

On the price of stability of some simple graph-based hedonic games ^{*}

Christos Kaklamanis^{a,c}, Panagiotis Kanellopoulos^{b,c}, Konstantinos Papaioannou^a, Dimitris Patouchas^a

^aUniversity of Patras, 26504 Rion, Greece

^bUniversity of Essex, Colchester CO4 3SQ, UK

^cComputer Technology Institute and Press “Diophantus”, 26504 Rion, Greece

Abstract

We consider graph-based hedonic games such as simple symmetric fractional hedonic games and social distance games, where a group of utility maximizing players have hedonic preferences over the players’ set, and wish to be partitioned into clusters so that they are grouped together with players they prefer. The players are nodes in a connected graph and their preferences are defined so that shorter graph distance implies higher preference. We are interested in Nash equilibria of such games, where no player has an incentive to unilaterally deviate to another cluster, and we focus on the notion of the price of stability. We present new and improved bounds on the price of stability for several graph classes, as well as for a slightly modified utility function.

Keywords: Fractional hedonic games, social distance games, equilibria, price of stability

1. Introduction

Economic entities, be it individuals or corporations, interact frequently in the context of performing complex tasks or even in enjoying cultural activities. For example, people usually tend to like or dislike other people and, therefore, wish to socialize or distance themselves depending on the occasion. The choice about which party to attend, which restaurant to dine in, etc., usually depends also on the other participants that will be present. This process, sometimes also called *coalition formation*, and the corresponding behavior can be captured by the class of hedonic games, where participating agents have preferences over coalitions (or groups), and, based on these preferences, they behave accordingly when selecting which group to join. In such scenarios, when an agent has decided to join a specific group, e.g., for having dinner, her utility depends only

^{*}Preliminary versions of this paper appeared in the Proceedings of the 9th Symposium on Algorithmic Game Theory (SAGT 2016) [21] and the 11th Symposium on Algorithmic Game Theory (SAGT 2018) [22].

on the other agents in the same group and not on how the remaining agents have been grouped together. Such preferences are termed *hedonic preferences* and completely ignore inter-coalitional dependencies. Due to their simplicity, hedonic games can be used to model a large spectrum of activities (e.g., clustering in social networks [2], distributed task allocation for wireless agents [29], etc.).

Hedonic games can be very expressive and admit a large class of utility functions over coalitions. For instance, given a group, we may care about the sum of utility we obtain over all members of the group, or only care about the minimum or maximum utility (again, over all members of the group). In addition, we may prefer a smaller group, containing people that we value significantly, over a larger group, with the same set of preferred people as well as several other members that we are indifferent to; in this case, we are interested in the average utility we obtain. What constitutes an acceptable or desired solution in such games is a question that has also attracted significant attention. Clearly, a natural objective is to compute a solution that maximizes some global function over all participating agents. This solution, however, may leave several agents unsatisfied and they may not adhere to it, but, on the contrary, may choose to deviate to another group if this is to their best interest. Then, such deviations may incentivize further agents (or even groups of agents) to deviate on their own (or, respectively, in collaboration), and so on, until a group formation satisfying all agents is reached, if such a formation exists.

In this paper, we assume that all players are nodes in a connected graph and that their preferences are defined based on the graph structure; in principle, a shorter distance leads to higher preference. We are interested in stable states of the coalition formation process in such graph-based games, i.e., when all cluster members are satisfied with the current configuration and no one wishes to join a different cluster, and, in addition, in comparing a stable state to the optimal partition. We consider a model where the utility obtained by a player depends on the cluster distance to the cluster members; informally, we assume that people tend to prefer being in clusters with friends, or friends of friends, than with strangers. Classes of strategic games that capture this process include *simple symmetric fractional hedonic games* as well as is that of *social distance games*.

Related work. Hedonic games that rely on hedonic preferences were introduced by Drèze and Greenberg [14]. Bogomolnaia and Jackson [11] present sufficient conditions for the existence of core stable partitions in hedonic settings and also consider the weaker notion of individual stability, where no player can deviate to another cluster without either hurting itself or hurting a member of its new cluster. An important subclass is that of additively separable hedonic games (see [3, 18, 25]), where the total utility of each player is defined as the sum of utility it obtains from each player in its cluster.

In *fractional hedonic games*, the utility of each player is defined as the sum of utility it obtains from each player in its cluster divided by the cluster size. Aziz et al. [2] introduce the model, consider more general stability notions, such as core stability, and present positive results for several classes of graphs.

Aziz et al. [4] consider the computational complexity of computing partitions that maximize the social welfare, defined as the sum of the players' utilities, in fractional hedonic games, without caring about stability. They consider three different notions of social welfare (i.e., utilitarian, egalitarian, and Nash welfare) and show that maximizing social welfare is NP-hard even for the subclass of *simple symmetric fractional hedonic games*, where the utility obtained from a single player can be either 0 or 1 and the utility is symmetric. On the positive side, they present polynomial time algorithms with small constant approximation ratio for the notions of utilitarian and egalitarian social welfare and the class of simple symmetric fractional hedonic games.

Olsen [26], among other results, suggests an alternate utility function for fractional hedonic games, where the utility function of player i does not take i into account when averaging over the cluster size, i.e., each player is interested in the average utility obtained from all other players in the same cluster. Fractional hedonic games under Olsen's utility function are termed *modified fractional hedonic games* in [16] where they consider Pareto optimality in such games. Monaco et al. [23, 24] consider stable outcomes in modified fractional hedonic games and study their existence and impact on the social welfare.

Bilò et al. [10] consider the price of anarchy and stability in fractional hedonic games. They show that when the utility function may take negative values, Nash stable outcomes are not guaranteed to exist, but if all values are nonnegative, then the partition where all players are in the same cluster is Nash stable. For the last case, they show an upper bound of $O(n)$ on the price of anarchy, which is tight even for simple symmetric fractional hedonic games. Furthermore, they show a lower bound of $\Omega(n)$ on the price of stability for games played on weighted stars and non-negative utility functions. For the price of stability in simple symmetric fractional hedonic games, they show a lower bound of 2 for general graphs, an upper bound of 4 for triangle-free graphs, and almost tight bounds for the case of bipartite graphs. In particular, they present an upper bound of $6(3 - 2\sqrt{2}) \approx 1.0294$ and a lower bound of 1.003. In addition, Bilò et al. [10] observe that their upper bounds still hold for the utility function defined by Olsen. Further notions of stability in fractional hedonic games have been investigated by Brandl et al. [12].

In *social distance games*, the utility of each player depends on the average inverse distance from the other cluster players. This class of games was introduced by Brânzei and Larson [13] who proved that finding the optimal partition is NP-hard and designed a 2-approximation algorithm with respect to the social welfare. In addition, they studied stability using the notion of the core. The work that is most related to ours is by Balliu et al. [5] who considered the price of stability in social distance games and presented lower bounds of $6/5$ for general graphs, $169/160$ for bipartite graphs, as well as an upper bound of $(\sqrt{2} + 1)/2$ for graphs of girth 5. Moreover, they proved that social distance games do not admit a potential function, as best response dynamics may cycle, while computing the social welfare maximizing Nash stable partition is also NP-hard. In another paper, Balliu et al. [6] considered the price of Pareto optimality in social distance games and presented asymptotically tight bounds.

Apart from social distance games and simple, symmetric fractional hedonic games, Peters and Elkind [28] investigate the computational complexity of stability-related questions in hedonic games, while Peters [27] studies the computational complexity of questions related to dichotomous hedonic games, where each player either approves or disapproves a given coalition. Barrot et al. [7] consider a model where agents can be friends, enemies, or unknown and show that stable outcomes may not exist, while Barrot and Yokoo [8] explore the interplay of stability and envy-freeness. Feldman et al. [17] consider the non-cooperative version of hedonic clustering games, where they characterize Nash equilibria and provide upper and lower bounds on the price of anarchy and price of stability. Hoefer et al. [20] study hedonic games and characterize the structures based on which dynamic coalition formation can stabilize quickly, while Hoefer and Jiamjitrak [19] consider proportional allocation for profit sharing in hedonic games

Our contribution. We present new and improved bounds on the price of stability for the classes of simple symmetric fractional hedonic games and social distance games. For fractional hedonic games, we improve upon the lower bound of [10] and show a lower bound of $1 + \sqrt{6}/2 \approx 2.224$ for general graphs. To do this, we present a construction where the optimal partition consists of cliques of different sizes, while the only Nash stable partition is the grand coalition, where all players form a single cluster. Then, we consider games played on graphs of girth at least 5, i.e., graphs without cycles of size 3 and 4. We prove that the price of stability, for this class of graphs, is 1. This result complements a result of Bilò et al. [10] that there exists a bipartite graph with price of stability at least 1.003. Since bipartite graphs have no cycles of length 3 but may have cycles of length 4, we obtain a clear separation of which girth values lead to price of stability equal to 1. Our final result, for this class, concerns modified fractional hedonic games, where the average utility is computed with respect to the cluster size minus 1, i.e., each player only considers the average utility it obtains by the nodes it is grouped together with. We show that the price of stability of simple symmetric fractional hedonic games is 1; the previously known bounds were those obtained in [10] for the standard utility function.

For social distance games, we improve upon the lower bound of $6/5$ from [5] and present a lower bound of 2 for general graphs. The construction we use in the proof admits an optimal partition consisting of clusters of diameter 1, while in the only Nash stable partition almost all nodes are at distance 2 from each other. We also consider social distance games played on trees and we prove that the price of stability is 1. Our final result concerns modified social distance games, where again the utility is computed with respect to the cluster size minus 1, and we show that, under this utility function, the price of stability of social distance games is 1.

Roadmap. The remainder of the paper is structured as follows. We begin, in Section 2, by formally introducing two classes of simple graph-based hedonic games and, then, presenting the necessary definitions. Then, in Section 3, we present improved lower bounds on the price of stability, while Section 4 contains positive results for specific graphs. In Section 5 we consider the price of stability

under a modified utility function and we conclude with open problems in Section 6.

2. Preliminaries

We consider non-cooperative graph-based coalition formation strategic games played on a graph $G = (V, E)$ by a set N of n utility maximizing players. Each node in $V(G)$ corresponds to a strategic player and a *partition* (or *clustering*) of the game consists of a set $\mathcal{C} = \{C_1, C_2, \dots\}$ of clusters such that $\cup_i C_i = N$, and $C_i \cap C_j = \emptyset$ for any pair $i \neq j$, i.e., each player belongs to exactly one cluster. We let $C(i)$ denote the cluster that player i belongs in.

Given a partition \mathcal{C} , each player i has a utility function $u_i : N \rightarrow \mathbb{R}$ that denotes how much player i values each of the remaining players and depends also on \mathcal{C} . We denote by w_{ij} the value that player i obtains from player j and we consider only the symmetric case $w_{ij} = w_{ji}$; we set $w_{ii} = 0$ for any player i . The utility of player i is then defined as

$$u_i(C(i)) = \frac{\sum_{j \in C(i) \setminus \{i\}} w_{ij}}{|C(i)|},$$

i.e., each player only considers the players in her cluster.

In the class of *simple symmetric fractional hedonic games*, it holds that w_{ij} is either 0 or 1; recall that $w_{ij} = w_{ji}$. In particular, for any two players i and j that are in the same cluster in \mathcal{C} , we have $w_{ij} = 1$ if the edge (i, j) exists in the graph G , while if edge (i, j) does not exist in the graph G , then $w_{ij} = 0$. Let $\deg_G(i)$ denote the degree of node i in graph G and let $\deg_C(i)$ denote the number of neighbors of node i that belong to cluster C . Given a partition \mathcal{C} , the utility of player i that is in cluster $C(i)$ is then equal to

$$u_i(C(i)) = \frac{\deg_{C(i)}(i)}{|C(i)|}.$$

In the class of *social distance games*, for any two players i and j that are in the same cluster in \mathcal{C} , we have $w_{ij} = \frac{1}{d_{C(i)}(i, j)}$, where $d_{C(i)}(i, j)$ is the distance of players i and j in the subgraph defined by cluster $C(i)$. In case i and j are disconnected in cluster $C(i)$, then $d_{C(i)}(i, j) = \infty$. Hence, the utility of player i is given by

$$u_i(C(i)) = \frac{1}{|C(i)|} \sum_{j \in C(i) \setminus \{i\}} \frac{1}{d_{C(i)}(i, j)}.$$

Clearly, in both game classes, for any player i and any partition \mathcal{C} it holds that $0 \leq u_i(C(i)) \leq \frac{n-1}{n}$. The *social welfare* $\text{SW}(\mathcal{C})$ of partition \mathcal{C} is defined as the sum of the players' utility, i.e., $\text{SW}(\mathcal{C}) = \sum_i u_i(C(i))$. We denote by \mathcal{C}^* the partition that maximizes the social welfare. An equivalent way to define the social welfare for simple symmetric fractional hedonic games is by taking into account the number of edges $E(C)$ inside each cluster C . Hence, we obtain that, in these games, $\text{SW}(\mathcal{C}) = \sum_{C \in \mathcal{C}} \frac{2E(C)}{|C|}$.

Since each player is utility maximizing, given a partition \mathcal{C} , player i may deviate from its current cluster $C(i)$ in \mathcal{C} and join another cluster C' , if it holds that $u_i(C(i)) < u_i(C' \cup i)$. A player i is *Nash stable* if there is no cluster $C' \neq C(i)$ such that its utility improves by deviating to C' and, similarly, a cluster is *Nash stable* if all players in the cluster are Nash stable. A partition is a *Nash stable partition* if all clusters are Nash stable.

The *price of stability* PoS (introduced in [1]) denotes the best-case performance deterioration arising from the requirement that the resulting partition is Nash stable. Given a graph G , the corresponding game Γ_G and its set of Nash stable partitions \mathcal{C}_s , the price of stability for the game Γ_G is formally defined as $\text{PoS}(\Gamma_G) = \max_{\mathcal{C} \in \mathcal{C}_s} \frac{\text{SW}(\mathcal{C}^*)}{\text{SW}(\mathcal{C})}$. Similarly, the price of stability for a class of strategic games is defined as $\text{PoS} = \max_G \text{PoS}(\Gamma_G)$.

We also consider a variant of these graph-based hedonic games, where the single difference is that the utility of player i is now defined as

$$u'_i(C(i)) = \begin{cases} \frac{\sum_{j \in C(i) \setminus \{i\}} w_{ij}}{|C(i)|-1}, & \text{if } |C(i)| > 1; \\ 0, & \text{otherwise.} \end{cases}$$

This variant is motivated by similar considerations in Olsen [26] and gives rise to *modified fractional hedonic games* and *modified social distance games*.

3. Lower bounds for general graphs

In this section we present our lower bounds on the price of stability. We begin by presenting an improved lower bound of $1 + \sqrt{6}/2 \approx 2.224$ for simple symmetric fractional hedonic games, where our construction extends in a non-trivial way the graph used in the lower bound of [10]. Then, we prove a lower bound of 2 for social distance games; in both cases, we begin by describing a graph and, then, we show that, for this graph, the grand coalition is the only Nash stable partition while the optimal partition contains cliques of different sizes.

In our proofs, we exploit the following technical lemma; in any stable partition, any two players having the same closed neighborhood must belong to the same cluster. Note that given a node i , its neighborhood \mathcal{N}_i is the set $\{j : (i, j) \in E(G)\}$.

Lemma 1. *For any two players x, y such that $(x, y) \in E(G)$, if $\mathcal{N}_x \cup x = \mathcal{N}_y \cup y$, then, in any stable partition, x and y are in the same cluster, i.e., $C(x) = C(y)$.*

PROOF. Assume otherwise and consider a stable partition \mathcal{C} such that there exist two neighboring players x and y with $\mathcal{N}_x \cup x = \mathcal{N}_y \cup y$, and $C(x) \neq C(y)$. Let x_i (respectively, y_i) denote the number of players in $C(x)$ (respectively, in $C(y)$) that are at distance i from x (respectively, from y). Then, in simple symmetric fractional hedonic games, the utility of player x is $u_x(C(x)) = \frac{x_1}{1 + \sum_i x_i}$, and the utility of player y is $u_y(C(y)) = \frac{y_1}{1 + \sum_i y_i}$, while, in social distance games,

the utility of player x is $u_x(C(x)) = \frac{\sum_i x_i/i}{1+\sum_i x_i}$, and the utility of player y is $u_y(C(y)) = \frac{\sum_i y_i/i}{1+\sum_i y_i}$.

Since \mathcal{C} is stable, neither x nor y gain by deviating from $C(x)$ and $C(y)$. In particular, in simple symmetric fractional hedonic games, the utility of player x when joining cluster $C(y)$ is $u_x(C(y) \cup x) = \frac{1+y_1}{2+\sum_i y_i}$, and the utility of player y when joining cluster $C(x)$ is $u_y(C(x) \cup y) = \frac{1+x_1}{2+\sum_i x_i}$. Similarly, in social distance games, the utility of player x when joining cluster $C(y)$ is $u_x(C(y) \cup x) = \frac{1+\sum_i y_i/i}{2+\sum_i y_i}$, since $(x, y) \in E(G)$ and any player k (other than x, y) is equidistant from x and y , and the utility of player y when joining cluster $C(x)$ is $u_y(C(x) \cup y) = \frac{1+\sum_i x_i/i}{2+\sum_i x_i}$.

Hence, by the stability of \mathcal{C} , it holds that $u_x(C(x)) \geq u_x(C(y) \cup x)$ and $u_y(C(y)) \geq u_y(C(x) \cup y)$. The lemma follows by observing that, since $u_x(C(x)) < 1$ and $u_y(C(y)) < 1$, we obtain $u_x(C(y) \cup x) > u_y(C(y))$, and, similarly, $u_y(C(x) \cup y) > u_x(C(x))$. Hence, it must hold that

$$u_x(C(x)) \geq u_x(C(y) \cup x) > u_y(C(y)) \geq u_y(C(x) \cup y) > u_x(C(x)),$$

i.e., we reach a contradiction. \square

3.1. Fractional hedonic games

Let α be a positive integer and consider the following graph G_1 (see also Figure 1). It consists of $\alpha + 2$ cliques K^κ where $1 \leq \kappa \leq \alpha + 2$ and some additional nodes and edges to be detailed later. Clique K^1 contains $4(\sqrt{6}+1)\alpha^2$ nodes¹, clique K^2 contains $4\alpha^2 + 2$ nodes, while each remaining clique contains 4α nodes. There exist $4(\sqrt{6}+1)\alpha^2$ additional nodes where each of them has degree 1 and is connected to a node in K^1 so that no pair of additional nodes shares a neighbor. The total number of nodes is $n = 8(2 + \sqrt{6})\alpha^2 + 2$. There exist additional edges as follows: each node in K^2 is connected to any node in K^κ , for $\kappa \in [1, \dots, \alpha + 2]$. The total number of edges in G_1 is

$$\begin{aligned} E(G_1) &= 4(\sqrt{6}+1)\alpha^2 + \frac{4(\sqrt{6}+1)\alpha^2(4(\sqrt{6}+1)\alpha^2 - 1)}{2} + \frac{(4\alpha^2 + 2)(4\alpha^2 + 1)}{2} \\ &\quad + \alpha \frac{4\alpha(4\alpha - 1)}{2} + (4\alpha^2 + 2)(4(\sqrt{6}+1)\alpha^2 + 4\alpha^2) \\ &= 32(3 + \sqrt{6})\alpha^4 + 8\alpha^3 + 2(11 + 5\sqrt{6})\alpha^2 + 1, \end{aligned}$$

where, in the first equality, the first term is due to the edges connecting the additional nodes to the nodes in K^1 , the second term is due to the edges inside K^1 , the third term is due to the edges inside K^2 , the fourth term is due to the edges inside the remaining cliques, while the last term is due to edges connecting nodes of K^2 to nodes in other cliques.

¹Strictly speaking, $|K^1|$ should be either $\lceil 4(\sqrt{6}+1)\alpha^2 \rceil$ or $\lfloor 4(\sqrt{6}+1)\alpha^2 \rfloor$ but the proof still follows in the same way. We set $|K^1| = 4(\sqrt{6}+1)\alpha^2$ to keep the presentation cleaner.

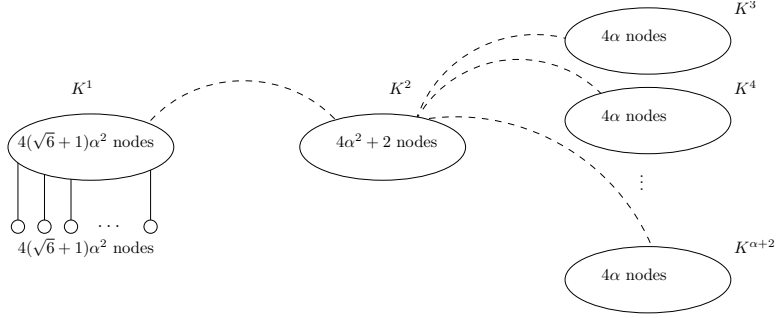


Figure 1: The graph G_1 used in the lower bound. Each bubble is a clique and dashed lines represent the edges connecting any node in K^2 to any node at another clique.

Consider the partition \mathcal{C} where each node in K^1 forms a cluster with its additional neighbor, nodes in K^2 and K^3 form a single cluster together, while each K^κ , for $\kappa \geq 4$, is a cluster. Then, the social welfare is

$$\begin{aligned}
\text{SW}(\mathcal{C}) &= \sum_{C \in \mathcal{C}} \frac{2E(C)}{|C|} \\
&= |K^1| + (|K^2| + |K^3| - 1) + \sum_{\kappa=4}^{\alpha+2} (|K^\kappa| - 1) \\
&= 4(\sqrt{6} + 1)\alpha^2 + 4\alpha^2 + 4\alpha + 1 + (\alpha - 1)(4\alpha - 1) \\
&= 4(3 + \sqrt{6})\alpha^2 - \alpha + 2.
\end{aligned}$$

Clearly, for the optimal partition \mathcal{C}^* it holds that $\text{SW}(\mathcal{C}^*) \geq \text{SW}(\mathcal{C})$, hence

$$\text{SW}(\mathcal{C}^*) \geq 4(3 + \sqrt{6})\alpha^2 - \alpha + 2. \quad (1)$$

Consider now the partition \mathcal{C}' where all nodes form a single cluster, i.e., the grand coalition. Its social welfare is

$$\begin{aligned}
\text{SW}(\mathcal{C}') &= \frac{2E(G_1)}{n} \\
&= \frac{32(3 + \sqrt{6})\alpha^4 + 8\alpha^3 + 2(11 + 5\sqrt{6})\alpha^2 + 1}{4(2 + \sqrt{6})\alpha^2 + 1}. \quad (2)
\end{aligned}$$

We now show that \mathcal{C}' is the only Nash stable partition; this key property will allow us to prove the main result of this section.

Lemma 2. \mathcal{C}' is the only Nash stable partition for graph G_1 .

PROOF. First, observe that, in any Nash stable partition, any additional node is in the same cluster as its neighbor in K^1 as, otherwise, its utility would be 0. Next, note that, due to Lemma 1, in any Nash stable partition and for each

$\kappa \geq 2$, all nodes in K^κ must be in the same cluster as they have the same closed neighborhood.

We now argue that all nodes in K^1 and the additional nodes must be in the same cluster as those in K^2 . Consider a node $i \in K^1$. The utility of i in any cluster containing only additional nodes and nodes in K^1 is exactly $1/2$. By deviating to the cluster C that contains the clique K^2 , the utility of i becomes at least $\frac{a_i+4\alpha^2+2}{2a_i+4\alpha^2+2+4\alpha^2+1} > 1/2$, where a_i is the number of nodes from K^1 in C . Indeed, the worst case is that C contains a_i nodes from K^1 , the corresponding a_i additional nodes, and all nodes in $\cup_{\kappa \geq 2} K^\kappa$. Hence, the nodes in $K^1 \cup K^2$ as well as the additional nodes are necessarily in the same cluster in a Nash stable partition.

We conclude the argument that \mathcal{C}' is the only Nash stable partition by showing that there cannot be another Nash stable partition where a node $i \in K^2$ is not together with all nodes in $\cup_{\kappa \geq 3} K^\kappa$. Assume otherwise that i is not in the same cluster as the nodes in $K^{\kappa'}$, for some $\kappa' \geq 3$. Then the utility of node $i \in K^2$ that belongs to cluster $C(i)$ is at most $u_i(C(i)) \leq \frac{4(\sqrt{6}+1)\alpha^2+4\alpha^2+1+4\alpha(\alpha-1)}{8(\sqrt{6}+1)\alpha^2+4\alpha^2+2+4\alpha(\alpha-1)} = \frac{4(3+\sqrt{6})\alpha^2-4\alpha+1}{8(2+\sqrt{6})\alpha^2-4\alpha+2}$. By deviating to the cluster containing the nodes in $K^{\kappa'}$, the utility of i is at least $\frac{4\alpha}{4\alpha+1} > u_i(C(i))$; a contradiction. Hence, we have proven that the only Nash stable partition is the grand coalition. \square

We are now ready to prove the main result of this section for simple symmetric fractional hedonic games.

Theorem 3. *The price of stability of simple symmetric fractional hedonic games is at least $1 + \frac{\sqrt{6}}{2} - \epsilon$ for $\epsilon > 0$.*

PROOF. By combining inequalities (1) and (2), we conclude that the price of stability for graph G_1 is

$$\begin{aligned} \text{PoS}(G_1) &= \frac{\text{SW}(\mathcal{C}^*)}{\text{SW}(\mathcal{C}')} \\ &\geq \frac{(4(3+\sqrt{6})\alpha^2 - \alpha + 2)(4(2+\sqrt{6})\alpha^2 + 1)}{32(3+\sqrt{6})\alpha^4 + 8\alpha^3 + 2(11+5\sqrt{6})\alpha^2 + 1} \\ &= \frac{16(12+5\sqrt{6})\alpha^4 - 4(2+\sqrt{6})\alpha^3 + 4(7+3\sqrt{6})\alpha^2 - \alpha + 2}{32(3+\sqrt{6})\alpha^4 + 8\alpha^3 + 2(11+5\sqrt{6})\alpha^2 + 1} \\ &\geq 1 + \frac{\sqrt{6}}{2} - \epsilon \end{aligned}$$

as α tends to infinity, where ϵ is an arbitrarily small positive number. \square

3.2. Social distance games

Let α be a positive integer. Consider the following graph G_2 that is also presented in Figure 2. There exist two sets S, S' of α nodes each, where each node $i \in S$ is connected only to the corresponding node $i' \in S'$. There also

exist $\alpha/2$ cliques K^j , $1 \leq j \leq \alpha/2$, each of size α . Any node in clique K^1 is connected to all nodes in $S' \cup (\bigcup_{i \geq 2} K^i)$. Therefore, the total number of nodes in G_2 is $\alpha^2/2 + 2\alpha$.

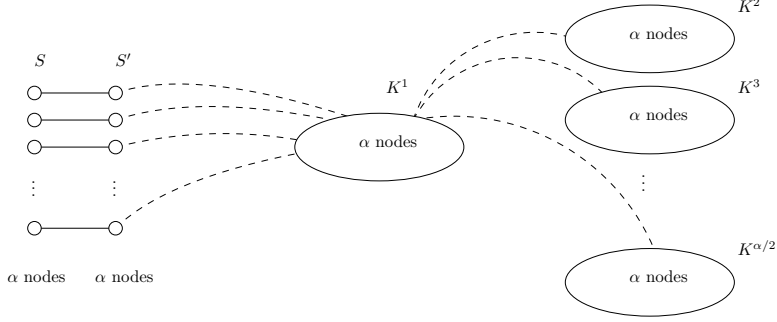


Figure 2: The graph G_2 used in the proof of the lower bound. Each bubble corresponds to a clique and each dashed line represents edges adjacent to all clique nodes.

We first argue about the social welfare in the optimal partition and then we argue that the grand coalition is the only Nash stable partition. Consider the partition \mathcal{C} where each player $i \in S$ is paired with the neighboring player $i' \in S'$, cliques K^1 and K^2 form a cluster, while each clique K^j , for $3 \leq j \leq \alpha/2$ forms a cluster. The social welfare of this partition is

$$\begin{aligned} \text{SW}(\mathcal{C}) &= \alpha + (2\alpha - 1) + (\alpha/2 - 2)(\alpha - 1) \\ &= \alpha^2/2 + \alpha/2 + 1, \end{aligned}$$

where the first term in the first equality is due to the α clusters containing the players in $S \cup S'$, the second term is due to the cluster $K^1 \cup K^2$, while the last term is due to the remaining $(\alpha/2 - 2)$ clusters, each containing a single clique of size α . Clearly, for the optimal partition \mathcal{C}^* it holds that $\text{SW}(\mathcal{C}^*) \geq \text{SW}(\mathcal{C})$, hence, it holds that

$$\text{SW}(\mathcal{C}^*) \geq \alpha^2/2 + \alpha/2 + 1. \quad (3)$$

Consider now the partition \mathcal{C}' where all players form a single cluster, i.e., the grand coalition. Its social welfare is

$$\begin{aligned} \text{SW}(\mathcal{C}') &= \frac{\alpha(1 + \frac{\alpha}{2} + \frac{\alpha^2/2-1}{3} + \frac{\alpha-1}{4}) + \alpha(\alpha + 1 + \frac{\alpha^2/2-1}{2} + \frac{\alpha-1}{3})}{\alpha^2/2 + 2\alpha} \\ &+ \frac{\alpha(\alpha^2/2 + \alpha - 1 + \frac{\alpha}{2}) + (\alpha^2/2 - \alpha)(2\alpha - 1 + \frac{\alpha^2/2-\alpha}{2} + \frac{\alpha}{3})}{\alpha^2/2 + 2\alpha} \\ &= \frac{\alpha(2\alpha^2 + 9\alpha + 5)/12 + \alpha(3\alpha^2 + 16\alpha + 2)/12 + \alpha(\alpha^2 + 3\alpha - 2)/2}{\alpha^2/2 + 2\alpha} \\ &+ \frac{(\alpha^2/2 - \alpha)(3\alpha^2 + 22\alpha - 12)/12}{\alpha^2/2 + 2\alpha} \end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha(3\alpha^3 + 38\alpha^2 + 30\alpha + 14)/24}{\alpha^2/2 + 2\alpha} \\
&= \frac{3\alpha^3 + 38\alpha^2 + 30\alpha + 14}{12\alpha + 48}, \tag{4}
\end{aligned}$$

where, in the first equality, the first term is due to the α players in S , the second term is due to the α players in S' , the third term is due to the α players in clique K^1 , while the last term is due to the $(\alpha^2/2 - \alpha)$ players in the remaining cliques.

Lemma 4. *C' is the only Nash stable partition for G_2 .*

PROOF. First, observe that in any stable partition, any player $i \in S$ must be in the same cluster as its neighbor $i' \in S'$. If this is not the case, then $u_i(C(i)) = 0$ as there is no path inside $C(i)$ connecting i to any other player in $C(i)$, while by joining cluster $C(i')$ the utility of player i would be strictly positive. Then, due to Lemma 1, observe that, in any stable partition, all players in a clique K^j , where $1 \leq j \leq \alpha/2$, must belong in the same cluster, as by the construction of graph G_2 , any pair of players in K^j has the same closed neighborhood.

We continue by showing that any player in S' belongs to the same cluster as the players in clique K^1 . Assume otherwise and consider such a player $i' \in S'$. By the discussion above, i' is in the same cluster as its neighboring player $i \in S$ and, by our assumption that i' is not in the same cluster as clique K^1 , i' has no path inside $C(i')$ connecting it to any other player in $C(i')$, hence $u_{i'}(C(i')) \leq 1/2$. Consider the cluster C that contains the players of clique K^1 and let us assume that C contains also κ players of S' (but different than i'), where $0 \leq \kappa < \alpha$ and the players of λ additional cliques $K^{j>1}$, where $0 \leq \lambda < \alpha/2$; by the discussion above, C also contains κ players of S . Then, the utility of player i' when joining cluster C is

$$u_i(C \cup i') = \frac{\alpha + (\kappa + \lambda\alpha)/2 + \kappa/3}{(\lambda + 1)\alpha + 2\kappa + 1},$$

which is strictly greater than $1/2$ for any $\kappa < \alpha$.

So far we have established that, in any Nash stable partition, the players of S , S' , and K^1 are necessarily in the same cluster, while the players of any clique $K^{j>1}$ are also together in a cluster; note that we have not yet ruled out the possibility that different cliques belong to different clusters. We conclude the argument that the grand coalition is the only Nash stable partition by showing that if there exists a clique $K^{j>1}$ that is not in the same cluster as the clique K^1 , then any player in K^1 has an incentive to deviate and join the same cluster as the players of $K^{j'}$. Indeed, in this case, the utility of any player k in K^1 is maximized whenever $C(k)$ contains, apart from players in $S \cup S' \cup K^1$, all players except those in $K^{j'}$ and, hence, it holds that

$$\begin{aligned}
u_k(C(k)) &\leq \frac{(\alpha/2 - 2)\alpha + 2\alpha - 1 + \frac{\alpha}{2}}{(\alpha/2 - 2)\alpha + 3\alpha} \\
&= \frac{\alpha^2 + \alpha - 2}{\alpha^2 + 2\alpha}
\end{aligned}$$

$$= \frac{\alpha - 1}{\alpha},$$

where the first term in the inequality is due to the distance to the other cliques, the last term is due to the distance to players in S while the remaining terms are due to the distance to players in S' and the remaining players of K^1 . By deviating to cluster $C_{j'}$ containing at least the players of clique $K^{j'}$, player k would obtain utility $u_k(C_{j'} \cup k) \geq \frac{\alpha}{\alpha+1}$,² i.e., strictly greater utility than $u_k(C(k))$. This concludes the argument that the grand coalition is the only Nash stable partition. \square

We are now ready to prove the main result of this section.

Theorem 5. *The price of stability of social distance games is at least $2 - \epsilon$ for $\epsilon > 0$.*

PROOF. By combining inequalities (3) and (4), we conclude that the price of stability for graph G_2 is

$$\begin{aligned} \text{PoS}(G_2) &= \frac{\text{SW}(\mathcal{C}^*)}{\text{SW}(\mathcal{C}')} \\ &\geq \frac{\alpha^2/2 + \alpha/2 + 1}{(3\alpha^3 + 38\alpha^2 + 30\alpha + 14)/(12\alpha + 48)} \\ &= \frac{6\alpha^3 + 30\alpha^2 + 36\alpha + 48}{3\alpha^3 + 38\alpha^2 + 30\alpha + 14} \\ &\geq 2 - \epsilon, \end{aligned}$$

as α tends to infinity, where ϵ is an arbitrarily small positive number. \square

4. Positive results for special cases

This section contains our upper bounds on the price of stability of simple symmetric fractional hedonic games (see Section 4.1) and social distance games (in Section 4.2). Recall that Bilò et al. [10] have shown that for graphs of girth 3 and 4 the price of stability can be strictly greater than 1 for the case of fractional hedonic games, so our first result below, see Theorem 9, is essentially the best possible with respect to girth.

4.1. Fractional hedonic games and girth at least 5

We consider the class of triangle-free and quadrilateral-free graphs, i.e., when there are no cycles of length 3 or 4, and we will make use of the following result by Dutton and Brigham [15] that upper-bounds the number of edges in a graph of girth at least 5.

² $u_k(C_{j'} \cup k)$ obtains its minimum value when $C_{j'}$ contains a single clique, i.e., $K^{j'}$.

Lemma 6 (Theorem 4 in [15]). *The maximum number of edges in a graph with n nodes and girth at least 5 is $\frac{n\sqrt{n-1}}{2}$.*

Bilò et al. [10] have shown that the price of stability is 1 for the case of trees, and, furthermore, that any optimal partition consists of clusters with diameter either 1 or 2, i.e., either single edges or non-degenerate stars. Therefore, in the following, we focus on graphs having at least 5 nodes which are not trees. We now prove the key lemma that there exists at least one optimal partition where all clusters are stars. In particular, we show that starting from an optimal partition having at least one non-star cluster, we can transform it to another partition with the same social welfare and strictly more star clusters. By repeatedly performing such transformations, we end up with an optimal partition where all clusters are stars. Then, we argue that this optimal partition must also be Nash stable; clearly, this implies a price of stability of 1.

Lemma 7. *Given a graph $G = (V, E)$ with girth at least 5, there exists an optimal partition where all clusters are stars.*

PROOF. Assume otherwise and consider an optimal partition \mathcal{C}^* having a cluster C that is not a star. Since C is not a star and the girth is at least 5, it holds that $|C| \geq 4$. We claim that C necessarily contains an edge (u, v) that is not a bridge, as, otherwise, C is a tree and, by [10], a star. We decompose C into two clusters $C_1 = \{u, v\}$ and $C_2 = C \setminus \{u, v\}$.

Let c_2 and e_2 be the number of nodes and edges, respectively, in C_2 and note that $c_2 \geq 2$ as $|C| \geq 4$. Similarly, let $\mathcal{N}_u = \{j \in C_2 : (u, j) \in E\}$ and $\mathcal{N}_v = \{j \in C_2 : (v, j) \in E\}$ and, furthermore, let $e_{12} = |\mathcal{N}_u| + |\mathcal{N}_v|$ be the total number of edges with one endpoint in C_1 and the other in C_2 . Then, the total utility of players in C is $u(C) = \frac{2+2e_2+2e_{12}}{2+c_2}$, as C_1 contains two nodes and a single edge. Similarly, the total utility of players in C_1 is $u(C_1) = 1$ and the total utility of players in C_2 is $u(C_2) = \frac{2e_2}{c_2}$.

Let $\Delta = u(C_1) + u(C_2) - u(C)$ and note that it suffices to prove that $\Delta \geq 0$. It holds that

$$\begin{aligned} \Delta &= 1 + \frac{2e_2}{c_2} - \frac{2 + 2e_2 + 2e_{12}}{2 + c_2} \\ &= \frac{2c_2 + c_2^2 + 4e_2 + 2e_2c_2 - 2c_2 - 2e_2c_2 - 2e_{12}c_2}{c_2(2 + c_2)} \\ &\geq \frac{c_2^2 + 4c_2 - 4 - 2e_{12}c_2}{c_2(2 + c_2)} \\ &\geq \frac{c_2^2 + 2c_2 - 2e_{12}c_2}{c_2(2 + c_2)}, \end{aligned}$$

where the first inequality holds since $e_2 \geq c_2 - 1$ and the last inequality holds since $c_2 \geq 2$.

Clearly, $\Delta \geq 0$ when $e_{12} \leq 1 + c_2/2$, so let us assume that $e_{12} > 1 + c_2/2$. First, observe that, since graph G has girth at least 5 and the edge $(u, v) \in E$,

clearly any two nodes in \mathcal{N}_u (respectively, in \mathcal{N}_v) are at distance at least 3. Similarly, any node in \mathcal{N}_u is at distance at least 2 from any node in \mathcal{N}_v . Since C_2 is connected, this implies that there exist at least $e_{12}/2$ nodes in $C_2 \setminus (\mathcal{N}_u \cup \mathcal{N}_v)$ having at least one neighbor in $\mathcal{N}_u \cup \mathcal{N}_v$. Let \mathcal{D} be a decomposition of C_2 into tree clusters so that each tree cluster contains at most one endpoint (out of the e_{12} endpoints in C_2) connected to the same node in C_1 and at most two endpoints in total. In the worst case, this decomposition contains $e_{12}/2$ tree clusters, each of utility at least 1. Consider now the partition \mathcal{C}' that, instead of cluster C that contains $c_2 + 2$ nodes, includes cluster C_1 and the tree clusters in \mathcal{C}' . It holds that

$$\begin{aligned} u(\mathcal{C}') - u(\mathcal{C}^*) &\geq 1 + \frac{e_{12}}{2} - \frac{2 + 2e_{12} + 2e_2}{2 + c_2} \\ &> 1 + \frac{2 + c_2}{4} - \sqrt{c_2 + 1} \\ &\geq 0, \end{aligned}$$

where the second inequality holds since $e_{12} > 1 + c_2/2$ and, by Lemma 6, $1 + e_{12} + e_2 \leq \frac{(2+c_2)\sqrt{c_2+1}}{2}$, while the last inequality holds since $c_2 \geq 2$. Hence, assuming $e_{12} > 1 + c_2/2$ contradicts the optimality of \mathcal{C}^* .

Overall, we have shown that a non-star cluster C can be split into two clusters C_1 and C_2 where C_1 is a star, without decreasing the social welfare. By iterating this process, we can reach an optimal partition where all clusters are stars. \square

The following result of Bilò et al. [9] states that in a triangle-free graph, the partition, that maximizes the social welfare among partitions consisting only of stars, is stable. Clearly, the optimal partition obtained by Lemma 7 is such a partition.

Lemma 8 (Lemma 3 in [9]). *Let G be a triangle-free graph, then any optimal star partition is stable.*

By combining Lemmas 7 and 8, we obtain the main result of this section.

Theorem 9. *The price of stability of simple symmetric fractional hedonic games on graphs of girth at least 5 is 1.*

4.2. Social distance games in trees

We now consider the case on social distance games where the graph G is a tree, and we prove that there exists an optimal partition that is stable. We call a partition *compact* if it consists only of stars and paths of 4 nodes and we begin by showing that at least one optimal partition is compact.

Lemma 10. *In social distance games on trees, there exists a compact optimal partition.*

PROOF. Consider a cluster C in the optimal partition that is neither a star nor a path of 4 nodes and let h be the depth of the subtree T_C defined by cluster C . By selecting a node i in depth $h - 1$ that has at least one neighbor of depth h , we can decompose T_C into a subtree T_1 that is a star having i as the center node and i 's descendants in T_C as leaf nodes, and another subtree T_2 containing the remaining nodes. Let $a \geq 1$ be the number of node i 's neighbors in T_1 , let j be the node in subtree T_2 that is the neighbor of node i in T_C , and $b \geq 1$ be the number of additional nodes in subtree T_2 .

Since C is part of an optimal partition, the utility $u(C)$ is at least the utility of the clusters defined by subtree T_1 and the optimal decomposition of T_2 into stars, i.e.,

$$u(C) \geq \frac{a}{a+1} + \frac{a}{2} + \frac{b+1}{2}, \quad (5)$$

where the first two terms are due to the cluster of subtree T_1 , while the last term holds since T_2 contains $b+1$ players and by Corollary 1 in [13] that bounds from below the social welfare in an optimal partition. Next, observe that the utility of cluster C is maximized when player j is directly connected to all remaining b players in T_2 , as in this way the number of node pairs inside T_2 at distance 2 is maximized, while the number of node pairs at distance 1 remains the same. Let C' be a cluster with this property, i.e., j is directly connected in C' to all b remaining players of subtree T_2 ; clearly, $u(C') \geq u(C)$ which, due to (5), implies that it must hold that $u(C') \geq \frac{a}{a+1} + \frac{a+b+1}{2}$, i.e.,

$$\begin{aligned} & \frac{a(1 + \frac{a}{2} + \frac{b}{3}) + a + 1 + \frac{b}{2} + b + 1 + \frac{a}{2} + b(1 + \frac{b}{2} + \frac{a}{3})}{a + b + 2} - \frac{a}{a+1} - \frac{a+b+1}{2} \\ &= \frac{3a^2 + 4ab + 15a + 3b^2 + 15b + 12}{6(a+b+2)} - \frac{a^2 + ab + 4a + b + 1}{2(a+1)} \\ &= \frac{3b + 3 - a^2b - ab}{3(a+1)(a+b+2)} \geq 0. \end{aligned} \quad (6)$$

Observe that (6) holds only when $a = 1$ and $b \geq 1$, or when $a = 2$ and $b = 1$. In the latter case, (6) holds with equality and, hence, without loss of generality we can decompose cluster C into two star clusters, having nodes i and j as center nodes, without decreasing the social welfare.

So, in the following, we assume $a = 1$ and distinguish on whether j is directly connected to all remaining b players of T_2 or not. In the first case, it holds that T_2 is a star centered at j and, since C is part of the optimal partition, $u(C)$ should be at least the utility of the two star clusters centered at i and j , i.e., since $a = 1$, we obtain that $u(C) \geq 1 + b/(b+1) + b/2$ which implies

$$\begin{aligned} & \frac{(1 + 1/2 + b/3) + (2 + b/2) + (b+1 + 1/2) + b(1 + b/2 + 1/3)}{b+3} - 1 - \frac{b}{b+1} - \frac{b}{2} \\ &= \frac{b^2/2 + 19b/6 + 5}{b+3} - \frac{b^2 + 5b + 2}{2(b+1)} \end{aligned}$$

$$= \frac{2-b}{3(b+1)} \geq 0. \quad (7)$$

Observe that (7) can only hold if $b = 1$, i.e., for a path of 4 nodes, or if $b = 2$ when again, as (7) holds with equality in this case, we can decompose C into two star clusters without decreasing the social welfare.

It remains to consider the case where in cluster C , player j is directly connected to $b - c$ nodes of subtree T_2 , for some $c < b$, while the remaining c nodes are at distance at least 2 from j ; again, the utility in this case is maximized when these c nodes are directly connected to a neighbor of j in T_2 and is

$$\begin{aligned} u(C) &= \frac{1 + \frac{1}{2} + \frac{b-c}{3} + \frac{c}{4} + 2 + \frac{b-c}{2} + \frac{c}{3} + b - c + 1 + \frac{c+1}{2}}{b+3} \\ &+ \frac{(b-c-1)(1 + \frac{b-c}{2} + \frac{c+1}{3}) + c + 1 + \frac{b-c}{2} + \frac{1}{3} + c(1 + \frac{c}{2} + \frac{bc}{3} + \frac{1}{4})}{b+3} \\ &= \frac{3b^2 - 2bc + 19b + 2c^2 - c + 30}{6(b+3)}. \end{aligned}$$

Since C is in an optimal partition and we have assumed $a = 1$, by (5) we obtain

$$u(C) \geq 1 + \frac{b+1}{2} = \frac{b+3}{2},$$

that is, it must hold

$$\frac{3b^2 - 2bc + 19b + 2c^2 - c + 30}{6(b+3)} - \frac{b+3}{2} = \frac{2c^2 + b + 3 - 2bc - c}{6(b+3)} \geq 0,$$

which, since $b > c$, can only hold when $b = 2$ and $c = 1$ or $b = 3$ and $c \leq 2$. When $b = 2$ and $c = 1$, we have $u(C) = \frac{77}{30}$, while by considering the two star clusters that include player i and its neighbor, and player j and two additional nodes, we obtain utility $\frac{8}{3} > \frac{77}{30}$; a contradiction. Similarly, if $b = 3$ and $c = 1$, then $u(C) = \frac{109}{36}$, while for the two clusters obtained by decomposing C into a path of 4 nodes and a path of 2 nodes, we have utility $\frac{19}{6} > \frac{109}{36}$. Finally, when $b = 3$ and $c = 2$, without loss of generality we can decompose C into two star clusters without decreasing the social welfare.

By repeatedly applying this process, we obtain a partition of at least the initial social welfare, where each cluster is either a path of 4 nodes or a star. \square

Clearly, if the compact optimal partition is stable, then the price of stability is 1. Starting from an unstable compact partition, we can obtain another compact partition with at least the same social welfare and fewer paths of 4 nodes.

Lemma 11. *A compact partition \mathcal{C} is either stable or it can be turned into another compact partition \mathcal{C}' , where $\text{SW}(\mathcal{C}) \leq \text{SW}(\mathcal{C}')$ and the number of paths of 4 nodes is strictly less in \mathcal{C}' than in \mathcal{C} .*

PROOF. Consider a compact partition \mathcal{C} . We argue about the stability of \mathcal{C} and we will show that when \mathcal{C} is unstable, then we can obtain another compact partition \mathcal{C}' satisfying the lemma.

In star clusters, each player that is a leaf has utility $1/2$ and by joining a path of 4 nodes would obtain utility $5/12$, if its neighbor is a path endpoint, or utility $7/15$ if its neighbor is not a path endpoint. Similarly, by joining a star, it would obtain utility of $1/2$, so no leaf player has any incentive to deviate. Each player that is a center in a star cluster has utility at least $1/2$ and, as in the case of leaf nodes, has no incentive to deviate.

In paths of 4 nodes, each player that is not a path endpoint has utility $5/8$ and would obtain strictly less utility by any deviation. A player that is a path endpoint obtains utility $11/24$ and has no incentive to deviate by connecting to an endpoint of another path of 4 nodes, or to a leaf in a star cluster of at least 3 players. When such a player has an incentive to connect as a leaf to a center of a star cluster, then by allowing this deviation we obtain another compact partition with strictly greater social welfare and a reduced number of paths of 4 nodes, as the original path has now become a star with 3 players. When a player that is a path endpoint wishes to deviate and connect to a non-endpoint node in a path of 4 nodes, then we can rearrange these two clusters, that are both paths of 4 nodes, into three star clusters of size 2, 3, and 3 with the same social welfare; again, we obtain a compact partition that satisfies the lemma. \square

By combining Lemmas 10 and 11 we obtain the following theorem, as there exists an optimal partition \mathcal{C}^* that is compact and, in case it is unstable, we can obtain another compact optimal partition with strictly fewer paths of nodes 4; clearly, this process will eventually halt at a stable compact optimal partition.

Theorem 12. *The price of stability of social distance games on trees is 1.*

5. Modified graph-based hedonic games

This section contains our results for the class of modified fractional hedonic games (in Section 5.1) and modified social distance games (in Section 5.2). These classes are based on alternative utility functions, first considered by Olsen [26], that favor coalitions of smaller size.

5.1. Fractional hedonic games

Recall that the utility of a player i belonging to cluster $C(i)$ is $u'_i(C(i)) = \frac{\deg_{C(i)}(i)}{|C(i)|-1}$. We show that, under this utility function, the price of stability of simple symmetric fractional hedonic games is 1 by arguing about the structure of the optimal partition. In particular, we show that there exists an optimal partition satisfying a desirable structure, and then we argue that this partition is Nash stable. We begin with a technical lemma.

Lemma 13. *For any integers x, y, z such that $1 \leq x \leq y \leq z - 1$, it holds that $z^2 - 3z + xy - x^2 + x + y - yz + 2 \geq 0$.*

PROOF. Fix z and let $f(x, y) = -x^2 + xy + x + (1 - z)y + z^2 - 3z + 2$. It suffices to prove that $f(x, y) \geq 0$ for any $1 \leq x \leq y \leq z - 1$. The derivative with respect to y is $f'_y(x, y) = x + 1 - z$. Note that $f'_y(x, y) < 0$ whenever $x < z - 1$ and $f'_y(x, y) = 0$ only when $x = z - 1$. In both cases, the value $y = z - 1$ minimizes $f(x, y)$. Then, $f(x, z - 1) = -x^2 + zx + 1 - z$ with $f(1, z - 1) = 0$ and $f(z - 1, z - 1) = 0$. The proof follows since $f(x, z - 1)$ is increasing up to $x = z/2$ and then becomes decreasing. \square

The next lemma specifies the structure of an optimal partition. In particular, we show that there exists an optimal partition where each cluster C is either a singleton cluster (when $|C| = 1$), a path of two nodes (when $|C| = 2$), a star with 2 leaves or a triangle (when $|C| = 3$), or a star (whenever $|C| \geq 4$).

Lemma 14. *There exists an optimal partition C^* where each cluster C with $|C| \geq 4$ is a star.*

PROOF. Consider a cluster C with $k \geq 4$ nodes that is not a star. Let ℓ be the number of edges in C ; then, the social welfare is $\text{SW}(C) = \frac{2\ell}{k-1}$. Let i be a node in C with the minimum degree, solving ties arbitrarily, and let $x = \deg(i)$. Let j be the neighbor of i with the minimum degree among all i 's neighbors, again solving ties arbitrarily, and let $y = \deg(j)$.

We argue that we can split C into two clusters, i.e., $C_1 = \{i, j\}$ and $C_2 = C \setminus C_1$ without decreasing the social welfare. The total social welfare of the two clusters is $\text{SW}' = \text{SW}(C_1) + \text{SW}(C_2) = 2 + \frac{2\ell'}{k-3}$, where ℓ' is the number of edges in cluster C_2 . It holds that $\ell = \ell' + x + y - 1$, as i (respectively, j) has $x - 1$ (respectively, $y - 1$) neighbors in C_2 while the edge (i, j) also exists in C but not in C_2 .

We now provide a lower bound on ℓ' based on x, y and k . By the definitions of x and y , it holds that C contains x nodes with degree at least y and $k - x$ nodes with degree at least x . Hence, we obtain that $\ell \geq \frac{xy + (k-x)x}{2}$. Since $\ell' = \ell - x - y + 1$, we obtain that

$$\ell' \geq \frac{kx + xy - x^2 - 2x - 2y + 2}{2}. \quad (8)$$

It suffices to prove that $\text{SW}' \geq \text{SW}(C)$. We have

$$\begin{aligned} \text{SW}' - \text{SW}(C) &= \frac{2k + 2\ell' - 6}{k - 3} - \frac{2\ell' + 2x + 2y - 2}{k - 1} \\ &= \frac{2k^2 - 6k + 4\ell' - 2kx - 2ky + 6x + 6y}{k^2 - 4k + 3} \\ &\geq \frac{2k^2 - 6k + 2kx + 2xy - 2x^2 - 4x - 4y + 4 - 2kx - 2ky + 6x + 6y}{k^2 - 4k + 3} \\ &= \frac{2k^2 - 6k + 2xy - 2x^2 + 2x + 2y - 2ky + 4}{k^2 - 4k + 3} \\ &\geq 0. \end{aligned}$$

The first inequality follows by using (8), while the last inequality holds due to Lemma 13 (by setting $z = k$) and since $k \geq 4$.

By repeating this process as long as there exists a non-star cluster of size at least 4, we obtain an optimal solution with the desired properties. \square

We now show that there exists an optimal partition that is Nash stable, hence the price of stability is 1.

Theorem 15. *The price of stability of simple symmetric modified fractional hedonic games is 1.*

PROOF. Consider an optimal partition satisfying the properties of Lemma 14. Clearly, any node in a triangle and any node that is a root in a star is satisfied since its utility is 1. The only players that may wish to deviate are nodes in singleton clusters or leaves in a star. Observe that there cannot be an edge connecting two nodes from the set of leaves and singletons, as these two nodes would form a new cluster and the social welfare would strictly increase; a contradiction to our original assumption. Similarly, there cannot be an edge connecting a leaf or singleton node i to a node j that belongs to a cluster C forming a triangle, as then the social welfare would strictly increase by creating cluster $\{i, j\}$ and reducing C to $C \setminus \{j\}$.

We first let all singleton clusters deviate and join their preferred star. Since the social welfare of any star cluster C with k nodes is $\text{SW}(C) = \frac{2(k-1)}{(k-1)} = 2$ irrespective of the number of leaves, these deviations do not decrease the social welfare. Then, all possible subsequent deviating moves (which can be made only by leaves) lead to partitions where the number of triangles and the number of stars remains the same, but the structure of these stars may change as the leaves deviate. Observe that a deviating move of node i from star $C(i)$ to another star C , requires that $|C(i)| > |C| + 1$ and strictly decreases the maximum size among these two star clusters, i.e., $|C(i)|$ in this example. Therefore, by considering the lexicographic order π of all star-clusters in the partition \mathcal{C} based on the number of nodes, from the minimum to the maximum, we observe that any deviating move that leads to partition \mathcal{C}' satisfies $\pi(\mathcal{C}) < \pi(\mathcal{C}')$ and, hence, this process is guaranteed to end. Furthermore, any deviating move does not decrease the social welfare, as the new clusters remain stars. The theorem follows. \square

5.2. Social distance games

Our results for the class of modified social distance games follow along similar lines as that in Section 5.1. In our proofs, we exploit the following technical lemma that argues about the social welfare in a partition that consists only of stars.

Lemma 16. *Any partition where m star clusters span a set of n nodes has total utility $m + n/2$.*

PROOF. Consider the m star clusters C_1, C_2, \dots, C_m and let n_i , $1 \leq i \leq m$, denote the number of nodes in cluster C_i . For a given cluster C_i , the cluster

center has utility 1 while each of the remaining $n_i - 1$ nodes has utility $(1 + (n_i - 2)/2)/(n_i - 1) = n_i/(2n_i - 2)$. By summing over all players in C_i , we obtain that $u(C_i) = 1 + n_i/2$. The lemma follows by summing over all m star clusters and since $\sum_i n_i = n$. \square

We first show that there exists an optimal partition where each cluster is a triangle, a star or a single disconnected node; we begin by an optimal partition that may not exhibit these properties and we show how to transform it into another optimal partition that satisfies them.

Lemma 17. *There exists an optimal partition \mathcal{C}^* where each cluster C with $|C| \geq 4$ is a star.*

PROOF. Consider an optimal partition \mathcal{C}^* and a cluster C with $|C| \geq 4$ nodes that is not a star. Let \mathcal{D} be a decomposition of cluster C into a collection of disjoint sets $\{T, P, I\}$ where T contains x disjoint cliques of size 3, i.e., triangles, P contains y disjoint cliques of size 2, i.e., pairs, and I contains z isolated nodes, so that z is the minimum among all such decompositions. Clearly, $|C| = 3x + 2y + z$.

We now argue about the structure inside cluster C . First, observe that there cannot be an edge connecting a node belonging to set T in \mathcal{D} to a node belonging to set I in \mathcal{D} , or connecting two nodes belonging to set I in \mathcal{D} , as then another decomposition \mathcal{D}' exists with strictly fewer isolated nodes than \mathcal{D} . Furthermore, each isolated node in \mathcal{D} may be directly connected to at most one node of each pair in \mathcal{D} as otherwise we could form a triangle and reduce the number of isolated nodes. In addition, for each pair in \mathcal{D} , at most one node may be directly connected to any isolated node, as otherwise we could split the pair, connect its endpoints to the isolated nodes and again reduce the number of isolated nodes. Let P' be the set of the $y' \leq y$ pairs in \mathcal{D} where exactly one endpoint has at least one isolated node as neighbor, while, for the remaining $y - y'$ pairs, no endpoint has an isolated node as neighbor. Furthermore, for any pair where one endpoint is directly connected to an isolated node, the other endpoint cannot be connected to a triangle node, as then we could again reduce the number of isolated nodes by rearranging the triangle, the pair, and the isolated node. Finally, any isolated node i is directly connected to at least one node (which, by the discussion above, must be a pair node), as otherwise the cluster $C \setminus \{i\}$ has strictly greater utility than C , contradicting the optimality of \mathcal{C}^* .

By the last observation about isolated nodes, it follows that we can decompose cluster C into smaller clusters such that each triangle in T forms a cluster, while there are also y star clusters that contain the nodes in P as well as the isolated nodes, i.e., $2y + z$ nodes in total. By Lemma 16, we have that the total utility $u(\mathcal{D})$ of the new set of clusters is

$$u(\mathcal{D}) = 3x + 2y + z/2. \tag{9}$$

Let us now argue about the utility of cluster C . We upperbound the utility of each player i in C according to its type in the decomposition \mathcal{D} . If i is in a

triangle,

$$\begin{aligned} u_i(C) &\leq \frac{3x + y' + 2(y - y') - 1 + (y' + z)/2}{3x + 2y + z - 1} \\ &= \frac{3x + 2y - y'/2 + z/2 - 1}{3x + 2y + z - 1}, \end{aligned} \quad (10)$$

since, by the discussion above, i can be directly connected to at most the nodes in triangles, one endpoint for each of y' pairs, both endpoints for the remaining $(y - y')$ pairs, while it can have distance at least 2 from the remaining $y' + z$ nodes.

If i is an isolated node, then

$$\begin{aligned} u_i(C) &\leq \frac{y' + (y' + 2(y - y') + 3x + z - 1)/2}{3x + 2y + z - 1} \\ &= \frac{3x/2 + y + y'/2 + z/2 - 1/2}{3x + 2y + z - 1}, \end{aligned} \quad (11)$$

since it can be directly connected to at most y' nodes and has distance at least 2 to all remaining nodes.

If i is a node in the y' pairs and has isolated nodes as neighbors, then the utility is

$$u_i(C) \leq \frac{3x + 2y + z - 1}{3x + 2y + z - 1}, \quad (12)$$

as it may be directly connected to all remaining nodes, while if i is an endpoint in one of the y' pairs and does not have isolated nodes as neighbors, then the utility is

$$u_i(C) \leq \frac{2y - 1 + (3x + z)/2}{3x + 2y + z - 1}, \quad (13)$$

since i can be directly connected only to the remaining nodes belonging to pairs in \mathcal{D} .

Finally, if i is part of the $y - y'$ pairs where no endpoint has an isolated node as neighbor, then the utility is

$$u_i(C) \leq \frac{3x + 2y - 1 + z/2}{3x + 2y + z - 1}, \quad (14)$$

as it may be directly connected to all remaining nodes apart from those that are isolated in \mathcal{D} .

The total utility of cluster C can be bounded from above by using inequalities (10)-(14) as

$$u(C) = \sum_{i \in C} u_i(C)$$

$$\begin{aligned}
&= \sum_{i \in T} u_i(C) + \sum_{i \in P'} u_i(C) + \sum_{i \in P \setminus P'} u_i(C) + \sum_{i \in I} u_i(C) \\
&\leq \frac{3x(3x + 2y - y'/2 + z/2 - 1) + y'(3x + 2y + z - 1 + 2y - 1 + (3x + z)/2)}{3x + 2y + z - 1} \\
&\quad + \frac{2(y - y')(3x + 2y - 1 + z/2) + z(3x/2 + y + y'/2 + z/2 - 1/2)}{3x + 2y + z - 1} \\
&= \frac{9x^2 + 4y^2 + z^2/2 + 12xy - 3xy' + 3xz + 2yz + y'z - 3x - 2y - z/2}{3x + 2y + z - 1}.
\end{aligned} \tag{15}$$

The proof follows as by combining (9) and (15) we obtain $u(\mathcal{D}) \geq u(C)$, since $(3x + 2y + z/2)(3x + 2y + z - 1) \geq 9x^2 + 4y^2 + z^2/2 + 12xy - 3xy' + 3xz + 2yz + y'z - 3x - 2y - z/2$ for any $y' \leq y$, i.e., we can decompose cluster C into triangles and stars without losing social welfare; clearly, we can repeatedly perform this process until all clusters are as desired. \square

We are now ready to prove the main result of this section, i.e., that there exists a stable optimal partition.

Theorem 18. *The price of stability of modified social distance games is 1.*

PROOF. Consider an optimal partition \mathcal{C}^* that exhibits the properties of Lemma 17. Clearly, if \mathcal{C}^* is stable, the theorem follows, so we assume that \mathcal{C}^* is unstable and we show how to modify it in order to obtain a stable optimal partition.

First, observe that any player that is in a triangle cluster and any player that is a root in a star cluster is satisfied since its utility is 1. Therefore, the players that may wish to deviate from \mathcal{C}^* are those that are either in singleton clusters, or leaves in a star. We now argue that players in singleton clusters are disconnected in graph G and, hence, obtain utility equal to 0 in any partition; so, they have no incentive to deviate and can be ignored, without loss of generality. Indeed, if i is connected by an edge (i, j) to another player j that is in a singleton cluster, then we can merge the two singleton clusters $\mathcal{C}^*(i)$ and $\mathcal{C}^*(j)$ to a single cluster and strictly increase the social welfare; a contradiction since \mathcal{C}^* is an optimal partition. A similar reasoning applies if player i is connected by an edge (i, k) to a player k that is in a triangle cluster or that is a leaf in a star cluster, as then we can replace $\mathcal{C}^*(i)$ and $\mathcal{C}^*(k)$ by $\mathcal{C}^*(i) \cup \{k\}$ and $\mathcal{C}^*(k) \setminus \{k\}$ and obtain strictly greater social welfare. Finally, if i is connected to a player ℓ that is root in a star cluster, we can merge clusters $\mathcal{C}^*(i)$ and $\mathcal{C}^*(\ell)$ and obtain strictly greater social welfare as $u(\mathcal{C}^*(i)) = 0$, $u(\mathcal{C}^*(\ell)) = 1 + |\mathcal{C}^*(\ell)|/2$, while $u(\mathcal{C}^*(\ell) \cup \{i\}) = 1 + (|\mathcal{C}^*(\ell)| + 1)/2$.

We now consider the players that are leaves in star clusters. As in the case of singleton clusters, a leaf player i cannot be connected by an edge to a triangle cluster or to another leaf player, as again we would obtain strictly greater welfare by letting player i and its neighbor form a new cluster. Hence, i can only deviate by becoming a leaf to another star cluster. By Lemma 16, the social welfare of any set of disjoint star clusters spanning a given set of nodes does not depend

on how these stars are actually formed, and, hence, any such deviation does not reduce the social welfare. Observe that such a deviating move of player i from the star $C^*(i)$ to another star C , requires that $|C^*(i)| > |C| + 1$. Therefore, by considering the lexicographic order π of all star clusters in the partition \mathcal{C}^* based on the number of nodes, from the minimum to the maximum, we observe that any deviating move that leads to a partition \mathcal{C}' satisfies $\pi(\mathcal{C}^*) < \pi(\mathcal{C}')$ and, hence, this process is guaranteed to end as we can apply this reasoning to any subsequent deviation. This concludes the proof of the theorem. \square

6. Conclusions

We have presented new bounds on the price of stability of some graph-based hedonic games such as simple symmetric fractional hedonic games and social distance games. In particular, we show improved lower bounds for the general case as well as positive results for special cases and variants.

Clearly, the most important open question concerns the upper bound for general graphs, as no non-trivial bound is known. We conjecture that the price of stability is constant in both cases but the proof of such a claim remains elusive. Interestingly, for social distance games no better bounds are known even for the restricted case of bipartite graphs; we remark that for the class of simple symmetric fractional hedonic games, the price of stability for bipartite graphs is a very small constant [10].

References

- [1] Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, E., Wexler, T., and Roughgarden, T.: The price of stability for network design with fair cost allocation. *SIAM Journal on Computing* 38(4): 1602–1623, 2008.
- [2] Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., and Peters, D.: Fractional hedonic games. *ACM Transactions on Economics and Computation* 7(2): 6:1–6:29, 2019.
- [3] Aziz, H., Brandt, F., and Seedig, H.: Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence* 195: 316–334, 2013.
- [4] Aziz, H., Gaspers, S., Gudmundsson, J., Mestre, J., and Täubig, H.: Welfare maximization in fractional hedonic games. In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 461–467, 2015.
- [5] Balliu, A., Flammini, M., Melideo, G. and Olivetti, D.: Nash stability in social distance games. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 342–348, 2017.

- [6] Balliu, A., Flammini, M., and Olivetti, D.: On Pareto optimality in social distance games. In Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI), pages 349–355, 2017.
- [7] Barrot, N., Ota, K., Sakurai, Y., and Yokoo, M.: Unknown agents in friends oriented hedonic games: Stability and complexity. In Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI), pages 1756–1763, 2019.
- [8] Barrot, N., and Yokoo, M.: Stable and envy-free partitions in hedonic games. In Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI), pages 67–73, 2019.
- [9] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, and L. Moscardelli. On the price of stability of fractional hedonic games. In Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pages 1239–1247, 2015.
- [10] Bilò, V., Fanelli, A., Flammini, M., Monaco, G., and Moscardelli, L.: Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. *Journal of Artificial Intelligence Research*, 62: 315–371, 2018.
- [11] Bogomolnaia, A., and Jackson, M.O.: The stability of hedonic coalition structures. *Games and Economic Behavior* 38(2): 201–230, 2002.
- [12] Brandl, F., Brandt, F., and Strobel, M. Fractional hedonic games: Individual and group stability. In Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), pages 1219–1227, 2015.
- [13] Brânzei, S., and Larson, K.: Social distance games. In Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI), pages 91–96, 2011.
- [14] Drèze, J.H., and Greenberg, J.: Hedonic coalitions: optimality and stability. *Econometrica*, 48(4): 987–1003, 1980.
- [15] Dutton, R. D., and Brigham, R. C.: Edges in graphs with large girth. *Graphs and Combinatorics*, 7(4):315–321, 1991.
- [16] Elkind, E., Fanelli, A., and Flammini, M.: Price of Pareto optimality in hedonic games. In Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI), pages 475–481, 2016.
- [17] Feldman, M, Lewin-Eytan, L., and Naor, J.: Hedonic clustering games. *ACM Transactions on Parallel Computing*, 2(1), Article 4, 2015.
- [18] Gairing, M., and Savani, S.: Computing stable outcomes in hedonic games. *Mathematics of Operations Research* 44(3): 1101–1121, 2019.

- [19] Hoefer, M., and Jiamjitrak, W.: On proportional allocation in hedonic games. In Proceedings of the 10th International Symposium on Algorithmic Game Theory (SAGT), pages 307–319, 2017.
- [20] Hoefer, M., Vaz, D., and Wagner, L.: Hedonic coalition formation in networks. In Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI), pages 929–935, 2015.
- [21] Kaklamanis, C., Kanellopoulos, P., and Papaioannou, K.: The price of stability of simple symmetric fractional hedonic games. In Proceedings of the 9th Symposium on Algorithmic Game Theory (SAGT), pages 220–232, 2016.
- [22] Kaklamanis, C., Kanellopoulos, P., and Patouchas, D.: On the price of stability of social distance games. In Proceedings of the 11th Symposium on Algorithmic Game Theory (SAGT), pages 125–136, 2018.
- [23] Monaco, G., Moscardelli, L., and Velaj, Y.: Stable outcomes in modified fractional hedonic games. In Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), pages 937–945, 2018.
- [24] Monaco, G., Moscardelli, L., and Velaj, Y.: On the performance of stable outcomes in modified fractional hedonic games with egalitarian social welfare. In Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), pages 873–881, 2019.
- [25] Olsen, M.: Nash stability in additively separable hedonic games and community structures. *Theory of Computing Systems* 45(4): 917–925, 2009.
- [26] Olsen, M.: On defining and computing communities. In Proceedings of the 18th Computing: Australasian Theory Symposium (CATS), pages 97–102, 2012.
- [27] Peters, D.: Complexity of hedonic games with dichotomous preferences. In Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI), pages 579–585, 2016.
- [28] Peters, D., and Elkind, E.: Simple causes of complexity in hedonic games. In Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), pages 617–623, 2015.
- [29] Saad, W., Han, Z., Basar, T., Debbah, M., Hjørungnes, A.: Hedonic coalition formation for distributed task allocation among wireless agents. *IEEE Transactions on Mobile Computing* 10(9): 1327–1344, 2011.