.

By Lemma 1, this algorithm can also be used to solve FriHG-IS-PA-AddAg. By Observation 6, we then can solve the remaining FriHG-IR-NA-AddAg and FriHG-IS-NA-AddAg. However, the latter two problems can be solved more efficiently by searching for a single "ρ-shaped" subgraph containing x. □

Brandt et al. [7] show that determining the existence of an NS partition is NP-hard for FriHG, and thus FriHG-NS-$\mathcal{G}$-$\mathcal{A}$ is NP-hard for every $\mathcal{A} \in \{\text{AddAg}, \text{DelAg}\}, \mathcal{G} \in \{\text{NA}, \text{PA}\}$ by Observation 5. Moreover, if an instance of FriHG is a DAG, then by Observation 8 we cannot hope to have any pair of agents in the same coalition in a stable partition. However, we obtain a non-trivial–although simple–algorithm when the preferences are symmetric for adding agents, whereas deleting agents is immune (see Proposition 1).

Proposition 3 (★). *For symmetric preferences,* FriHG-NS-NA-AddAg *and* FriHG-NS-PA-AddAg *are polynomial-time solvable.*

Proof sketch. For symmetric friendship relations, the following partition is NS: Put all agents who have at least one friend together, while all agents without any friends form singleton coalitions. Due to this, the problem of ensuring x and y are together reduces to adding at most two agents to the original graph and checking whether afterwards x and y will each have a friend. □

For core stability, we can solve the control goal PA efficiently by reducing to finding a minimum-weight subgraph where x and y are mutually reachable. This is connected to the 2-DSN problem, which admits a polynomial-time algorithm [21, 26]:

Directed Steiner Network (2-DSN)
Input: A directed graph $G = (V, A)$ with an arc-weighting function $\omega: A \to \mathbb{R}$, two pairs $(s_1, t_1), (s_2, t_2)$, and $\delta \in \mathbb{R}$.
Question: Is there a subgraph $H = (V', A')$ of G which contains a path from s_1 to t_1 and from s_2 to t_2, respectively, such that $\sum_{a \in A'} \omega(a) \leq \delta$?

Theorem 2 (★). FriHG-CS-NA-AddAg *and* FriHG-CS-PA-AddAg *are polynomial-time solvable.*

Proof Sketch. By Observation 7, two agents x and y can be in the same coalition in a CS partition if and only if they are in the same strongly connected component of the friendship graph, i.e., if there is a path from x to y and vice versa. This is closely related to the polynomial-time solvable 2-DSN-problem [21, 26].

The main difference is that 2-DSN has arc-weights, whereas we have vertex weights. To address this, we replace every agent in $\mathcal{W}$ with two agents, one taking care of the in-arcs and the other the out-arcs. Let $((\mathcal{V} = \mathcal{U} \cup \mathcal{W}, \mathcal{F}), x, y, k)$ be an instance of FriHG-CS-PA-AddAg. We will transform this instance into an instance $I = (G, \omega, (s_1, t_1), (s_2, t_2))$ of 2-DSN as follows:

– For every $i \in \mathcal{U}$, we add to G a vertex u_i.
– For every $i \in \mathcal{W}$, we add to G two vertices, w_i^1 and w_i^2.

The idea is that w_i^1 inherits the in-arcs of $i \in \mathcal{W}$ and w_i^2 the out-arcs. Formally, we construct the following arcs with weights and add them to G.

– For every arc $(i, j) \in A(\mathcal{F})$, we add to G an arc (u, v) with weight 0, where • if $i \in \mathcal{U}$, then $u := u_i$, otherwise $u := w_i^2$ and • if $j \in \mathcal{U}$, then $v := u_j$, otherwise $v := w_j^1$.
– For every $i \in \mathcal{W}$, we construct the arc (w_i^1, w_i^2) with weight 1.

We set $(s_1, t_1) := (u_x, u_y)$ and $(s_2, t_2) := (u_y, u_x)$. Moreover, we set $\delta := k$. The correctness is deferred to the appendix. □

We finally look into enforcing that grand coalition is stable. For every S $\in$ {IR, IS, NS, CS}, we discover an algorithm for FriHG-S-GR-AddAg that runs in time $|\mathcal{W}|^k |\mathcal{V}|^{O(1)}$. However, we show that the problem is W[2]-hard wrt. k, i.e., it is unlikely to admit an algorithm that runs in time $f(k) \cdot |\mathcal{V}|^{O(1)}$, where f is some computable function. The result is through a straightforward reduction from Set Cover, which is W[2]-hard wrt. the set cover size h [18].

Theorem 3 (★). *For every* S $\in$ {IR, IS, NS, CS}, FriHG-S-GR-AddAg *is W[2]-hard wrt. k even when the preference graph is symmetric and in* XP *wrt. k.*

In contrast to FriHG-S-GR-AddAg, FriHG-S-GR-DelAg is solvable in polynomial time for every S $\in$ {IR, IS, NS, CS}. For each stability concept, we can determine in polynomial time a maximum subset of agents such that the grand coalition partition of it is stable. For IR, IS, and NS, we obtain this set by recursively removing agents who have no friends, as they cannot be contained in an IR partition. For CS, it is known that every coalition in a stable partition is a strongly connected component of the friendship graph. Thus we only need to keep a largest strongly connected component.

Proposition 4 (★). *For every stability concept* S ∈ {IR, IS, NS, CS}, FriHG-S-GR-DelAg *is polynomial-time solvable.*

5 ADDITIVE PREFERENCES

In this section, we consider the case with additive preferences. It turns out that most of our control problems remain intractable even in very restricted cases such as when the budget is zero or the preference graph is a DAG or symmetric. As in the previous section, we first consider NA, then PA, and finally GR.

First, we show that even the most basic stability requirement IR becomes difficult to determine once we enforce some control goal.

Theorem 4. *For each control action* $\mathcal{A} \in \{\text{AddAg}, \text{DelAg}\}$, AddHG-IR-NA-$\mathcal{A}$ *is* NP-*complete; NP-hardness remains even if the budget is* $k = 0$ *and the preference graph has only one feedback arc.*

Proof. The NP-containment follows directly from Observation 2. Moreover, since $k = 0$, both control problems reduce to checking whether there is an IR partition with agent x not alone in his coalition. It remains to show the NP-hardness for the mentioned restriction. In the remainder of the proof, we will focus on this restricted variant and reduce from the following well-known NP-complete problem [24].

Restricted Exact Cover by 3-Sets (RX3C)

Input: A $3\hat{n}$-element set $\mathcal{E} = [3\hat{n}]$ and a collection $\mathcal{S} = \{S_1, \ldots, S_{3\hat{n}}\}$ of 3-element subsets of $\mathcal{E}$ such that each element appears in *exactly* three members of $\mathcal{S}$.

Question: Does $\mathcal{S}$ contain an *exact cover* for $\mathcal{E}$, i.e., a subcollection $\mathcal{K} \subseteq \mathcal{S}$ such that $|\mathcal{K}| = \hat{n}$ and $\cup_{S_j \in \mathcal{K}} S_j = \mathcal{E}$?

Let $I = (\mathcal{E} = [3\hat{n}], \mathcal{S})$ be an instance of RX3C. We construct an AddHG-instance whose preference graph has one feedback arc. The set of agents is $\mathcal{U} = \{x, y_0, y_1\} \cup \{u_i \mid i \in \mathcal{E}\} \cup \{s_j \mid S_j \in \mathcal{S}\}$, where

- x is the special agent not to be alone,
- y_0, y_1 are auxiliary-agents,
- each element $i \in \mathcal{E}$ has one *element-agent* u_i,
- each set $S_j \in \mathcal{S}$ has one *set-agent* s_j.

We construct the utilities as follows; the unmentioned utilities are 0:

- For each agent $a \in \mathcal{U} \setminus \{x\}$, set $\mu_a(x) = -1$.
- For each element $i \in [3\hat{n}-1]$, set $\mu_{u_i}(u_{i+1}) = 1$, for element-agent $u_{3\hat{n}}$, set $\mu_{u_{3\hat{n}}}(y_0) = 1$.
- For auxiliary-agent y_0, set $\mu_{y_0}(y_1) = 1$, for y_1, set $\mu_{y_1}(u_1) = 1$.
- For each element $i \in [3\hat{n}]$, set $\mu_{u_i}(y_1) = -1$.
- For each set $S_j \in \mathcal{S}$ and all elements $i \in S_j$, set $\mu_{u_i}(s_j) = 1$.
- For each set $S_j \in \mathcal{S}$, set $\mu_{s_j}(y_1) = 1$.
- For all $S_j, S_\ell \in \mathcal{S}$ with $j < \ell$ and $S_j \cap S_\ell \neq \emptyset$, set $\mu_{s_j}(s_\ell) = -1$.

To complete the construction, we define x as the agent not being alone. Before we show the correctness, we first observe that by deleting the arc (y_1, u_1) we obtain a DAG with following topological order: $u_1, \ldots, u_{3\hat{n}}, s_1, \ldots, s_{3\hat{n}}, y_0, y_1, x$.

Claim 4.1 (★). *If* $\mathcal{K}$ *is an exact cover for* I*, then the following partition* Π *with* $\Pi(x) = \{x, u_1, u_2, \ldots, u_{3\hat{n}}, y_0, y_1\} \cup \{s_j \mid S_j \in \mathcal{K}\}$ *and* $\Pi(s_j) = \{s_j\}$ *for all* $S_j \notin \mathcal{K}$ *is* IR.

Claim 4.2. *If* Π *is an* IR *partition with* $|\Pi(x)| \geq 2$*, then the sets corresponding to the set-agents in* $\Pi(x)$ *form an exact cover of* I.

Proof. Let $C := \Pi(x)$ be the coalition containing special agent x with $|C| \geq 2$. We want to show that the sets corresponding to set-agents contained in C form an exact cover of I. We start by showing that auxiliary-agent y_1 must be in C.

If $s_j \in C$ for some $S_j \in \mathcal{S}$, then he needs to be in the same coalition with y_1 for C to be IR, because he has negative utility towards x and the only positive utility is towards y_1. The same argument holds if $y_0 \in C$: He has negative utility towards x and his only positive utility towards y_1.

If an element-agent $u_i \in C$ for some $i \in [3\hat{n}]$, then there is a set-agent s_j with $i \in S_j$ such that $s_j \in C$, which implies $y_1 \in C$, or $u_{i+1} \in C$; otherwise u_i has negative utility towards C. By repeating this argument, either a set-agent–and thus y_1–must be in C, or $u_{3\hat{n}} \in C$. However, if $u_{3\hat{n}}$ is in C, then either a set-agent or y_0 must be in C; in both cases, this implies y_1 must also be in C.

Thus $y_1 \in C$ and hence $u_1 \in C$, because y_1 has negative utility towards x and only positive utility towards u_1. Agent u_1 has -1 utility towards x and y_1 in C, so u_1 needs at least one set-agent s_j with $1 \in S_j$ to not deviate. However, there can be at most one $s_j \in C$ with $1 \in S_j$: Suppose, towards a contradiction, that there are $s_j, s_{j'} \in C, j < j'$ with $1 \in S_j$ and $1 \in S_{j'}$. Then s_j has negative utility towards x and $s_{j'}$. But he has only one positive out-arc to y_1, so C would not be IR anymore, contradiction.

Hence we have exactly one set-agent $s_j \in C$ with $1 \in S_j$. Thus $u_2 \in C$ for C to be IR. Similarly to element 1, there must be exactly one $s_j \in C$ with $2 \in S_j$. Same arguments hold for $u_3, \ldots, u_{3\hat{n}}$.

To summarize, we need for every element-agent u_i for $i \in [3\hat{n}]$ at least one set-agent s_j with $i \in S_j$ which implies we have a set cover of all sets corresponding to the set-agents contained in C. As we have discussed, the sets must also be disjoint, so the set cover is also an exact-cover. (end of the proof of Claim 4.2) ⋄

The correctness follows immediately from Claims 4.1 and 4.2. □

The above result is tight since for DAGs or symmetric preferences, we can solve the problem in polynomial time. In both cases we need to find only one agent $a \in \mathcal{U}$, such that x has non-negative utility towards agent a or vice versa:

Proposition 5 (★). *For DAGs,* AddHG-IR-NA-AddAg *and hence* AddHG-CS-NA-AddAg *is polynomial-time solvable. For symmetric preferences,* AddHG-IR-NA-AddAg *is polynomial-time solvable.*

Next, we consider the two more stringent stability concepts IS and NS. It is known that determining the existence of NS and IS partitions is NP-hard on AddHG [29]. Hence, by Observation 3, the control problems with goals NA and PA are NP-hard as well. We discover that the control problems remain NP-hard even when the preference graph is a DAG or symmetric.

Theorem 5 to Theorem 8 are all shown via reductions from RX3C, using approaches that are similar in structure but distinct in their technical details.

Theorem 5 (★). *For each stability concept* S ∈ {IS, NS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-NA-$\mathcal{A}$ *and* AddHG-S-PA-$\mathcal{A}$ *are NP-complete; NP-hardness remains even if the budget is* $k = 0$ *and the preference graph is a DAG.*

Hardness for IS and NS extend to symmetric preferences.

Theorem 6. *For each stability concept* S ∈ {IS, NS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-NA-$\mathcal{A}$ *is NP-complete; NP-hardness remains even if the budget is* $k = 0$ *and the preference graph is symmetric.*

Proof. The NP-containment follows directly from Observation 2. As before, we will reduce from RX3C to show the NP-hardness for the mentioned control problem.

Let $I = (\mathcal{E} = [3\hat{n}], \mathcal{S})$ be an instance of RX3C. Let us construct a hedonic game instance with additive and symmetric preferences as follows. The set of agent is $\mathcal{U} = \{x, y\} \cup \{u_i \mid i \in \mathcal{E}\} \cup \{s_j \mid S_j \in \mathcal{S}\} \cup \{d_\ell \mid \ell \in [2\hat{n}]\}$, where

- x is the special agent and y an auxiliary-agent,
- each element $i \in \mathcal{E}$ has one *element-agent* u_i,
- each set $S_j \in \mathcal{S}$ has one *set-agent* s_j,
- and auxiliary-agents d_ℓ for $\ell \in [2\hat{n}]$.

We construct the symmetric utilities as follows; the unmentioned utilities are 0:

- Set $\mu_x(y) = \mu_y(x) = 1$.
- For each agent $a \in \mathcal{U} \setminus \{x, y\}$, set $\mu_x(a) = \mu_a(x) = -1$.
- For each $S_j \in \mathcal{S}$, set $\mu_{s_j}(y) = \mu_y(s_j) = 1$.
- For each set $S_j \in \mathcal{S}$ and all elements $i \in S_j$, set $\mu_{s_j}(u_i) = \mu_{u_i}(s_j) = 1$.
- For each set $S_j \in \mathcal{S}$ and every auxiliary-agent d_ℓ for $\ell \in [2\hat{n}]$, set $\mu_{s_j}(d_\ell) = \mu_{d_\ell}(s_j) = 3$.

The sketch of the preference graph can be found in the full version of the paper [12]. To complete the construction, let agent x be the special agent not to be alone.

In the forward direction of the correctness proof, we show that if I has an exact cover, we can construct a NS partition with $|\Pi(x)| \geq 2$, which is, by Observation 1, also IS. In the backward direction, we assume the partition Π being IS–an even weaker assumption than NS–and show that this implies the existence of an exact cover for I.

Claim 6.1 (★). *If* $\mathcal{K}$ *is an exact cover for* I*, then the following partition* Π *with* $\Pi(x) = \{x, y\}$, $\Pi(s_j) = \{s_j, u_{j_1}, u_{j_2}, u_{j_3}\}$ *for* $S_j = \{j_1, j_2, j_3\} \in \mathcal{K}$ *and* $\Pi(s_j) = \{s_j, d_\ell\}$ *for* $S_j \notin \mathcal{K}$ *and* $\ell \in [2\hat{n}]$ *is* NS.

Claim 6.2 (★). *If* Π *is an* IS *partition with* $|\Pi(x)| \geq 2$*, then* $y \in \Pi(x)$ *and all set-agents which are in a coalition with their three element-agents form an exact cover of* I.

The correctness follows immediately from Claims 6.1 and 6.2. □

In contrast to Proposition 5, PA remains NP-hard on DAGs or symmetric preferences, even when no agent can be added or deleted.

Theorem 7 (★). *For each stability concept* S ∈ {IR, CS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-PA-$\mathcal{A}$ *is* NP-*complete; NP-hardness remains even if the budget is* $k = 0$ *and the preference graph is acyclic with maximum vertex degree nine.*

The hardness remains for symmetric preferences.

Theorem 8 (★). *For each stability concept* S ∈ {IR, IS, NS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-PA-$\mathcal{A}$ *is* NP-*hard even if the budget is* $k = 0$ *and the preference graph is symmetric.*

Finally, we consider the control goal GR: Making the grand coalition partition stable. We already know from the previous section that for every S ∈ {IR, IS, NS, CS}, FriHG-S-GR-AddAg is NP-hard and W[2]-hard wrt. k even for symmetric preferences. We strengthen this result by showing the same for the DelAg case.

Theorem 9 (★). *For each stability concept* S ∈ {IR, IS, NS, CS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-GR-$\mathcal{A}$ *is in* XP *and* W[2]*-hard wrt.* k *when the preference graph is a DAG.*

Theorem 10 (★). *For each stability concept* S ∈ {IR, IS, NS} *and each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-S-GR-$\mathcal{A}$ *is* W[2]*-hard and in* XP *wrt.* k*. The hardness holds even when the preference graph is symmetric.*

For core stability, Theorem 9 shows that AddHG-CS-GR-AddAg and AddHG-CS-GR-DelAg admit a polynomial-time algorithm when the preference graph is a DAG and the budget is a constant. However, in contrast with IR, IS, and NS, the problems for CS are in general coNP-complete even when $k = 0$. The hardness holds even when the preferences are symmetric.

Theorem 11 (★). *For each control action* $\mathcal{A}$ ∈ {AddAg, DelAg}*,* AddHG-CS-GR-$\mathcal{A}$ *is* coNP-*complete. It remains* coNP-*hard even when* $k = 0$ *and the preference graph is symmetric.*

6 CONCLUSION

Motivated by control in other computational social choice problems, we introduce control in hedonic games. We study three control goals: ensuring an agent is not alone (NA), ensuring a pair of agents is together (PA), and ensuring the grand coalition is stable (GR), combined with two control actions: adding and deleting agents. We present a complete complexity picture for these control goals and actions across four stability concepts and two preference representations–FriHG and AddHG.

Our work opens several potential directions for future research. First, alternative control actions remain unexplored. In Stable Roommates and Marriage settings, a common control action is removing acceptability–making previously acceptable pairs unacceptable to each other. Analogously, one could study removing friendship relations in FriHG or adjusting utility values in AddHG. Second, inspired by destructive control in voting [25], one could study destructive control in hedonic games: ensuring a specific agent remains isolated or preventing a specific pair from being in the same coalition. Third, other solution concepts merit investigation, such as Pareto optimal or strictly core stable partitions. Finally, other simple compact preference representations warrant exploration, such as fractional hedonic games [1], anonymous preferences [6] or B- and W-preferences [9]. We conjecture that the ideas of many of our hardness reductions could work also for fractional hedonic games, since Observation 5 extends to any compact preference representation. However, anonymous, B- and W-preferences have a different structure and therefore likely require different ideas.

ACKNOWLEDGMENTS

This work and all authors were supported by the Vienna Science and Technology Fund (WWTF) [10.47379/VRG18012].

REFERENCES

[1] Haris Aziz, Florian Brandl, Felix Brandt, Paul Harrenstein, Martin Olsen, and Dominik Peters. 2019. Fractional Hedonic Games. *ACM Transactions on Economics and Computation* 7, 2 (2019).

[2] Haris Aziz and Rahul Savani. 2016. Hedonic Games. In *Handbook of Computational Social Choice*, Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia (Eds.). Cambridge University Press, 356–376.

[3] Suryapratim Banerjee, Hideo Konishi, and Tayfun Sönmez. 2001. Core in a Simple Coalition Formation Game. *Social Choice and Welfare* 18, 1 (2001), 135–153.

[4] John Bartholdi III, Craig A. Tovey, and Michael A. Trick. 1992. How hard is it to control an election? *Mathematical and Computer Modelling* 16, 8/9 (1992), 27–40.

[5] Niclas Boehmer, Robert Bredereck, Klaus Heeger, and Rolf Niedermeier. 2021. Bribery and Control in Stable Marriage. *Journal of Artificial Intelligence Research* 71 (2021), 993–1048.

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

[7] Felix Brandt, Martin Bullinger, and Leo Tappe. 2024. Stability based on single-agent deviations in additively separable hedonic games. *Artificial Intelligence* 334 (2024), 104160.

[8] Martin Bullinger, Edith Elkind, and Jörg Rothe. 2016. Cooperative Game Theory. In *Economics and Computation, An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Jörg Rothe (Ed.). Springer, 135–193.

[9] Katarína Cechlárová and Antonio Romero-Medina. 2001. Stability in coalition formation games. *International Journal of Game Theory* 29 (2001), 487–494.

[10] Georgios Chalkiadakis, Edith Elkind, and Michael J. Wooldridge. 2011. *Computational Aspects of Cooperative Game Theory.* Morgan & Claypool Publishers.

[11] Jiehua Chen, Gergely Csáji, Sanjukta Roy, and Sofia Simola. 2023. Hedonic Games With Friends, Enemies, and Neutrals: Resolving Open Questions and Fine-Grained Complexity (AAMAS 2023). 251–259.

[12] Jiehua Chen, Jakob Guttmann, Merisa Mustajbašić, and Sofia Simola. 2026. *Control in Hedonic Games.* Technical Report. arXiv:2602.18506 [cs.GT].

[13] Jiehua Chen, Christian Hatschka, and Sofia Simola. 2026. Computational Social Choice: Parameterized Complexity and Challenges. *Computer Science Review* 61 (2026).

[14] Jiehua Chen, Joanna Kaczmarek, Paul Nüsken, Jörg Rothe, Ildikó Schlotter, and Tessa Seeger. 2025. Control in Computational Social Choice. In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI 2025).* 10391–10399.

[15] Jiehua Chen and Ildikó Schlotter. 2025. *Control in Stable Marriage and Stable Roommates: Complexity and Algorithms.* Technical Report. arXiv:2502.01215.

[16] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. 2015. *Parameterized Algorithms.* Springer.

[17] Dinko Dimitrov, Peter Borm, Ruud Hendrickx, and Shao Chin Sung. 2006. Simple priorities and core stability in hedonic games. *Social Choice and Welfare* 26, 2 (2006), 421–433.

[18] Rodney G. Downey and Michael R. Fellows. 2013. *Fundamentals of Parameterized Complexity.* Springer.

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

[20] Piotr Faliszewski and Jörg Rothe. 2016. Control and Bribery in Voting. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia (Eds.). Cambridge University Press, Chapter 7, 146–168.

[21] Jon Feldman and Matthias Ruhl. 2006. The Directed Steiner Network Problem is Tractable for a Constant Number of Terminals. *SIAM J. Comput.* 36, 2 (2006), 543–561.

[22] Robert W. Floyd. 1962. Algorithm 97: Shortest path. *Commun. ACM* 5, 6 (June 1962), 345.

[23] Steven D. Galbraith. 2012. *Mathematics of Public Key Cryptography.* Cambridge University Press.

[24] Teofilo F. Gonzalez. 1985. Clustering to Minimize the Maximum Intercluster Distance. *Theoretical Computer Science* 38 (1985), 293–306.

[25] Edith Hemaspaandra, Lane A. Hemaspaandra, and Jörg Rothe. 2007. Anyone but Him: The Complexity of Precluding an Alternative. *Artificial Intelligence* 171, 5-6 (2007), 255–285.

[26] Chung-Lun Li, S Thomas McCormick, and David Simchi-Levi. 1992. The point-to-point delivery and connection problems: complexity and algorithms. *Discrete Applied Mathematics* 36, 3 (1992), 267–292.

[27] Rolf Niedermeier. 2006. *Invitation to Fixed-Parameter Algorithms.* Oxford University Press.

[28] Dominik Peters. 2017. Precise complexity of the core in dichotomous and additive hedonic games. In *Proceedings of the International Conference on Algorithmic Decision Theory (ADT 2017).* Springer, 214–227.

[29] Shao-Chin Sung and Dinko Dimitrov. 2010. Computational Complexity in Additive Hedonic Games. *European Journal of Operational Research* 203, 3 (2010), 635–639.

[30] Gerhard J. Woeginger. 2013. Core Stability in Hedonic Coalition Formation. In *Proceedings of the 39th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM 2013) (Lecture Notes in Computer Science, Vol. 7741).* 33–50.

[31] Gerhard J. Woeginger. 2013. A hardness result for core stability in additive hedonic games. *Mathematical Social Science* 65, 2 (2013), 101–104.