Strategyproof Mechanisms for Additively Separable and Fractional Hedonic Games

Michele Flammini   MICHELE.FLAMMINI@GSSI.IT
Gran Sasso Science Institute
L'Aquila, Italy

Bojana Kodric   BOJANA.KODRIC@GSSI.IT
Gran Sasso Science Institute
L'Aquila, Italy

Gianpiero Monaco   GIANPIERO.MONACO@UNIVAQ.IT
University of L'Aquila
L'Aquila, Italy

Qiang Zhang   CSQZHANG@GMAIL.COM
Borealis AI
Toronto, Canada.

Abstract

Additively separable hedonic games and fractional hedonic games have received considerable attention in the literature. They are coalition formation games among selfish agents based on their mutual preferences. Most of the work in the literature characterizes the existence and structure of stable outcomes (i.e., partitions into coalitions) assuming that preferences are given. However, there is little discussion of this assumption. In fact, agents receive different utilities if they belong to different coalitions, and thus it is natural for them to declare their preferences strategically in order to maximize their benefit. In this paper we consider strategyproof mechanisms for additively separable hedonic games and fractional hedonic games, that is, partitioning methods without payments such that utility maximizing agents have no incentive to lie about their true preferences. We focus on social welfare maximization and provide several lower and upper bounds on the performance achievable by strategyproof mechanisms for general and specific additive functions. In most of the cases we provide tight or asymptotically tight results. All our mechanisms are simple and can be run in polynomial time. Moreover, all the lower bounds are unconditional, that is, they do not rely on any computational complexity assumptions.

1. Introduction

Teamwork and coalition or group formation has been an important and widely investigated issue in computer science research. In many economic, social and political situations, individuals carry out activities in groups rather than by themselves. In these scenarios, it is of crucial importance to consider the satisfaction of the members of the groups. For example, the utility of an individual in a group sharing a resource depends both on the consumption level of the resource and on the identity of the members in the group; similarly, the utility for a party belonging to a political coalition depends both on the party program and on the identity of its members.

Hedonic games, introduced by Dréze and Greenberg (1980), model the formation of coalitions (groups) of agents. These are games in which agents have preferences over the set of all possible agent coalitions, and the utility of an agent merely depends on the composition of the coalition she belongs to.

In this paper we consider *additively separable hedonic games* (ASHGs), which constitute a natural and succinctly representable class of hedonic games. Each agent in an ASHG has a value for every other agent, and the utility she ascribes to a given coalition is simply the sum of the values she assigns to its members. Additive separability satisfies a number of desirable properties (Aziz, Brandt, & Seedig, 2013b) and ASHG are equivalent to the non-transferable utility generalization of graph games studied by Deng and Papadimitriou (1994). They have also inspired several related coalition formation models. Among these, we further consider *fractional hedonic games* (FHGs), introduced by Aziz, Brandl, Brandt, Harrenstein, Olsen, and Peters (2019), which are similar to ASHGs, with the difference that the utility of each agent is divided by the size of her coalition. This allows to model behavioral dynamics in social environments that are not captured by ASHGs: one usually prefers having a couple of good friends in a coalition composed by few other people rather than being part of a crowded coalition populated by uninteresting agents. FHGs are well-suited to model, for instance, the formation of social groups or of political parties.

Coalition formation in ASHGs and FHGs has received considerable attention from the perspective of coalition stability, i.e., core, Nash equilibria, and other suitable equilibrium notions, or from a classical offline optimization point of view in which solutions are not necessarily stable (see the Related Work section), but rather try to maximize suitable social welfare criteria. However, little emphasis has been placed onto the mechanism design setting in which agents have private preferences. A major challenge in this respect is to design algorithms that work well even when the input is reported by selfish agents aiming only at maximizing their personal utility. An interesting approach is to use strategyproof mechanisms (Dughmi & Ghosh, 2010; Procaccia & Tennenholtz, 2013), that is, algorithms (not using payments) where agents have no incentive to lie about their true preferences for increasing their personal benefit.

1.1 Our Contribution

We present strategyproof mechanisms for ASHGs and FHGs, both for general and for specific valuation functions. In particular, we consider: i) *general valuations* where the additive valuations among agents can get any values; ii) *non-negative valuations*, where they can only get positive values; iii) *duplex valuations*, where they can only belong to the set $\{-1, 0, 1\}$ (each agent i can express for any other agent j if she is an enemy, neutral or a friend), and finally iv) *simple valuations*, where they can only take values in $\{0, 1\}$ (each agent can express only a sort of "I like" opinion for any other agent). The latter setting has been also considered in other papers, since it models a basic economic scenario referred to in the literature as Bakers and Millers (Aziz et al., 2019; Bilò, Fanelli, Flammini, Monaco, & Moscardelli, 2015, 2018). See Section 2 for more details about the considered valuations.

We focus on the classical utilitarian social welfare, that is the sum of individual utilities of the agents for their assigned coalitions, and provide several lower and upper bounds on the performance achievable by strategyproof mechanisms. To this aim, we start by

providing the result of Theorem 1, which shows an upper bound of 2 for the approximation ratio achieved by a maximum weight matching on symmetric FHGs. This result is both interesting in its own right, improving over the result of Aziz, Gaspers, Gudmundsson, Mestre, and Täubig (2015), and also used later on in the paper. Then, for general valuations, in Theorem 2 we prove that there is no randomized strategyproof mechanism always returning a solution with non-negative expected social welfare and with a bounded approximation ratio both for ASHGs and FHGs, even for the simple case of only 3 agents. For non-negative valuations, we note that the mechanism returning the grand coalition is strategyproof and optimal for ASHGs, while Proposition 1 shows that the approximation ratio of the same mechanism is $n/2$ for FHGs. In Theorem 3, we further prove that no deterministic strategyproof mechanism can achieve an approximation ratio better than $n/2$ for FHGs, meaning that the mechanism returning the grand coalition is in fact the best possible one. For duplex valuations, we present several lower bounds for ASHGs and FHGs. Theorem 4 and Theorem 5 state that no deterministic mechanism has approximation ratio less than $n - 2$ for ASHGs or $2 - \epsilon$ for FHGs, respectively. Theorem 6 and Theorem 7 show that no randomized strategyproof mechanism can achieve an approximation of $2 - \epsilon$ for ASHGs or $3/2 - \epsilon$ for FHGs, respectively. Such lower bounds are proven via a suitable example in which one agent is valued positively by almost all the other agents, making it harder for the mechanisms to be immune to her possible manipulations. On the positive side, we present a deterministic strategyproof mechanism $\mathcal{M}_3$ with approximation $O(n^2)$ for ASHGs and $O(n)$ for FHGs (Theorem 8). The mechanism returns pairs of agents and puts all other agents that it did not manage to pair into singleton coalitions. Furthermore, we give a randomized strategyproof mechanism $\mathcal{M}_4$ with approximation $O(n)$ for ASHGs (Theorem 9), which tries to imitate a randomly chosen perfect or near-perfect matching on the whole set of agents, and a randomized strategyproof mechanism $\mathcal{M}_5$ with approximation 8 for FHGs (Theorem 10), which is a variant of $\mathcal{M}_3$ that considers the agents in an order chosen uniformly at random from the set of all possible orderings of n elements. For simple valuations, Theorem 11 first proves that no deterministic strategyproof mechanism can achieve approximation ratio less than $6/5$ for FHGs, by considering a simple example with 7 agents in a cycle, and then Theorem 12 shows that returning a maximum matching yields a strategyproof mechanism with approximation ratio 2. Both Theorem 10 and Theorem 12 use Theorem 1 from Section 3 as a building block.

Our results are summarized in Table 1. As it can be noticed, in most of the cases (except in the case of duplex valuations) we provide tight or asymptotically tight results. We remark that the lower bounds on the performance of randomized strategyproof mechanisms also hold for deterministic ones, and similarly, the upper bounds on the performance of deterministic strategyproof mechanisms also hold for randomized ones.

We finally point out that, while on the one hand all our mechanisms are simple and can be run in polynomial time, on the other hand all lower bounds (some of them randomized) are unconditional, that is, they do not rely on any computational complexity assumptions.

1.2 Related Work

The optimization problem of partitioning agents into coalitions so as to maximize the social welfare is a major research challenge in AI and it has been extensively investigated in

the field of multi-agent systems under the name of Coalition Structure Generation (CSG). Many works characterize the computational complexity of finding optimal solutions, focus on providing efficient algorithms, prove hardness results and give suitable approximations under different assumptions or variants of the problem (Rahwan, Michalak, Wooldridge, & Jennings, 2015).

In the literature, a significant stream of research considers hedonic games (see Aziz & Savani, 2016), and in particular ASHGs, from a strategic cooperative point of view (Banerjee, Konishi, & Sönmez, 2001; Bogomolnaia & Jackson, 2002; Elkind & Wooldridge, 2009), with the purpose of characterizing the existence and the properties of stable coalition structures such as the core, and from a non-cooperative point of view (Bloch & Diamantoudi, 2011; Feldman, Lewin-Eytan, & Naor, 2015), with special focus on pure Nash equilibria. Computational complexity issues related to the problem of computing stable outcomes have been considered by Aziz et al. (2013b), Gairing and Savani (2019), Peters (2016), Peters and Elkind (2015) and Woeginger (2013). Concerning FHGs, Aziz et al. (2019) give some properties guaranteeing the (non-)emptiness of the core. Moreover, Brandl, Brandt, and Strobel (2015) study the computational complexity of understanding the existence of core and individually stable outcomes. Finally, Carosi, Monaco, and Moscardelli (2019) study the local core stability in FHGs. From a non-cooperative point of view, Bilò et al. (2018) study the existence, efficiency and computational complexity of Nash equilibria. Further results on the price of stability for specific FHGs have been presented by Peters (2016) and for some simple graph-based hedonic games by Kaklamanis, Kanellopoulos, Papaioannou, and Patouchas (2021). Other stability notions have also been investigated, for example Aziz, Brandt, and Harrenstein (2013a), Elkind, Fanelli, and Flammini (2020) focus on Pareto optimality. Flammini, Monaco, Moscardelli, Shalom, and Zaks (2018) consider the online scenario for both ASHGs and FHGs. Olsen (2012) considers a variant of FHGs called modified fractional hedonic games, where the utility of each agent in a coalition structure is equal to the sum of the weights of the incident edges in the coalition she belongs to, divided by the size of the coalition minus 1. Monaco, Moscardelli, and Velaj (2020) consider Nash and core stable outcomes for modified fractional hedonic games and provide bounds on their performance. Finally, Bullinger (2020) gives algorithms for finding Pareto optimal solutions in ASHGs, FHGs and modified FHGs.

The design of truthful mechanisms *with money*, that is of algorithms that use payments to convince the selfish agents to reveal the truth and then compute the outcome on the basis of their reported values, has been studied in numerous scenarios. However, there are settings where monetary transfers are not feasible, because of either ethical or legal issues (Nisan, Roughgarden, Tardos, & Vazirani, 2007), or practical matters in enforcing and collecting payments (Procaccia & Tennenholtz, 2013). A growing stream of research thus focuses on the design of the more applicable truthful mechanisms *without money*, often called strategyproof mechanisms, that lead agents to report their true preferences without resorting to payments.

Along these lines, Wright and Vorobeychik (2015) investigate strategyproof mechanisms for ASHGs. They only consider positive preferences. Under this assumption, a trivial optimal strategyproof mechanism just puts all the agents in the same grand coalition. Therefore, they assume coalition size constraints and (approximate) envy-freeness. Their main contribution is a mechanism that, despite not having theoretical guarantees, achieves a good

experimental performance. We point out that in this paper we focus on theoretical results concerning ASHGs and FHGs, for which, to the best of our knowledge, no strategyproof mechanism has been proposed in the scientific literature, yet.

Vallée, Bonnet, Zanuttini, and Bourdon (2014) consider classical hedonic games with general preference relationships, and characterize the conditions on the game structure that allow rational false-name manipulations. However, they do not provide mechanisms. Aziz et al. (2013a) show that the serial dictatorship mechanism is Pareto optimal and strategyproof for general hedonic games when appropriate restrictions are imposed onto the agents. Finally, Rodríguez-Álvarez (2009) studies properties of strategyproof mechanisms that return core stable solutions for hedonic games.

1.3 Paper Organization

The paper is organized as follows. In Section 2, we formally describe the problems and introduce some useful definitions. In Section 3, we present a result that is somewhat orthogonal to the rest of the paper, as it gives an approximation bound for a non-truthful mechanism. This result, in fact, is used later in the paper. The studies of the performance of strategyproof mechanisms are presented in Section 4, 5, 6, and 7, which address, respectively, general, non-negative, duplex and simple valuations. Finally, in Section 8, we provide some concluding remarks and list interesting open problems.

		[-1,1]	[0,1]	{-1,0,1}	{0,1}
ASHGs	L.B.	Unbounded*	1	$\Omega(n), 2 - \epsilon^*$	1
	U.B.			$O(n^2), O(n)^*$	
FHGs	L.B.	Unbounded*	$\frac{n}{2}$	$2 - \epsilon, \frac{3}{2} - \epsilon^*$	$\frac{6}{5}$
	U.B.		$\frac{n}{2}$	$O(n), 8^*$	2

Table 1: Our results for the different cases. Randomized mechanisms are denoted by *. L.B. stands for lower bounds. U.B. stands for upper bounds.

2. Preliminaries

In additively separable hedonic games (ASHGs) and fractional hedonic games (FHGs), we are given a set $N = \{1, \ldots, n\}$ of selfish agents. The outcome of the game is a partition of the agents into disjoint coalitions $\mathcal{C} = \{C_1, C_2, \ldots\}$, where each coalition C_j is a subset of agents and each agent is in exactly one coalition. Let $\mathscr{C}$ be the collection of all the possible outcomes. Given a partition $\mathcal{C} \in \mathscr{C}$, we denote by $|\mathcal{C}|$ the number of its coalitions and by $\mathcal{C}^i$ the coalition of $\mathcal{C}$ containing agent i. Similarly, given a coalition C, we let $|C|$ be the size or number of agents in C. The *grand coalition* is the outcome in which all the agents are in the same coalition, i.e., $|\mathcal{C}| = 1$. A *singleton coalition* is any coalition C such that

$|C| = 1$. We assume that each agent has a privately known valuation $v_i : N \to \mathbb{R}$, mapping every agent to a real (possibly negative) value. In ASHGs, for any $\mathcal{C} \in \mathscr{C}$, the preference or utility of agent i is $u_i(\mathcal{C}) = \sum_{j \in \mathcal{C}^i} v_i(j)$, that is, it is additively induced by her valuation function. Similarly, in FHGs, for any $\mathcal{C} \in \mathscr{C}$, the utility of agent i is $u_i(\mathcal{C}) = \frac{\sum_{j \in \mathcal{C}^i} v_i(j)}{|\mathcal{C}^i|}$.

We are interested in four basic classes of valuation functions. Namely, for any pair of agents $i, j \in N$, we consider:

- *General valuations*: $v_i(j) \in [-1, 1]$;

- *Non-negative valuations*: $v_i(j) \in [0, 1]$;

- *Duplex valuations*: $v_i(j) \in \{-1, 0, 1\}$;

- *Simple valuations*: $v_i(j) \in \{0, 1\}$.

Furthermore, a valuation function is *symmetric* if and only if $v_i(j) = v_j(i)$, for any $i, j \in N$. We always assume that $v_i(i) = 0$ for every $i \in N$. Notice that any valuation function can be represented by using values in the range $[-1, 1]$.

2.1 Graph Representation

ASHGs and FHGs have a very intuitive graph representation. In fact, any instance of these games can be expressed by a weighted directed graph $G = (V, E, w)$, where nodes in V represent the agents, and arcs or directed edges are associated with non null valuations. Namely, if $v_i(j) \neq 0$, then E contains the arc (i, j) of weight $w(i, j) = v_i(j)$. As an example, in case of simple valuations, if $(i, j) \notin E$ then $v_i(j) = 0$, while if $(i, j) \in E$ then $w(i, j) = v_i(j) = 1$. Throughout the paper we will sometimes describe an instance of the considered game by its graph representation.

Given a weighted graph $G = (V, E, w)$, where w is the weight function of the edges, we denote by $w(E)$ the sum of the weights of the edges belonging to E, i.e., $w(E) = \sum_{\{i,j\} \in E} w(i, j)$.

2.2 Mechanisms

Agents are self-interested entities. Thus, they may strategically misreport their valuation functions in order to maximize their utilities. Let $\mathbf{d}$ denote the preferences (valuation functions) declared by all the agents.

A deterministic mechanism $\mathcal{M}$ maps every set (or list) of preferences $\mathbf{d}$ to a set of disjoint coalitions $\mathcal{M}(\mathbf{d}) \in \mathscr{C}$. We denote by $\mathcal{M}^i(\mathbf{d})$ the coalition assigned to agent i by $\mathcal{M}$. The utility of agent i is given by $u_i(\mathcal{M}(\mathbf{d}))$. Let $\mathbf{d}_{-i}$ be the valuation functions declared by all agents except agent i and let d_i be a possible declaration of valuation function by i. A deterministic mechanism $\mathcal{M}$ is *strategyproof* if for any $i \in N$, any list of preferences $\mathbf{d}_{-i}$, any v_i and any d_i, it holds that $u_i(\mathcal{M}(\mathbf{d}_{-i}, v_i)) \geq u_i(\mathcal{M}(\mathbf{d}_{-i}, d_i))$. In other words, a strategyproof mechanism prevents any agent i from benefiting by declaring a valuation different from v_i, whatever the other declared valuations are.

A randomized mechanism $\mathcal{M}$ maps every set of agents' preferences $\mathbf{d}$ to a distribution Δ over the set of all the possible outcomes $\mathscr{C}$. The expected utility of agent i is given by

$\mathbb{E}[u_i(\mathcal{M}(\mathbf{d}))] = \mathbb{E}_{\mathcal{C} \sim \Delta}[u_i(\mathcal{C})]$. A randomized mechanism $\mathcal{M}$ is *universally strategyproof* if it is a distribution over deterministic strategyproof mechanisms. This means that the mechanism is strategyproof even if the outcomes of the random coin flips are known to the agents. On the other hand, a randomized mechanism $\mathcal{M}$ is *strategyproof in expectation* if for any $i \in N$, any preferences $\mathbf{d}_{-i}$, any v_i and any d_i, $\mathbb{E}[u_i(\mathcal{M}(\mathbf{d}_{-i}, v_i))] \geq \mathbb{E}[u_i(\mathcal{M}(\mathbf{d}_{-i}, d_i))]$. Note that universal strategyproofness implies strategyproofness in expectation. In this paper, all presented randomized mechanisms will be universally strategyproof and all lower bounds for strategyproof randomized mechanisms will hold for strategyproofness in expectation.

We say that a deterministic mechanism $\mathcal{M}$ is *acceptable* if it always guarantees a non negative sum of the players' utilities (i.e., non negative social welfare, as defined in the following paragraph), i.e., $\sum_{i=1}^{n} u_i(\mathcal{M}(\mathbf{d})) \geq 0$ for any possible list of preferences $\mathbf{d}$. Similarly, a randomized mechanism $\mathcal{M}$ is *acceptable* if $\mathbb{E}[\sum_{i=1}^{n} u_i(\mathcal{M}(\mathbf{d}))] \geq 0$ holds for every $\mathbf{d}$. In the following, we will always implicitly restrict our attention to acceptable mechanisms. In fact, a simple acceptable strategyproof mechanism for all the considered classes of valuations can be trivially obtained by putting every agent into a separate singleton coalition, regardless of all the declared valuations.

2.3 Mechanism Performance

We are interested in strategyproof mechanisms that perform well with respect to the goal of maximizing the classical utilitarian social welfare, that is, the sum of the utilities achieved by all the agents. Namely, the social welfare of a given outcome $\mathcal{C}$ is $\mathrm{SW}(\mathcal{C}) = \sum_{i \in N} u_i(\mathcal{C})$. We denote by $\mathrm{SW}(C) = \sum_{i \in C} u_i(\mathcal{C})$ the overall social welfare achieved by the agents belonging to a given coalition C. We measure the performance of a mechanism by comparing the social welfare it achieves with the optimal one. More precisely, the approximation ratio of a deterministic mechanism $\mathcal{M}$ is defined as

$$r^{\mathcal{M}} = \sup_{\mathbf{d}} \frac{\mathrm{OPT}(\mathbf{d})}{\mathrm{SW}(\mathcal{M}(\mathbf{d}))},$$

where $\mathrm{OPT}(\mathbf{d})$ is the social welfare achieved by an optimal set of coalitions in the instance induced by $\mathbf{d}$. For randomized mechanisms, the approximation ratio is computed with respect to the expected social welfare, that is,

$$r^{\mathcal{M}} = \sup_{\mathbf{d}} \frac{\mathrm{OPT}(\mathbf{d})}{\mathbb{E}[\mathrm{SW}(\mathcal{M}(\mathbf{d}))]}.$$

If a set of preferences $\mathbf{d}$ is such that $\mathrm{OPT}(\mathbf{d}) = 0$, any mechanism $\mathcal{M}$ returns an optimal coalition structure. Therefore, we define the approximation ratio for such an instance to be 1. On the other hand, for a fixed mechanism $\mathcal{M}$, if there exists a set of preferences $\mathbf{d}$ such that the (expected) social welfare of the outcome $\mathcal{M}(\mathbf{d})$ is 0, while $\mathrm{OPT}(\mathbf{d}) \neq 0$, we say that the approximation ratio $r^{\mathcal{M}}$ is unbounded.

In all of the presented mechanisms, possible ties will be resolved by considering the ordering in which the mechanism is processing the agents. More specifically, all of the mechanisms that do not return the grand coalition, will be creating coalitions of size at most 2. Therefore, when a possible tie appears, i.e., the i-th agent could be placed in a coalition both with the j-th and the k-th agent, the mechanism will pair it with the j-th agent if and only if $j < k$.

3. Maximum Matching Approximation for Symmetric FHGs

Even though we are generally interested in strategyproof mechanisms, before starting to analyze the aforementioned valuation function classes, in this section we prove a result of somewhat different flavor. Namely, we give a bound on the approximation ratio achieved by a maximum weight matching mechanism in symmetric fractional hedonic games. We will use this result in Sections 6 and 7 for giving bounds on truthful mechanisms, but we also believe it has applicability beyond mechanism design.

First, notice that every weighted undirected graph $G = (V, E, w)$ where $w : E \to \mathbb{R}$ corresponds to an instance of a fractional hedonic game with a symmetric valuation function (for any two agents i and j, it holds that $v_i(j) = v_j(i)$) and vice versa. We can show that the outcome induced by a maximum weight matching in an undirected weighted graph G is a 2-approximation of the maximum social welfare of the corresponding symmetric FHG. This improves the result of Aziz et al. (2015), where a bound of 4 is proven. We state the matching mechanism in terms of undirected weighted graphs.

Mechanism $\mathcal{M}_1$. *Given an undirected weighted graph $G = (V, E, w)$ where $w : E \to \mathbb{R}$, the mechanism performs as follows:*
1 Create a complete graph $G' = (V, E', w')$ by adding edges of weight 0 to G.
2 Consider any fixed numbering of the nodes in V, where $|V| = n$, and represent each matching as a binary vector $(x_{12}, x_{13}, \ldots, x_{23}, x_{24}, \ldots, x_{n-1n})$ in $\{0, 1\}^{\binom{n}{2}}$, and let $\prec$ be the lexicographic order on these vectors.
3 Return the $\prec$-minimal matching from the set $\mathrm{argmax}_{M \in \mathcal{M}} \sum_{\{i,j\} \in M} w'(i,j)$.

We note here that Dughmi and Ghosh (2010) showed that such a $\prec$-minimal matching can be found in polynomial time. This is precisely the reason why we introduce the binary vector representation of matchings and use it as a tie-breaking rule in $\mathcal{M}_1$. A run of the matching mechanism on an example instance can be seen in Figure 1.
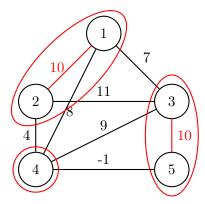

Figure 1: The outcome of Mechanism $\mathcal{M}_1$ on an example instance. The edges chosen by the mechanism and the corresponding coalitions are depicted in red. The solution returned by the mechanism has value 20, while the optimum, achieved by the coalition structure $\{\{1, 2, 4\}, \{3, 5\}\}$, is $\frac{74}{3}$.

Theorem 1. *For symmetric fractional hedonic games, returning the coalitions induced by Mechanism $\mathcal{M}_1$ is a 2-approximation of the maximum social welfare.*

Proof. Let m be the matching computed by Mechanism $\mathcal{M}_1$ and $\mathcal{C}^m$ the coalitions induced by m. Let $\mathcal{C}^* = \{C_1^*, \ldots, C_p^*\}$ be an optimal coalition structure (we assume all coalitions in $\mathcal{C}^*$ to have strictly positive social welfare, indeed we can ignore coalitions in $\mathcal{C}^*$ whose social welfare is zero). Let $m' = m_1' \cup \ldots \cup m_p'$ where m_h', $1 \leq h \leq p$, is a maximum matching in the graph induced by the vertices of C_h^*. Let $\mathcal{C}^{m'}$ be the coalitions induced by m'. Let A_h be the vertices matched in m_h' and $B_h = C_h^* \setminus A_h$. Notice that B_h is an independent set in C_h^* and that $|A_h|$ is an even number.

Lemma 1. *For any $h = 1, \ldots, p$ and any edge $\{i,j\} \in m_h'$, if $|B_h| > 0$, then*

$$\sum_{b \in B_h} (w(i,b) + w(j,b)) \leq w(i,j)(|B_h| + 1) \ .$$

Proof. First notice that, for any $b \in B_h$, it holds that $w(i,b) \leq w(i,j)$ and $w(j,b) \leq w(i,j)$, since otherwise we can get a better matching by removing the edge $\{i,j\}$ from m_h' and adding a new edge having weight strictly greater than $w(i,j)$. We now distinguish two cases depending on the size of B_h. If $|B_h| = 1$, then the claim easily follows from the observation that $w(i,b) \leq w(i,j)$ and $w(j,b) \leq w(i,j)$. If $|B_h| > 1$, then suppose that $\sum_{b \in B_h}(w(i,b) + w(j,b)) > w(i,j)(|B_h| + 1)$. This implies that there are two distinct edges $\{i,b\}$ and $\{j,b'\}$ for some $b, b' \in B_h, b \neq b'$ such that $w(i,b) + w(j,b') > w(i,j)$, which contradicts the fact that m_h' is a maximum matching in C_h^*. $\square$

Let $\widehat{E_h}$ be the set of edges of the graph induced by the vertices of A_h minus the edges belonging to the matching m_h'. Moreover, let $w(\widehat{E_h}) = \sum_{\{i,j\} \in \widehat{E_h}} w(i,j)$.

Lemma 2. *For any $h = 1, \ldots, p$, it holds that $w(\widehat{E_h}) \leq w(m_h')(|A_h| - 2)$.*

Proof. Let us consider the graph G_{A_h} induced by the vertices of A_h and suppose that G_{A_h} is complete (if it is not complete, we can simply add edges with weight zero). Note that there exists a set of $|A_h| - 1$ disjoint perfect matchings on G_{A_h} that contains m_h' (see, e.g., Anderson, 2001, Theorem 8.1). Let us denote this set by M. Then, notice that $\frac{1}{|A_h|-1} \sum_{\{i,j\} \in A_h} w(i,j)$ is the average weight of a matching from M. Therefore, it has to hold that $\sum_{\{i,j\} \in A_h} w(i,j) \leq w(m_h') \cdot (|A_h| - 1)$. The claim now follows by using the fact that $w(\widehat{E_h}) = \sum_{\{i,j\} \in A_h} w(i,j) - w(m_h')$. $\square$

Then, when $|B_h| > 0$, by using Lemma 1 and Lemma 2, we can bound the social welfare of C_h^*, for any $h = 1, \ldots, p$:

$$\begin{aligned}
\text{SW}(C_h^*) &= \\
&= \frac{1}{|C_h^*|} \left(\sum_{\{i,j\} \in m_h'} (w(i,j) + \sum_{b \in B_h} w(i,b) + w(j,b)) + w(\widehat{E_h}) \right) \\
&\leq \frac{1}{|C_h^*|} \left(w(m_h') + w(m_h')(|B_h| + 1) + w(m_h')(|A_h| - 2) \right) \\
&= w(m_h').
\end{aligned}$$

Similarly, when $|B_h| = 0$ we can see that $SW(C_h^*) \leq w(m_h')$. Therefore, overall we have that $SW(\mathcal{C}^*) \leq w(m')$. Since by the choice of m we have $w(m) \geq w(m')$, then the social welfare of $\mathcal{C}^m$ is

$$\mathrm{SW}(\mathcal{C}^m) = \frac{w(m)}{2} \geq \frac{w(m')}{2} \geq \frac{\mathrm{SW}(\mathcal{C}^*)}{2}.$$

$\square$

Computing the maximum social welfare in symmetric fractional hedonic games is an NP-hard problem (Aziz et al., 2015), and we are not aware of stronger inapproximability results. Therefore, the 2-approximation algorithm of Theorem 1 is not tight in this sense. Lastly, we note that 2 is the best approximation achievable by using matchings when dealing with the problem of computing the maximum social welfare in symmetric fractional hedonic games. This can be easily verified by considering a complete graph on n nodes. In the grand coalition, each node has utility $\frac{n-1}{n}$ (consider large n), while in a matching each node has a utility of at most $\frac{1}{2}$.

4. General Valuations

In this section, we return to our main topic of strategyproof mechanisms and consider the setting where agents have general valuations. We are able to prove that there is no randomized strategyproof mechanism with bounded approximation ratio both for ASHGs and FHGs. Clearly, the theorem applies also to deterministic mechanisms, since they are special cases of randomized ones.

Theorem 2. *For general valuation functions, there is no randomized strategyproof acceptable mechanism with bounded approximation ratio both for ASHGs and FHGs.*

Proof. We first prove the theorem for ASHGs and then show that the same arguments directly apply also to FHGs.

(a) Instance I_1 (b) Instance I_2

Figure 2: The lower bound instance for general valuations.

Let $\mathcal{M}$ be a given randomized strategyproof mechanism. Provided that $\mathcal{M}$ is strategyproof, we implicitly assume that the agents' declared preferences $\mathbf{d}$ correspond to the true valuation functions. Let us then consider the instance I_1 depicted in Figure 2a, and let p be the probability that $\mathcal{M}$ returns an outcome for I_1 where agents 2 and 3 are together in the same coalition. Then, the expected social welfare is

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}(\mathbf{d}^1))] \leq p(\epsilon - 0.1) + (1 - p)\epsilon = \epsilon - 0.1p \,,$$

while the optimal solution has social welfare ϵ (notice that the presence of agent 1 is necessary to guarantee a strictly positive value of the optimal outcome in instance I_1). Therefore,

the randomized mechanism has bounded approximation ratio only if $\epsilon - 0.1p > 0$; that implies $p < 10\epsilon$. Let us now consider the instance I_2 depicted in Figure 2b, and let q be the probability that mechanism $\mathcal{M}$ returns an outcome where agents 2 and 3 are together in the same coalition. Then the expected social welfare is

$$\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}^2))] \leq 0.9q + (1-q)\epsilon \ .$$

We notice that $\mathcal{M}$ can be strategyproof only if $p \geq q$, otherwise agent 2 in I_2 could improve her utility by declaring value -1 for agent 3, since in such a case she would get utility $-p\epsilon > -q\epsilon$. The optimal solution of instance I_2 has value 0.9. Thus, the approximation ratio of $\mathcal{M}$ is

$$\sup_{\mathbf{d}} \frac{\text{OPT}(\mathbf{d})}{\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}))]} \geq \frac{\text{OPT}(\mathbf{d}^2)}{\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}^2))]} \geq \frac{0.9}{0.9q + (1-q)\epsilon} \geq \frac{0.9}{0.9p + \epsilon} > \frac{0.9}{0.9 \cdot 10\epsilon + \epsilon} = \frac{0.9}{10\epsilon} \ .$$

As ϵ can be arbitrarily small, we can then conclude that $\mathcal{M}$ has an unbounded approximation ratio. The claim then follows by the arbitrariness of $\mathcal{M}$.

Now, by using the same notation and the same arguments for FHGs, from instance I_1 we learn that the expected social welfare of $\mathcal{M}$ is

$$\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}^1))] \leq p \cdot \frac{\epsilon - 0.1}{3} + (1-p) \cdot \frac{\epsilon}{2} \ .$$

The optimal solution now has social welfare $\frac{\epsilon}{2}$ but we conclude the same as before that for a bounded approximation ratio it has to hold that $\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}^1))] > 0$, which is equivalent to $p < \frac{15\epsilon}{5\epsilon+1}$. Furthermore, it again has to hold that $p \geq q$ in order for $\mathcal{M}$ to be strategyproof. Finally, from instance I_2 and being interested in $0 < \epsilon < 0.3$ such that $\frac{0.9-\epsilon}{2} > \frac{0.9}{3}$, we conclude that the approximation ratio of $\mathcal{M}$ is

$$\sup_{\mathbf{d}} \frac{\text{OPT}(\mathbf{d})}{\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}))]} \geq \frac{\text{OPT}(\mathbf{d}^2)}{\mathbb{E}[\text{SW}(\mathcal{M}(\mathbf{d}^2))]} \geq \frac{0.3}{q\frac{0.9-\epsilon}{2} + (1-q)\frac{\epsilon}{2}} \geq \frac{0.6}{q(0.9 - 2\epsilon) + \epsilon} > \frac{6 + 30\epsilon}{145\epsilon - 250\epsilon^2}.$$

Again, $\mathcal{M}$ has an unbounded approximation ratio because ϵ can be arbitrarily small. □

5. Non-Negative Valuations

In this section, we consider the setting where agents have non-negative valuations. Let us first present a simple optimal mechanism for non-negative valuations in ASHGs.

Mechanism $\mathcal{M}_2$. *Given as input a list of agents' valuations* $\mathbf{d} = \langle d_1, ..., d_n \rangle$*, the mechanism outputs the grand coalition, i.e.* $\mathcal{M}(\mathbf{d}) = \{\{1, \ldots, n\}\}$*.*

It is trivial to see that, in ASHGs with non-negative valuations, the above mechanism $\mathcal{M}_2$ is acceptable, strategyproof, and achieves the optimal social welfare. Therefore, we now focus on FHGs. We are able to show that any deterministic strategyproof mechanism cannot have an approximation ratio better than $\frac{n}{2}$.

Theorem 3. *For FHGs with non-negative valuations, no deterministic strategyproof acceptable mechanism can achieve approximation ratio r, with $r < \frac{n}{2}$.*

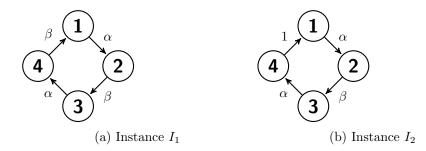(a) Instance I_1 (b) Instance I_2

Figure 3: The lower bound instances for non-negative valuations with 4 agents.

Proof. Assume $\frac{1}{n} \gg \alpha \gg \beta$. Let us consider the instance I_1 (see Figure 3a) with an even number n of agents, where the valuation functions are as follows:

- for any $i = 1, 3, \ldots, n-1$, $v_i(j) = \alpha$ if $j = i+1$ and $v_i(j) = 0$ otherwise;

- for any $i = 2, 4, \ldots, n-2$, $v_i(j) = \beta$ if $j = i+1$ and $v_i(j) = 0$ otherwise;

- $v_n(1) = \beta$ and $v_n(j) = 0$ for any $j \neq 1$.

The optimal outcome is given by the set of coalitions $\mathcal{C} = \{C_1, C_2, \ldots, C_{\frac{n}{2}}\}$, where $C_j = \{2j-1, 2j\}$ for any $j = 1, \ldots, \frac{n}{2}$, and achieves social welfare $\frac{n}{4}\alpha$. We now show that any deterministic strategyproof mechanism with an approximation ratio lower than $\frac{n}{2}$ has to output the grand coalition. In fact, since the grand coalition has social welfare $\frac{\alpha+\beta}{2}$, and its approximation ratio tends to $\frac{n}{2}$ when β/α tends to 0, this is enough to prove the claim.

Assume then that a deterministic strategyproof mechanism $\mathcal{M}$ with an approximation ratio strictly less than $\frac{n}{2}$ outputs an outcome different from the grand coalition. In this case, there must be at least one agent k having null utility, since we can find agents k and $k+1$ that are not in the same coalition. But then, k might try to improve her utility by declaring $v_k(k+1) = 1$, creating instance I_2 (see Figure 3b). Note that in this case in order to achieve an approximation less than $\frac{n}{2}$, $\mathcal{M}$ must return an outcome in which agents k and $k+1$ are in the same coalition. Hence, agent k indeed improves her utility by declaring $v_k(k+1) = 1$. Therefore, to maintain strategyproofness, $\mathcal{M}$ has to output the grand coalition also when the input is I_1. Thus, the theorem is proven. $\qquad\square$

Given the above result, it is easy to show that returning the grand coalition is the best we can do.

Proposition 1. *For FHGs with non-negative valuations, Mechanism $\mathcal{M}_2$ is a deterministic strategyproof acceptable mechanism with approximation ratio $\frac{n}{2}$.*

Proof. As valuations are non-negative and Mechanism $\mathcal{M}_2$ always outputs the grand coalition, the mechanism is clearly acceptable and strategyproof. Let us now focus on its approximation ratio for the social welfare. On the one hand, given any $\mathbf{d}$, $\texttt{OPT}(\mathbf{d}) \leq \frac{1}{2}\sum_{i \in N}\sum_{j \in N} v_i(j)$. This holds because any coalition with positive social welfare in the optimal coalition structure consists of at least two agents. Otherwise, the coalition has zero social welfare since $v_i(i) = 0$ for any $i \in N$. On the other hand, the grand coalition has social welfare equal to $\frac{1}{n}\sum_{i \in N}\sum_{j \in N} v_i(j)$. The approximation ratio follows. $\qquad\square$

6. Duplex Valuations

In this section, we consider the setting where agents have duplex valuations. We first present deterministic lower bounds for ASHGs and FHGs.

Theorem 4. *For ASHGs with duplex valuations, no deterministic strategyproof acceptable mechanism has approximation ratio less than $n - 2$.*

Proof. Let us consider the instance I_1 depicted in Figure 4a, where the valuations of the n agents are as follows:

- for $i = 1, \ldots, n - 2$, $v_i(j) = 1$ if $j = n - 1$ and $v_i(j) = 0$ otherwise;

- $v_{n-1}(j) = 1$ if $j = n$ and $v_{n-1}(j) = -1$ otherwise;

- $v_n(j) = -1$ for $j = 1, \ldots, n - 2$ and $v_n(n - 1) = 0$.
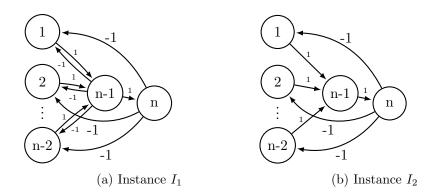

(a) Instance I_1 (b) Instance I_2

Figure 4: The lower bound instance for duplex valuations.

In any optimal outcome agents $n - 1$ and n are in the same coalition, and all other agents are in other coalitions. The resulting social welfare is 1, and in particular it is due to agent $n-1$ having utility 1. It is easy to see that any mechanism having a bounded approximation ratio has to return the optimal outcome, as any other solution would have social welfare of at most zero. Let us now consider the instance I_2 depicted in Figure 4b, where agent $n-1$ is the only one with a different valuation function with respect to I_1, that is $v_{n-1}(n) = 1$ and $v_{n-1}(j) = 0$ for $j \neq n$. Any strategyproof mechanism with bounded approximation ratio for I_2 has to put agents $n - 1$ and n in the same coalition, otherwise $n - 1$ would have zero utility and could increase her utility by declaring her valuation function as it is in instance I_1. Moreover, any outcome in which $n - 1$ and n are together has social welfare 1, because adding any further agent to this coalition does not change its social welfare and the social welfare of all coalitions that do not contain $n-1$ and n is 0. However, the optimal outcome, in which $1, 2, \ldots, n - 1$ are all together in the same coalition and agent n is alone, achieves social welfare $n - 2$. This proves the $n - 2$ lower bound for any deterministic strategyproof mechanism. □

Theorem 5. *For FHGs with duplex valuations, no deterministic strategyproof acceptable mechanism can achieve approximation $2 - \epsilon$, for any $\epsilon > 0$.*

Proof. The proof is analogous to the proof of Theorem 4. Let us again consider the instances depicted in Figure 4. From instance I_1 we see that any mechanism having bounded approximation ratio has to return $n-1$ and n in the same coalition. From instance I_2, on the other hand, we see that returning $n-1$ and n in the same coalition achieves social welfare $\frac{1}{2}$, while the optimal solution puts $1, 2, \ldots, n-1$ all together in the same coalition, achieving social welfare $\frac{n-2}{n-1}$. It follows that by increasing the value of n, the approximation ratio tends to 2, thus proving the theorem. □

We now turn our attention to randomized mechanisms and give lower bounds on the performance of randomized strategyproof mechanisms for ASHGs and FHGs.

Theorem 6. *For ASHGs with duplex valuations, no randomized strategyproof acceptable mechanism can achieve approximation $2 - \epsilon$, for any $\epsilon > 0$.*

Proof. Let us consider the instance I_1 depicted in Figure 4a. Let p be the probability that a randomized mechanism returns the outcome where agents $n-1$ and n are together in the same coalition and no other agent is with them in this coalition. Notice that in such a case agent $n-1$ has expected utility equal to p. Denote the outcome of the randomized mechanism on I_1 by rm . Then the expected social welfare in this case is such that $\mathbb{E}[rm] \leq p$. Let us now consider the instance I_2 depicted in Figure 4b. Let q be the probability that agents $n-1$ and n are together in the same coalition (possibly with other agents). Notice that the social welfare of any outcome where agents $n-1$ and n are together is always 1, independently from the coalitions that the other agents are members of. Moreover, notice that in such a case agent $n-1$ has expected utility equal to q. On the other hand, the mechanism does not put agents $n-1$ and n together in the same coalition with probability $1-q$. In such a case, i.e., with probability $1-q$, the social welfare is at most equal to $n-2$. Denote by rm' the outcome of the randomized mechanism on I_2. It turns out that the expected social welfare in this case is such that $\mathbb{E}[rm'] \leq q + (1-q)(n-2)$. We notice that the mechanism is strategyproof only if $q \geq p$. In fact, if $p > q$, then agent $n-1$ can improve her utility in I_2 by declaring value $v_{n-1}(j) = -1$, for every $j = 1, \ldots, n-2$, and $v_{n-1}(n) = 1$ (thus reconstructing the instance I_1), since in such a case she would get expected utility $p > q$. Therefore, the expected social welfare of the mechanism on I_1 is maximized when $p = q$, because $\mathbb{E}[rm] \leq p \leq q$.

We notice that 1 is the optimal value for the instance depicted in Figure 4a, and $n-2$ is the optimal value for the instance depicted in Figure 4b. Therefore, the expected approximation ratio of the randomized mechanism is $\frac{1}{\mathbb{E}[rm]}$ for the instance depicted in Figure 4a, and $\frac{n-2}{\mathbb{E}[rm']}$ for the instance depicted in Figure 4b. The best approximation ratio that a strategyproof mechanism can achieve is therefore obtained by setting the value of q such as to minimize the maximum of these two quantities. This is achieved by equalizing the expected approximation ratio of the mechanism on both instances (where we, as previously remarked, set $p = q$). We have that

$$\frac{1}{\mathbb{E}[rm]} = \frac{n-2}{\mathbb{E}[rm']} \implies \frac{1}{q} = \frac{n-2}{q+(1-q)(n-2)} \implies q = \frac{q+(1-q)(n-2)}{n-2} \implies \frac{(q-1)(n-2)}{q(3-n)} = 1 \implies$$
$q = \frac{n-2}{2n-5}.$

It follows that for large value of n, q tends to $\frac{1}{2}$ and the approximation ratio tends to 2, thus proving the theorem. □

Theorem 7. *For FHGs with duplex valuations, no randomized strategyproof acceptable mechanism can achieve approximation $3/2 - \epsilon$, for any $\epsilon > 0$.*

Proof. The proof is similar to the proof of Theorem 6. Let us again first consider instance I_1 in Figure 4a and let us denote by p the probability that a randomized mechanism returns an outcome where agents $n-1$ and n are together in a coalition and no other agent is with them in this coalition. Then, we can see that the expected utility of agent $n-1$ is $p/2$ and, if we again denote by rm the outcome of the mechanism on I_1, $\mathbb{E}[rm] \leq p/2$. Now, let us turn our attention to instance I_2 in Figure 4b and let us denote by q the probability that agents $n-1$ and n are in a coalition together. The utility of agent $n-1$ depends on the number of agents in his coalition. If we denote by q_0 the probability that agents $n-1$ and n are alone in the coalition, q_1 the probability that there is one additional agent with them in the coalition, up to q_{n-2} the probability that all of the other $n-2$ agents are in a coalition together with agents $n-1$ and n, then $q = q_0 + \cdots + q_{n-2}$. Furthermore, if we denote by rm' the outcome of the mechanism on I_2, the expected utility of the agent $n-1$ is $\mathbb{E}[u_{n-1}(rm')] = \frac{q_0}{2} + \frac{q_1}{3} + \cdots + \frac{q_{n-2}}{n}$. In order for the mechanism to be strategyproof, it has to hold that $\mathbb{E}[u_{n-1}(rm')] \geq \mathbb{E}[u_{n-1}(rm)] = p/2$. The expected social welfare of the randomized mechanism on I_1 is therefore maximized for $\frac{p}{2} = \frac{q_0}{2} + \frac{q_1}{3} + \cdots + \frac{q_{n-2}}{n} \leq \frac{q_0 + q_1 + \cdots + q_{n-2}}{2} = \frac{q}{2}$. On the other hand, we have that

$$\mathbb{E}[rm'] \leq \frac{q_0}{2} + \frac{q_1}{3} + \cdots + \frac{q_{n-2}}{n} + (1-q)\frac{n-2}{n-1} \leq \frac{q}{2} + (1-q)\frac{n-2}{n-1}.$$

As in the proof of Theorem 6, by equalizing the expected approximation ratio of the mechanism on both instances, since $\mathbb{E}[rm] \leq p/2 \leq q/2$, we arrive at

$$\frac{\frac{1}{2}}{\mathbb{E}[rm]} = \frac{\frac{n-2}{n-1}}{\mathbb{E}[rm']} \implies \frac{1}{q} = \frac{\frac{n-2}{n-1}}{\frac{q}{2} + (1-q)\frac{n-2}{n-1}} \implies q = \frac{2n-4}{3n-7}.$$

It follows that as n increases, q tends to $2/3$ and thus the approximation ratio tends to $3/2$. $\qquad \square$

On the positive side, we will now present a deterministic strategyproof acceptable mechanism $\mathcal{M}_3$ with approximation $O(n^2)$ for ASHGs and $O(n)$ for FHGs. We doubt the existence of deterministic strategyproof acceptable mechanisms with approximation ratio $O(n)$ for ASHGs and $O(1)$ for FHGs. Some discussion supporting this view is provided after the analysis of $\mathcal{M}_3$. Closing the gap for deterministic mechanisms and duplex valuations remains one of the main open problems.

The following definition is crucial for mechanism $\mathcal{M}_3$.

Definition 1. *Given $\mathbf{d} = \langle d_1, ..., d_n \rangle$ declared by the set of agents N, we say that an agent $i \in N$ is a sink if there is no agent $j \in N$ such that $d_i(j) = 1$ and $d_j(i) \neq -1$.*

The idea of the mechanism $\mathcal{M}_3$ is as follows. It considers the agents in an arbitrary ordering. If the considered agent i has value 1 for some other agent j, such that j also has value 1 for i, or j is a sink, or j is before i in the ordering, then it returns agents i and j together in a coalition. If, after considering all the agents, the mechanism does not create a coalition for a specific agent, then it returns such agents in singleton coalitions.

Some example instances that invoke the just described cases, together with the outcomes produced by $\mathcal{M}_3$, are shown in Figure 5.

Mechanism $\mathcal{M}_3$. *Given any declared valuation $\mathbf{d} = \langle d_1, ..., d_n \rangle$, the mechanism performs as follows:*
1 Consider any ordering of the agents and, for the sake of simplicity, let i be the i-th agent in that ordering.
2 Compute S, the set of sinks.
3 For $i = 1$ to n and i not placed in a coalition yet:

a If there exists unpaired $j \in N$ such that $d_i(j) = 1 \wedge d_j(i) = 1$: put agents i and j together into a coalition.

b Else, if there exists unpaired $j \in N$ such that $d_i(j) = 1 \wedge d_j(i) = 0 \wedge j$ is a sink: put agents i and j together into a coalition.

c Else, if there exists unpaired $j \in N$ such that $d_i(j) = 1 \wedge d_j(i) = 0 \wedge j < i$: put agents i and j together into a coalition.

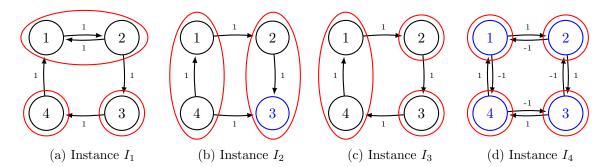*4 Put every agent that was not placed into a coalition in step 3 in a singleton coalition.*

(a) Instance I_1 (b) Instance I_2 (c) Instance I_3 (d) Instance I_4

Figure 5: Outcomes of Mechanism $\mathcal{M}_3$ on different instances, where on I_1, I_2, I_3 and I_4 lines 3a and 4, 3b and 3c, 3c and 4, and only 4 of $\mathcal{M}_3$ were applied, respectively. The sink nodes are depicted in blue and the coalitions returned by $\mathcal{M}_3$ in red.

First, let us prove that Mechanism $\mathcal{M}_3$ returns an outcome with positive social welfare whenever the optimal solution does so as well.

Lemma 3. *Given the valuations $\mathbf{d} = \langle d_1, ..., d_n \rangle$ declared by the agents, if there exists an agent i that is not a sink, then Mechanism $\mathcal{M}_3$ returns an outcome where two agents are put together in the same coalition, thus yielding positive social welfare.*

Proof. Only two scenarios are possible after running $\mathcal{M}_3$: i) agent i is put together with another agent, achieving a positive social welfare, or ii) i is put alone. The second case implies that, for any agent j such that $d_i(j) = 1$ and $d_j(i) \neq -1$, agent j is not a sink and she appears after i in the ordering. Thus we can now consider agent j as the new focus of our attention that is not a sink and apply the same argument as above. Once we in this manner arrive at the last agent k in the ordering that is not a sink, we know that she will

be successfully matched. This is indeed so, because there exists ℓ such that $d_k(\ell) = 1$ and $d_\ell(k) \neq -1$. If ℓ is before k, they will be matched. But also if ℓ is after k, meaning that ℓ must be a sink, they will again be matched. □

Theorem 8. *For ASHGs and FHGs with duplex valuations, Mechanism $\mathcal{M}_3$ is a deterministic strategyproof acceptable mechanism. The approximation ratio is $O(n^2)$ for ASHGs with duplex valuations, and $O(n)$ for FHGs with duplex valuations.*

Proof. No agent i is put in a coalition together with another agent j if there is a value of -1 between them, that is, if $d_i(j) = -1$ or $d_j(i) = -1$. This implies that no agent has negative utility in the outcome returned by $\mathcal{M}_3$, i.e., $\mathcal{M}_3$ is acceptable.

Strategyproofness is shown by case distinction. If we assume that agent i gets positive utility when declaring her valuation truthfully, then i cannot obtain better utility by declaring $d_i \neq v_i$, as utility of $\frac{1}{2}$ is the best she can obtain.

Assume now that agent i gets utility zero when declaring her valuation truthfully. If i is a sink, there is no way of improving her utility. Therefore, let us consider the case where i is not a sink. In this case, there exists an agent j such that $d_i(j) = 1$ and $d_j(i) \geq 0$. Let us first assume that i is in a singleton coalition in the outcome returned by Mechanism $\mathcal{M}_3$. From here, it follows that j was either paired already when the mechanism considered i in step 3 or $i < j$. In both cases, i cannot get paired with j by modifying her valuation as the problem is the ordering in which the mechanism is processing the agents in step 3 and she has no control over it. Let us now assume that i is in a coalition with agent k such that $d_i(k) = 0$ in the outcome returned by Mechanism $\mathcal{M}_3$. Since $d_i(k) = 0$ and $i \notin S$, the coalition had to be formed in step 3c. But this means that k was inspected after i in step 3, so it is not the case that i is not matched to j because of k. In turn, by misreporting $d_i(k) = -1$, i can at most achieve being in a singleton coalition. However, also in this case her utility stays zero.

We now turn to the approximation ratio of the mechanism. Notice that, given the valuations declared by the agents, if all the agents are sinks, then the optimal solution has social welfare zero and also $\mathcal{M}_3$ returns the outcome where each agent is in a singleton coalition. If the optimal solution has positive social welfare (and thus there exists an agent that is not a sink), then by Lemma 3, we know that our mechanism returns an outcome with social welfare at least 1 for ASHGs, and at least $\frac{1}{2}$ for FHGs. The theorem follows by noticing that any agent can get utility at most $n-1$ for ASHGs and at most 1 for FHGs. □

We point out that, if we consider ASHGs, there exists an instance and an ordering of the agents for that instance, such that the optimal solution has value $\Theta(n^2)$, while $\mathcal{M}_3$ only puts two agents in a coalition in the last iteration of the *For* loop (see Figure 6). Clearly, $\mathcal{M}_3$ could perform more *For* loops and allow iterative creation of new sinks in order to match more than one pair of agents. However, in such a case we can show that the mechanism is not strategyproof anymore. In fact, consider instance I_3 in Figure 5c. If the mechanism iterates the *For* loop, it would return in the first iteration agents $\{4, 1\}$ in a coalition, and then, in a second iteration of the *For* loop, agents $\{2, 3\}$ together. Notice that agent 1 has utility zero. However, agent 1 can improve her utility by declaring a further arc of weight -1 to agent 4. In fact, in this case, in the first iteration the mechanism would put agents $\{3, 4\}$ together, and then, in the second one, agents $\{1, 2\}$.
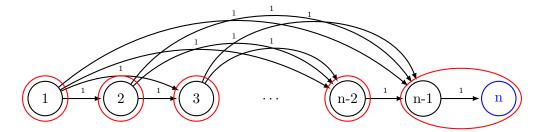
Figure 6: An instance in which each node $i \in \{1, \ldots, n-2\}$ has a positive valuation of 1 to all nodes in $\{i+1, \ldots, n-1\}$, and $n-1$ has a positive valuation of 1 to n. The sink nodes are depicted in blue and the coalitions returned by $\mathcal{M}_3$ in red.

We now present a simple randomized strategyproof acceptable mechanism $\mathcal{M}_4$ with approximation ratio $O(n)$ both for ASHGs and for FHGs. The idea of $\mathcal{M}_4$ is to pick a perfect or near-perfect matching M on the set of agents N uniformly at random, and then try to mimic M. More precisely, for every pair $\{i, j\} \in M$ a coalition $\{i, j\}$ is formed if and only if both $d_i(j) \neq -1$ and $d_j(i) \neq -1$. Figure 7 contains an example listing all the possible outcomes of Mechanism $\mathcal{M}_4$, when applied to the same instance.

Mechanism $\mathcal{M}_4$. *Given any declared valuation* $\mathbf{d} = \langle d_1, ..., d_n \rangle$, *the mechanism performs as follows:*
1 Consider any ordering of the agents and, for the sake of simplicity, let i be the i-th agent in that ordering.
2 Let M be a random perfect or near-perfect matching on the complete graph given by the set of the agents.
3 For $i = 1$ to n: If there exists $j \in N$ such that $\{i, j\} \in M \wedge d_i(j) \neq -1 \wedge d_j(i) \neq -1$: put agents i and j together into a coalition.
4 Put every agent that was not placed into a coalition in step 3 in a singleton coalition.

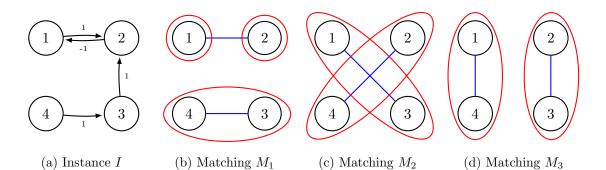(a) Instance I (b) Matching M_1 (c) Matching M_2 (d) Matching M_3

Figure 7: Different outcomes of Mechanism $\mathcal{M}_4$ on the same instance, depending on the randomly chosen perfect matching. The matched vertices are connected via a blue edge, while the coalitions returned by $\mathcal{M}_4$ are shown in red.

Theorem 9. *For ASHGs and FHGs with duplex valuations, Mechanism $\mathcal{M}_4$ is a randomized universally strategyproof acceptable mechanism. The approximation ratio is $O(n)$ for ASHGs and FHGs with duplex valuations.*

Proof. Mechanism $\mathcal{M}_4$ is obviously acceptable because no agent i is placed in a coalition with another agent j if $d_i(j) = -1$ or $d_j(i) = -1$. Strategyproofness follows from restricting the non-singleton coalitions of the outcome to subsets of the matching M. Indeed, since no agent can influence M, there is also no way for an agent to improve her utility by manipulating her true valuations even if she would know M in advance. Mechanism $\mathcal{M}_4$ is therefore universally strategyproof.

For the approximation ratio, let us focus on two agents i and j. If n is even, i is equally likely to be matched to any of the other $n - 1$ agents, so in particular also to j, and that is with probability $1/(n-1)$. If n is odd, i is unmatched with probability $1/n$ and matched with the remaining probability $(n-1)/n$. In particular, i is matched with any of the remaining $n - 1$ agents with probability $1/n$.

This in total means that if we consider any edge that contributes to the optimum, Mechanism $\mathcal{M}_4$ will also obtain its contribution with probability of at least $\frac{1}{n}$. Now, for any preference profile $\mathbf{d}$ in ASHGs, using linearity of expectation,

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}_4(\mathbf{d}))] = \sum_{e \in E} \Pr(e \in \mathcal{M}_4(\mathbf{d})) \cdot w(e) \geq \sum_{\substack{e \in E \\ w(e) \geq 1}} \frac{1}{n} \cdot w(e) = \frac{1}{n} \sum_{\substack{e \in E \\ w(e) \geq 1}} w(e) \geq \frac{1}{n} \mathrm{OPT}(\mathbf{d}) \ .$$

For FHGs, analogously we see that

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}_4(\mathbf{d}))] = \sum_{e \in E} \Pr(e \in \mathcal{M}_4(\mathbf{d})) \cdot \frac{w(e)}{2} \geq \sum_{\substack{e \in E \\ w(e) \geq 1}} \frac{1}{n} \cdot \frac{w(e)}{2} = \frac{1}{2n} \sum_{\substack{e \in E \\ w(e) \geq 1}} w(e) \geq \frac{1}{2n} \mathrm{OPT}(\mathbf{d}) \ .$$

$\square$

The fact that Mechanism $\mathcal{M}_4$ gives an $O(n)$-approximation for FHGs with duplex valuations is less interesting as Mechanism $\mathcal{M}_3$ already had the same approximation ratio for this setting and $\mathcal{M}_3$ is in addition deterministic.

Lastly, we present a randomized strategyproof acceptable mechanism $\mathcal{M}_5$ with approximation ratio $O(1)$ for FHGs with duplex valuations. The idea is to develop a randomized version of mechanism $\mathcal{M}_3$, for which we could not guarantee a better approximation ratio for every possible ordering of the agents. In mechanism $\mathcal{M}_5$ we will choose the ordering in which the agents are processed uniformly at random from the set of all possible permutations of n elements.

Mechanism $\mathcal{M}_5$. *Given any declared valuation $\mathbf{d} = \langle d_1, ..., d_n \rangle$, the mechanism performs as follows:*
1 Let Π be a permutation of n elements, chosen uniformly at random.
2 Run $\mathcal{M}_3$ for the ordering given by Π.

The additional permutation Π is the only source of randomness in mechanism $\mathcal{M}_5$ and without it, we would not be able to guarantee a constant approximation ratio for FHGs

with duplex valuations. More precisely, there exist instances for which we cannot guarantee a constant approximation ratio for every possible ordering of the agents. One such example is already given in Figure 6, in which the output of Mechanism $\mathcal{M}_5$ would coincide with the output of Mechanism $\mathcal{M}_3$ without the additional permutation Π. However, if we use an additional ordering based on Π, which is chosen uniformly at random, the outcome in the example instance given in Figure 6 changes (see Figure 8), and in fact the expected approximation ratio becomes at most 8, as we will see in Theorem 10.
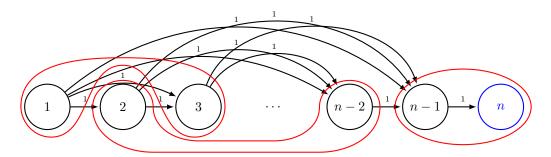

Figure 8: An instance in which each node $i \in \{1, \ldots, n-2\}$ has a positive valuation of 1 to all nodes in $\{i+1, \ldots, n-1\}$, and $n-1$ has a positive valuation of 1 to n. The sink nodes are depicted in blue and the coalitions returned by $\mathcal{M}_5$ in red, assuming that the order in which the nodes are processed, given by Π, is $3, n-2, 1, n-1, n, 2, \ldots$. For such Π, the first two pairs, $\{1, 3\}$ and $\{2, n-2\}$, are created via line $3c$ and the last pair, $\{n-1, n\}$, via line $3b$ of $\mathcal{M}_3$.

Theorem 10. *For FHGs with duplex valuations, Mechanism $\mathcal{M}_5$ is a randomized universally strategyproof acceptable mechanism. The approximation ratio is 8 for FHGs with duplex valuations.*

Proof. Mechanism $\mathcal{M}_5$ is acceptable exactly for the same reasons as mechanism $\mathcal{M}_3$.

For a fixed permutation Π, $\mathcal{M}_5$ coincides with $\mathcal{M}_3$, which is truthful by Theorem 8. Since Mechanism $\mathcal{M}_5$ is strategyproof even if the random permutation Π is fixed, and therefore known to the agents, it is universally strategyproof.

We will now prove the claimed approximation ratio of Mechanism $\mathcal{M}_5$ when applied to FHGs with duplex valuations in two steps. First, we will relate the expected social welfare of the coalition structure returned by Mechanism $\mathcal{M}_5$ to the social welfare of the optimal matching on the set of all agents in Lemma 4. Then, we will show that the social welfare of the optimal matching is related to the actual optimal social welfare $\text{OPT}(\mathbf{d})$ by using Theorem 1. By combining these two statements, we will arrive at the claimed result.

Lemma 4. *Given the valuations $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$ declared by the agents, the expected value of the outcome returned by Mechanism $\mathcal{M}_5$ is at least $1/4$ of the social welfare of the optimal matching $\text{OPT}_M(\mathbf{d})$, i.e.,*

$$\mathbb{E}[SW(\mathcal{M}_5)] \geq \frac{1}{4}\text{OPT}_M(\mathbf{d}) \ .$$

Proof. Let us denote by ℓ the number of pairs appearing in the optimal matching M. Then, $M = \{\{i_1, i'_1\}, \dots, \{i_\ell, i'_\ell\}\}$. Now, consider the contribution of the set $\{i_p, i'_p\}$ for some $p \in \{1, \dots, \ell\}$ to the social welfare of M, and denote it by c_p. Notice that $c_p \in \{\frac{1}{2}, 1\}, \forall p \in [\ell]$.

If $c_p = 1$, then $d_{i_k}(i'_k) = d_{i'_k}(i_k) = 1$, which means that Mechanism $\mathcal{M}_5$ will be able to put these agents in a coalition together independently of the pairwise order in which permutation Π places i_p and i'_p. If $c_p = 1/2$ and one of i_p, i'_p is a sink, then again as in the previous case Mechanism $\mathcal{M}_5$ will be able to create this pairing independently of the mutual order of i_p and i'_p in the permutation Π. Lastly, if the contribution of $\{i_p, i'_p\}$ to $\texttt{OPT}_M(\mathbf{d})$ is $1/2$ and both i_p and i'_p are not sinks, then let us assume w.l.o.g. that $d_{i_k}(i'_k) = 1$ and $d_{i'_k}(i_k) = 0$. In this case, in order for Mechanism $\mathcal{M}_5$ to be able to create this pair, permutation Π has to be such that $\Pi^{-1}(i'_k) < \Pi^{-1}(i_k)$.

Let us denote by X_p the random variable whose value is equal to the contribution of $\{i_p, i'_p\}$ to $\mathrm{SW}(\mathcal{M}_5(\mathbf{d}))$, where the individual contribution of each agent from $\{i_p, i'_p\}$ equals half of the value of the pair it is placed in. Note that i_p and i'_p may not be paired together, so $X_p \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$. Then,

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}_5(\mathbf{d}))] \geq \mathbb{E}\left[\sum_{p=1}^{\ell} X_p\right] = \sum_{p=1}^{\ell} \mathbb{E}[X_p],$$

where the inequality follows from $\mathcal{M}_5$ possibly having positive contributions from agents that are not in M, and the equality by linearity of expectation.

Now, let us take a look at $\mathbb{E}[X_p]$ for $p \in \{1, \dots, \ell\}$. If $c_p = 1$, then also $\mathbb{E}[X_p] \geq \frac{1}{2}$. This is so because $\mathcal{M}_5$ is able to pair i_p and i'_p independently of their mutual ordering in Π, and the case that they are not paired can only occur if at least one of them, w.l.o.g. i_p, is paired with another agent i_q such that $d_{i_p}(i_q) = d_{i_q}(i_p) = 1$. Consequently, the contribution of i_p is then $1/2$. If $c_p = 1/2$ and one of the two agents is a sink, we claim that $\mathbb{E}[X_p] \geq \frac{1}{4}$. Again, $\mathcal{M}_5$ is able to pair i_p and i'_p independently of their mutual ordering in Π. If the pair is established, the contribution of $\{i_p, i'_p\}$ to $\mathcal{M}_5((d))$ is $\frac{1}{2}$ and if it is not established, this means that $\mathcal{M}_5$ paired at least one of them, w.l.o.g. i_p, with some i_q and the contribution of $\{i_p, i'_p\}$ is therefore at least $\frac{1}{4}$. In conclusion, $\mathbb{E}[X_p] \geq \frac{1}{4}$. Note that for the two cases considered so far, we derived the expected value of X_p deterministically. Let us now consider the last case, where the contribution of $\{i_p, i'_p\}$ to $\mathrm{OPT}(\mathbf{d})$ is $\frac{1}{2}$ and none of the two agents is a sink. Let us w.l.o.g. assume that $d_{i_p}(i'_p) = 1$ and $d_{i'_p}(i_p) = 0$. In this case, $\mathcal{M}_5$ is able to pair them if and only if $\Pi^{-1}(i'_p) < \Pi^{-1}(i_p)$ and since Π is chosen uniformly at random, this occurs with probability $\frac{1}{2}$. Furthermore, we claim that $\mathbb{E}\left[X_p \mid \Pi^{-1}(i'_p) < \Pi^{-1}(i_p)\right] \geq \frac{1}{4}$. Similarly to the previously discussed cases, this is so because the case that $\Pi^{-1}(i'_p) < \Pi^{-1}(i_p)$ and the pair $\{i_p, i'_p\}$ is not created can only occur if, w.l.o.g., i_p is paired with some i_q and again this immediately implies that the contribution of $\{i_p, i'_p\}$ is at least $\frac{1}{4}$. In conclusion, $\mathbb{E}[X_p] \geq \mathbb{E}\left[X_p \mid \Pi^{-1}(i'_p) < \Pi^{-1}(i_p)\right] \geq \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$, where the first inequality follows by the law of total expectation.

We showed that for all $p \in \{1, \dots, \ell\}, \mathbb{E}[X_p] \geq \frac{1}{4}c_p$. Therefore,

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}_5(\mathbf{d}))] \geq \sum_{p=1}^{\ell} \mathbb{E}[X_p] \geq \sum_{p=1}^{\ell} \frac{1}{4}c_p = \frac{1}{4}\sum_{i=1}^{\ell} c_p = \frac{1}{4}\texttt{OPT}_M(\mathbf{d}).$$

$\square$

To see how the social welfare of the optimal matching $\mathtt{OPT}_M(\mathbf{d})$ relates to the optimal social welfare $\mathrm{OPT}(\mathbf{d})$, notice that every instance of FHG with duplex valuations can be seen as an undirected weighted graph with edge weights in $\{-2, -1, 0, 1, 2\}$. Formally, given a directed weighted graph $G = (V, E, w)$ corresponding to a FHG instance with duplex valuations, we create a complete undirected weighted graph $\tilde{G} = (V, \tilde{E}, \tilde{w})$ by setting $\tilde{w}(\{i, j\}) = w(i, j) + w(j, i)$. A matching m in $\tilde{G}$ naturally induces an outcome for fractional hedonic games, that is, any edge $\{i, j\} \in m$ induces the coalition $C_{i,j} = \{i, j\}$, and for any node i not matched in m we have the coalition $C_i = \{i\}$. Now, we can apply Theorem 1, to conclude that

$$\mathtt{OPT}_M(\mathbf{d}) \geq \frac{1}{2}\mathtt{OPT}(\mathbf{d}). \tag{1}$$

By combining Lemma 4 and Equation 1, we see that

$$\mathbb{E}[\mathrm{SW}(\mathcal{M}_5(\mathbf{d}))] \geq \frac{1}{4}\mathtt{OPT}_M(\mathbf{d}) \geq \frac{1}{4} \cdot \frac{1}{2}\mathtt{OPT}(\mathbf{d}) = \frac{1}{8}\mathtt{OPT}(\mathbf{d}) \ ,$$

and finally arrive at the claim of the theorem. $\square$

Note that since Mechanism $\mathcal{M}_5$ creates pairs, for ASHGs with duplex valuations it can achieve a social welfare of at most $O(n)$, while there are instances where the optimal coalition structure is the grand coalition with a social welfare of $O(n^2)$ (see Figure 8). In conclusion, a mechanism that creates pairs cannot have a better approximation ratio than $O(n)$ for ASHGs with duplex valuations and, therefore, $\mathcal{M}_5$ cannot asymptotically outperform $\mathcal{M}_4$ in this setting.

7. Simple Valuations

Exactly as in the case of non-negative valuations, for ASHGs with simple valuations, Mechanism $\mathcal{M}_2$ is acceptable and strategyproof and it also achieves the optimal social welfare. Therefore, we focus on FHGs. We first prove that no deterministic strategyproof mechanism can approximate the optimal social welfare with an approximation ratio that is better than $\frac{6}{5}$.

Theorem 11. *For FHGs with simple valuations, no deterministic strategyproof acceptable mechanism has approximation ratio less than $\frac{6}{5}$.*

Proof. Let us consider the instance I_1 depicted in Figure 9a. The reader can easily check (by considering all the possible coalitions) that an optimal solution has social welfare $\frac{5}{3}$. It is composed by three coalitions where two of them contain two consecutive agents and the remaining one contains three consecutive agents. For instance, an optimal solution could be $C_1 = \{1, 2\}, C_2 = \{3, 4\}, C_3 = \{5, 6, 7\}$. Notice that the grand coalition has social welfare 1. Therefore, a mechanism achieving an approximation better than $\frac{5}{3}$ has to return more than one coalition. In such a solution there always exists at least one agent, say agent k, having utility zero. Let us now consider the instance I_2 depicted in Figure 9b, where without loss of generality we suppose that $k = 2$. Again the reader can easily check (by considering all the possible coalitions) that the optimal solution has social welfare 2 and

consists of coalitions $C_1 = \{2, 3, 4\}, C_2 = \{5, 6\}, C_3 = \{1, 7\}$. Once again the reader can check that any solution where agents 2 and 3 are not in the same coalition (i.e., any solution where agent 2 has utility equal to 0 in instance I_1) can achieve a social welfare of at most $\frac{5}{3}$, and therefore an approximation not better than $\frac{6}{5}$. We conclude that any mechanism achieving an approximation ratio strictly better than $\frac{6}{5}$, in both instances I_1 and I_2, is not strategyproof.

(a) Instance I_1　　　　　　　　(b) Instance I_2

Figure 9: The lower bound instance for simple valuations.

We now show that Mechanism $\mathcal{M}_1$ is a deterministic strategyproof acceptable mechanism with nearly optimal social welfare. Given the preferences declared by the agents $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$, and the associated directed weighted graph representation $G = (V, E, w)$ (notice that since we are considering simple valuations, d_i represents (indeed is) the set E_i of arcs outgoing from node i in G), we construct a complete undirected weighted graph $\tilde{G} = (V, \tilde{E}, \tilde{w})$ by setting $\tilde{w}(\{i, j\}) = w(i, j) + w(j, i)$. This means that for each $\{i, j\} \in \tilde{E}$, we have that the weight $\tilde{w}(\{i, j\}) = 1$ if either $d_i(j) = 1$ (i.e., $(i, j) \in E$) or $d_j(i) = 1$ (i.e., $(j, i) \in E$) but not both, and $\tilde{w}(\{i, j\}) = 2$ if both $d_i(j) = 1$ and $d_j(i) = 1$, otherwise $\tilde{w}(\{i, j\}) = 0$. A matching m in $\tilde{G}$ naturally induces an outcome for fractional hedonic games, that is, any edge $\{i, j\} \in m$ induces the coalition $C_{i,j} = \{i, j\}$, and for any node i not matched in m we have the coalition $C_i = \{i\}$. Notice that the coalitions induced by the matching are such that each agent can have utility either $\frac{1}{2}$ or 0. We show that Mechanism $\mathcal{M}_1$ is strategyproof in this setting.

Theorem 12. *For FHGs with simple valuations, Mechanism $\mathcal{M}_1$ is a deterministic strategyproof acceptable mechanism with approximation ratio of 2.*

Proof. The approximation ratio follows from Theorem 1. Now, assume for a contradiction that $\mathcal{M}_1$ is not truthful. Then there exists agent i that benefits by declaring preferences different from the real ones, i.e., there exists $\tilde{G} = (V, \tilde{E}, \tilde{w})$ with $\tilde{E}$ induced by edges $E_{-i} \cup E_i$, and E_i' (inducing the graph $\tilde{G}' = (V, \tilde{E}', \tilde{w}')$ where the set of edges $\tilde{E}' = \tilde{E}_{-i} \cup \tilde{E}'_i$), violating truthfulness. Let $m = \mathcal{M}_1(\tilde{G})$ and $m' = \mathcal{M}_1(\tilde{G}')$. Agent i has utility zero in the coalitions induced by m, that is, if $\{i, j\} \in m$ then $(i, j) \notin E_i$. Yet agent i has utility $\frac{1}{2}$ in the coalitions induced by m'. It means that there exists $k \in V$ such that $(i, k) \in E_i$ and $\{i, k\} \in m'$. Moreover, since the input of the mechanism is a complete undirected weighted graph whose

weights are obtained from the preferences declared by the agents and agent i has utility $\frac{1}{2}$ in the coalitions induced by m', it follows that $(i, k) \in E_i \cap E_i'$. This implies that both m and m' are in $\mathcal{M}(\tilde{E}) \cap \mathcal{M}(\tilde{E}')$, where $\mathcal{M}(\tilde{E})$ and $\mathcal{M}(\tilde{E}')$ denote the set of all possible matchings on $\tilde{G} = (V, \tilde{E}, \tilde{w})$ and $\tilde{G}' = (V, \tilde{E}', \tilde{w}')$, respectively. Since the mechanism returns the maximum weight matching it follows that m and m' are optimal in both $\mathcal{M}(\tilde{E})$ and $\mathcal{M}(\tilde{E}')$. Recalling that $\mathcal{M}_1$ breaks ties consistently, this yields a contradiction, as needed. $\square$

We point out that when dealing with FHGs, it is natural to resort to matchings. Many papers (e.g. Aziz et al., 2015, 2019; Bilò et al., 2018) used them. The challenge is how to exploit their properties, and in this sense we make some steps forward. Indeed, we better exploit properties of maximum weight matchings. This is manifested by the fact that our analysis can be used to improve the 4-approximation (Theorem 7 of the paper by Aziz et al. (2015)) of maximum weight matching for symmetric valuations, i.e., undirected graph, to a 2-approximation (see Section 3). Another remark is that our results not only work for the approximation of asymmetric FHGs, i.e., directed graphs, but also include strategyproofness, which was not considered before for FHGs.

8. Conclusion and Future Work

Hedonic games have received considerable attention in the scientific community, but most of the literature assumes that agents' preferences are given. However, agents receive different utilities if they belong to different coalitions, and thus it is natural for them to declare their preferences strategically in order to maximize their benefit. In this paper we studied strategyproof mechanisms (without payments), that is, coalition formation algorithms in which agents have no incentive to lie about their true preferences, for two natural and succinctly representable classes of hedonic games with cardinal utilities, ASHGs and FHGs. We adopted the classical utilitarian social welfare, i.e., the sum of agents' utilities, and provided a number of lower and upper bounds on the performance achievable by deterministic and randomized strategyproof mechanisms.

We showed some differences between ASHGs and FHGs and shed some light onto the capabilities and limitations of strategyproof mechanisms for both game classes. In particular, we first provided quite negative results for general additive valuations for both ASHGs and FHGs, and then focused on more specific additive valuations, for which we were able to show a better performance. Despite the primarily theoretical interest of these specific valuations, they also model realistic scenarios, such as Bakers and Millers (Aziz et al., 2019; Bilò et al., 2018). We were mainly interested in designing deterministic strategyproof mechanisms with performance guarantees (i.e., in the worst-case analysis sense), and in some cases we provided results on the limitations of randomized strategyproof mechanisms. Moreover, for the case of duplex valuations, for which we were not able to find deterministic strategyproof mechanisms with an asymptotically tight performance with respect to our lower bounds, we also designed randomized strategyproof mechanisms with better performance, both for ASHGs and FHGs.

In summary, our results say that for both games, when considering general valuations, no randomized (and thus also no deterministic) strategyproof mechanism can achieve a bounded approximation ratio for all inputs. Moreover, when considering positive valuations,

using the trivial strategyproof mechanism which puts all the agents together into the same coalition is the best that can be done, both for ASHGs and FHGs. For the case of duplex valuations we were not able to close the asymptotic gaps between our upper and lower bounds. However, we provided randomized strategyproof mechanisms for both games with good, and in case of FHGs even asymptotically tight, performance. These mechanisms are in addition simple and efficient in terms of complexity, so we believe that they could be used in practice. Finally, for simple valuations, while for ASHGs the problem is trivial, for FHGs we showed that returning a particular maximum matching is both strategyproof and almost optimal. This last mechanism is also simple and efficient and thus we believe it is very suitable for practical use.

To the best of our knowledge, this is the first work that analyzes the performance of strategyproof mechanisms for ASHGs and FHGs. There are several open problems that still need to be addressed. First of all, some of the provided upper and lower bounds are not tight, as outlined in Table 1. In particular, it is intriguing to close the gap between upper and lower bounds of deterministic and randomized strategyproof mechanisms in the setting of duplex valuations, both for ASHGs and FHGs. Moreover, it would be valuable to understand whether randomized strategyproof mechanisms can achieve a significantly better performance than deterministic ones for FHGs with non-negative valuations. What is also important to understand is what happens when valuations are drawn from a distribution (in order to avoid the bad instances), or when there are size constraints on the coalitions. It would be also interesting to consider more general valuation functions, e.g. submodular. Another open research direction is considering strategyproof mechanisms whose outputs satisfy additional properties like Pareto optimality (Elkind et al., 2020) and envy-freeness (Barrot & Yokoo, 2019). Finally, we were mainly interested in analyzing the performance guaranteed by strategyproof mechanisms, focusing on and guided by the worst-case view. It would be also important to design and analyze strategyproof mechanisms that work well in, and are especially suited for, particular practical applications.

As a more general research direction, we see the study of strategyproof mechanisms for all the classes of hedonic games appearing in the literature.

Acknowledgments

References

Anderson, I. (2001). *A first course in discrete mathematics*. Springer undergraduate mathematics series. Springer.

Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., & Peters, D. (2019). Fractional Hedonic Games. *ACM Trans. Econom. Comput.*, *7*(2), 6:1–6:29.

Aziz, H., Brandt, F., & Harrenstein, P. (2013a). Pareto optimality in coalition formation. *Games Econom. Behav.*, *82*, 562–581.

Aziz, H., Brandt, F., & Seedig, H. G. (2013b). Computing desirable partitions in additively separable hedonic games. *Artif. Intell.*, *195*, 316–334.

Aziz, H., Gaspers, S., Gudmundsson, J., Mestre, J., & Täubig, H. (2015). Welfare Maximization in Fractional Hedonic Games. In *Proc. 24th Intl. Joint Conf. Artif. Intell. (IJCAI)*, pp. 461–467.

Aziz, H., & Savani, R. (2016). Hedonic Games. In *Handbook of Computational Social Choice*, pp. 356–376.

Banerjee, S., Konishi, H., & Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice & Welfare*, *18*(1), 135–153.

Barrot, N., & Yokoo, M. (2019). Stable and Envy-free Partitions in Hedonic Games. In Kraus, S. (Ed.), *Proc. 28th Intl. Joint Conf. Artif. Intell. (IJCAI)*, pp. 67–73.

Bilò, V., Fanelli, A., Flammini, M., Monaco, G., & Moscardelli, L. (2015). On the Price of Stability of Fractional Hedonic Games. In *Proc. 14th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 1239–1247.

Bilò, V., Fanelli, A., Flammini, M., Monaco, G., & Moscardelli, L. (2018). Nash Stable Outcomes in Fractional Hedonic Games: Existence, Efficiency and Computation. *J. Artif. Intell. Res.*, *62*, 315–371.

Bloch, F., & Diamantoudi, E. (2011). Noncooperative formation of coalitions in hedonic games. *Int. J. Game Theory*, *40*(2), 263–280.

Bogomolnaia, A., & Jackson, M. O. (2002). The Stability of Hedonic Coalition Structures. *Games Econom. Behav.*, *38*(2), 201–230.

Brandl, F., Brandt, F., & Strobel, M. (2015). Fractional hedonic games: Individual and group stability. In *Proc. 14th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 1219–1227.

Bullinger, M. (2020). Pareto-optimality in cardinal hedonic games. In *Proc. 19th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 213–221.

Carosi, R., Monaco, G., & Moscardelli, L. (2019). Local Core Stability in Simple Symmetric Fractional Hedonic Games. In *Proc. 18th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 574–582.

Deng, X., & Papadimitriou, C. H. (1994). On the Complexity of Cooperative Solution Concepts. *Math. Oper. Res.*, *19*(2), 257–266.

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

Dughmi, S., & Ghosh, A. (2010). Truthful assignment without money. In *Proc. 11th Conf. Econom. Comput. (EC)*, pp. 325–334.

Elkind, E., Fanelli, A., & Flammini, M. (2020). Price of pareto optimality in hedonic games. *Artif. Intell.*, *288*, 103357.

Elkind, E., & Wooldridge, M. J. (2009). Hedonic coalition nets. In *Proc. 8th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 417–424.

Feldman, M., Lewin-Eytan, L., & Naor, J. (2015). Hedonic Clustering Games. *ACM Trans. Parallel Computing*, *2*(1), 4:1–4:48.

Flammini, M., Monaco, G., Moscardelli, L., Shalom, M., & Zaks, S. (2018). Online Coalition Structure Generation in Graph Games. In *Proc. 17th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 1353–1361.

Flammini, M., Monaco, G., & Zhang, Q. (2017). Strategyproof Mechanisms for Additively Separable Hedonic Games and Fractional Hedonic Games. In *Proc. 15th Intl. Workshop Approx. and Online Algorithms (WAOA)*, pp. 301–316.

Gairing, M., & Savani, R. (2019). Computing Stable Outcomes in Symmetric Additively Separable Hedonic Games. *Math. Oper. Res.*, *44*(3), 1101–1121.

Kaklamanis, C., Kanellopoulos, P., Papaioannou, K., & Patouchas, D. (2021). On the price of stability of some simple graph-based hedonic games. *Theor. Comput. Sci.*, *855*, 1–15.

Monaco, G., Moscardelli, L., & Velaj, Y. (2020). Stable outcomes in modified fractional hedonic games. *Autonomous Agents and Multi-Agent Syst.*, *34*(1), 4.

Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. V. (2007). *Algorithmic game theory*. Cambridge University Press.

Olsen, M. (2012). On Defining and Computing Communities. In *Proc. 18th Conf. Computing: The Australasian Theory Symposium (CATS)*, pp. 97–102.

Peters, D. (2016). Graphical Hedonic Games of Bounded Treewidth. In *Proc. 30th Conf. Artificial Intelligence (AAAI)*, pp. 586–593.

Peters, D., & Elkind, E. (2015). Simple causes of complexity in hedonic games. In *Proc. 24th Intl. Joint Conf. Artif. Intell. (IJCAI)*, pp. 617–623.

Procaccia, A. D., & Tennenholtz, M. (2013). Approximate Mechanism Design without Money. *ACM Trans. Econom. Comput.*, *1*(4), 18:1–18:26.

Rahwan, T., Michalak, T. P., Wooldridge, M. J., & Jennings, N. R. (2015). Coalition structure generation: A survey. *Artif. Intell.*, *229*, 139–174.

Rodríguez-Álvarez, C. (2009). Strategy-proof coalition formation. *Int. J. Game Theory*, *38*(3), 431–452.

Vallée, T., Bonnet, G., Zanuttini, B., & Bourdon, F. (2014). A study of sybil manipulations in hedonic games. In *Proc. 13th Conf. Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 21–28.

Woeginger, G. J. (2013). Core Stability in Hedonic Coalition Formation. In *Proc. 39th Intl. Conf. Current Trends in Theory & Practice of Comput. Sci. (SOFSEM)*, pp. 33–50.

Wright, M., & Vorobeychik, Y. (2015). Mechanism Design for Team Formation. In *Proc. 29th Conf. Artificial Intelligence (AAAI)*, pp. 1050–1056.