On the Price of Stability of Fractional Hedonic Games

Vittorio Bilò
Univ. of Salento, Italy
vittorio.bilo@unisalento.it

Angelo Fanelli
CNRS, (UMR-6211), France
angelo.fanelli@unicaen.fr

Michele Flammini
Univ. of L'Aquila, Italy
GSSI Institute, L'Aquila, Italy
michele.flammini@univaq.it

Gianpiero Monaco
Univ. of L'Aquila, Italy
gianpiero.monaco@univaq.it

Luca Moscardelli
Univ. of Chieti-Pescara, Italy
luca.moscardelli@unich.it

ABSTRACT

We consider fractional hedonic games, where self-organized groups (or clusters) are created as a result of the strategic interactions of independent and selfish players and the happiness of each player in a group is the average value she ascribes to its members. We adopt Nash stable outcomes, that is states where no player can improve her utility by unilaterally changing her own group, as the target solution concept. We study the quality of the best Nash stable outcome and refer to the ratio of its social welfare to the one of an optimal clustering as to the price of stability. We remark that a best Nash stable outcome has a natural meaning of stability since it is the optimal solution among the ones which can be accepted by selfish users. We provide upper and lower bounds on the price of stability for games played on different network topologies. In particular, we give an almost tight bound (up to a 0.026 additive factor) for bipartite graphs and suitable bounds on more general families of graphs.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems; J.4 [**Computer Applications**]: Social and Behavioral Sciences - Economics.

General Terms

Economics, Theory and Algorithms

Keywords

Algorithmic game theory; Clustering formation; Nash stability; Fractional hedonic games.

1. INTRODUCTION

In many economical, social and political situations, individuals carry out activities in groups rather than alone and on their own. In these scenarios, understanding the "happiness" of each member of the group becomes of crucial importance. As examples, 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 trait and on the identity of its members.

Hedonic games, introduced in [12], describe the dependence of a player's utility (or payoff) on the identity of the members of her group. They are games in which players have preferences over the set of all possible player partitions (called clusterings). In particular, the utility of each player only depends on the composition or structure of the cluster she belongs to. In the literature, a significant stream of research considered this topic from a strategic cooperative point of view [8, 9, 13, 15], with the purpose of characterizing the existence and the properties of coalitional structures such as, for instance, the core. Nevertheless, studying strategic solutions under a non-cooperative scenario (such as, for instance, Nash equilibria) becomes of fundamental importance when considering huge environments (like the Internet) lacking a social planner or where the cost of coordination is tremendously high. Examples of non-cooperative studies on hedonic games, in which self-organized clusterings are obtained from the decisions taken by independent and selfish players, can be found in [7, 14, 15].

In this work, we consider the class of (symmetric) *fractional hedonic games* introduced in [1]. These games are defined by a graph in which nodes represent players and the weight of each edge measures the happiness of its two incident players when belonging to the same cluster. The utility that player i gets when belonging to cluster C is given by the total weight of edges which are incident to i and to some other player belonging to C (the total happiness of i in C) divided by the cardinality of C, i.e., the number of its nodes. The social welfare of a clustering is the sum of the players' utilities.

Fractional hedonic games model natural behavioral dynamics in social environments that are not captured by additive separable ones, that is, games in which the utility of a player is simply defined as her total happiness. In particular, fractional hedonic games defined on ***undirected and unweighted bipartite graphs*** suitably model a basic economic scenario in which each player can be considered as a buyer or a seller. There are only edges connecting buyers and sellers and every player sees a player of the same type as a market competitor. In a situation of free movement, each player prefers to be situated in a group (market) with a small number of competitors: Each buyer wants to be situ-

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

ated in a group with many sellers and few other buyers, thus maximising their ratio, in order to decrease the price of the good. On the other hand, a seller wants to be situated in a group maximizing the number of buyers against the number of sellers, in order to be able to increase the price of the good and gain a higher profit. This senario is referred to in [1] as *Bakers and Millers* and can be generalized to situations in which there are more than two types of players by means of k-partite graphs.

In this setting, a clustering is *Nash stable* (or it is a *Nash equilibrium*) if no player can improve her utility by unilaterally changing her own cluster. Our aim is to understand the performance of Nash stable clusterings. In particular, we study the quality of a best Nash stable outcome and refer to the ratio of its social welfare to the one of the socially optimal clustering as to the *price of stability* (a study on the price of stability for multi-agent systems can be found in [16]). A best Nash stable outcome has a natural meaning of stability, since it is an optimal solution among the ones which can be accepted by selfish players [3]. Moreover, in many networking applications and multi-agent systems, agents are never completely unrestricted; rather, they interact with an underlying protocol that essentially proposes a collective solution to all participants, each of whom can either accept it or defect from it. As a result, it is in the interest of the protocol designer to seek for a best solution at equilibrium. In fact, this can naturally be viewed as the optimum subject to the constraint that the solution has to be stable, with no agent having an incentive to unilaterally defect from it once it is offered.

Related Work. Hedonic games have been first formalized by Dréze and Greenberg [12] who analyzed them under a cooperative perspective. Properties guaranteeing the existence of core allocations (a core is a clustering structure in which no group of players has an incentive to form a different cluster) for games with additively separable utility have been studied by Banerjee, Konishi and Sönmez [9], while Bogomolnaia and Jackson [8] consider several forms of clustering stability like the core, the Nash and the individual stability. Ballester [5], Aziz, Brandt and Seedig [2] and Olsen [17] deal with computational complexity issues related to hedonic games, also considering additively separable utilities.

Bloch and Diamantoudi [7] study non-cooperative games of coalition formation and identify conditions for stable outcomes. In a similar way, Apt and Witzel [4] study how certain proposed rules can transform clusterings into other ones with specific properties. Feldman, Lewin-Eytan and Naor [14] investigate some interesting subclasses of hedonic games from a non-cooperative point of view, by characterizing Nash equilibria and providing upper and lower bounds on both the price of stability and the price of anarchy (that is the ratio between the quality of the worst Nash stable clustering and the socially optimal clustering).

Finally, hedonic games have also been considered by Charikar, Guruswami and Wirth [10] and by Demaine et al. [11] from a classical optimization point of view (i.e, where solutions are not necessarily stable).

Fractional hedonic games have been recently introduced by Aziz, Brandt and Harrenstein [1] again from the cooperative perspective. They prove that the core can be empty for games played on general graphs and that it is not empty for games played on some classes of undirected and unweighted graphs (that is, graphs with degree at most 2, multipartite complete graphs, bipartite graphs admitting a perfect matching and regular bipartite graphs). Olsen [18] investigates computational issues and the existence of Nash stable outcomes in the variant of fractional hedonic games in which the utility of player i in cluster C is defined as the total happiness of i in C divided by $|C|-1$, that is, the variant in which the player herself is not accounted to the population of the clustering. Although the difference between the two utility functions might seem "almost" negligible, the sets of Nash stable outcomes they induce in games played on a same graph are usually quite different.

Bilò et al. [6] consider Nash stable solutions in fractional hedonic games. They first show that, in presence of negative edge weights, these outcomes are not guaranteed to exist, while, if all edge weights are non-negative, the basic outcome in which all players belong to the same cluster is Nash stable. They provide an upper bound of $O(n)$ on the price of anarchy for games played on weighted graphs and show that it is asymptotically tight even for games played on unweighted paths; moreover, they also show a lower bound of $\Omega(n)$ on the price of stability for games played on weighted stars. As a consequence of these results, they pose the characterization of the price of stability for games played on undirected and unweighted graphs as an interesting and challenging open problem. As a partial answer to this question, they show that, for undirected and unweighted bipartite graphs, the price of stability is at least 1.002 and at most 2 and that, for undirected and unweighted trees, the price of stability is 1. Both upper bounds are obtained via polynomial time algorithms constructing Nash stable clusterings of optimal or almost optimal social value.

Our Contribution. In this paper, we focus on the price of stability of fractional hedonic games played on undirected and unweighted graphs. For general graph topologies, we give a lower bound of 2. Moreover, we provide an upper bound of 4 which holds under the assumption that the game possesses a 2-Strong Nash stable clustering, that is, a clustering such that no pair of players can improve their utility by simultaneously changing her own cluster. However, we show that there are games for which such a condition is not always guaranteed. We then focus on games played on specific graph topologies. In particular, for triangle-free graphs, we prove an upper bound of 4, while, for bipartite graphs, we give an upper bound of $6(3-2\sqrt{2}) \approx 1.0294$ and a lower bound of 1.003. We stress that our upper bounds on the price of stability directly extend also to the utility function considered by Olsen in [18].

Paper Organization. The paper is organized as follows. In Section 2, we formally define fractional hedonic games and give some preliminary results. The technical contributions of the paper are then presented in Sections 3, 4 and 5, which address, respectively, games played on general, triangle-free and bipartite graphs. Finally, some interesting open problems are stated in Section 6. Due to space constraints, some proofs are omitted.

2. DEFINITIONS AND NOTATION

For an integer $n > 0$, we denote with $[n]$ the set $\{1, \ldots, n\}$. For any set S, we refer to the number of elements in S as the size of S, and we denote it with $|S|$.

Let $G = (V, E)$ be a connected undirected graph with $n = |V|$. Given a subset of nodes $S \subseteq V$, $G_S = (S, E_S)$ is the subgraph of G induced by the set S, i.e., $E_S = \{\{u, v)\} \in$

$E : u,v \in S\}$. $N_u(S)$ denotes the neighbors of u in S, i.e., $N_u(S) = \{v \in S : \{u,v\} \in E\}$ and $E_u(S)$ the edges in E_S being incident to u, i.e., $E_u(S) = \{\{u,v\} \in E : \{u,v\} \in E_S\}$. A vertex cover of G is any subset of nodes $C \subseteq V$ such that each edge in E is incident to at least a node in C. A minimum vertex cover is a vertex cover of minimum size. An independent set of G is any subset of nodes $I \subseteq V$ such that, for every pair of nodes $u, v \in V$, there is no edge in E connecting them. It is obvious that, if C is a vertex cover of G then $V \setminus C$ is an independent set of G.

The *fractional hedonic game* induced by G, denoted as $\mathcal{G}(G)$, is the non-cooperative strategic game in which each node $u \in V$ is associated with a selfish player (or agent) and each player chooses to join a certain *cluster* (assuming that candidate clusters are numbered from 1 to n). Hence, a state of the game, that we will call in the sequel a *clustering*, is a partition of the agents into n clusters $\mathbf{C} = \{C_1, C_2, \dots, C_n\}$ such that $C_j \subseteq V$ for each $j \in [n]$, $\bigcup_{j\in[n]} C_j = V$ and $C_i \cap C_j = \emptyset$ for any $i, j \in [n]$ with $i \neq j$. Notice that some clusters may be empty. We denote by $C(u)$ the cluster in $\mathbf{C}$ chosen by agent u and by $\gamma(\mathbf{C}, u)$ the index of such a cluster, i.e., $C(u) = C_{\gamma(\mathbf{C},u)}$. In a clustering $\mathbf{C}$, the payoff (or utility) of agent u is defined as

$$p_u(\mathbf{C}) = \frac{|E_u(C(u))|}{|C(u)|}.$$

Each agent chooses the cluster she belongs to with the aim of maximizing her payoff. We denote by $(\mathbf{C}, u, j)$, the new clustering obtained from $\mathbf{C}$ by moving agent u from $C(u)$ to C_j; formally, $(\mathbf{C}, u, j) = \mathbf{C} \setminus \{\mathbf{C}(u), C_j\} \cup \{\mathbf{C}(u) \setminus \{u\}, C_j \cup \{u\}\}$. We say that an agent *deviates* if she changes the cluster she belongs to. Given a clustering $\mathbf{C}$, an *improving move* (or simply a *move*) for player u is a deviation to any cluster C_j that strictly increases her payoff, i.e., $p_u((\mathbf{C}, u, j)) > p_u(\mathbf{C})$. We say that an agent is *stable* if she cannot perform a move; a clustering is *Nash stable* (or is a *Nash equilibrium*) if every agent is stable. More generally, a clustering $\mathbf{C}$ is *k-Strong Nash stable* (or is a k-Strong Nash equilibrium) if, for each $\mathbf{C}'$ obtained from $\mathbf{C}$ when at most k players jointly change their strategies, it holds that $p_u(\mathbf{C}) \geq p_u(\mathbf{C}')$ for some u belonging to the set of deviating players, that is, there always exists a player not improving her utility after the joint collective deviation. By definition, a 1-Strong Nash stable clustering is a Nash stable clustering and, for each $k > 1$, each k-Strong Nash stable clustering is also a $(k-1)$-Strong Nash stable clustering. We denote with $\mathsf{NSC}_k(\mathcal{G}(G))$ the set of k-Strong Nash stable clusterings of $\mathcal{G}(G)$.

The *social welfare* of a clustering $\mathbf{C}$ is the summation of the players' payoffs, i.e., $\mathsf{SW}(\mathbf{C}) = \sum_{u\in V} p_u(\mathbf{C})$. We overload the social welfare function by applying it also to single clusters to obtain their contribution to the social welfare, i.e., $\mathsf{SW}(C_i) = \sum_{u\in C_i} p_u(\mathbf{C})$ so that $\mathsf{SW}(\mathbf{C}) = \sum_{i\in[n]} \mathsf{SW}(C_i)$. Notice that the following property, providing a simple formula to compute the social welfare of a clustering, holds:

Property 1. *Given any cluster C, $\mathsf{SW}(C) = \frac{2|E_C|}{|C|}$.*

In particular the above property states that the contribution of each cluster to the social value is given by twice the number of intra-cluster edges divided by the cardinality of the cluster. Moreover, if G_{C_i} is a (non-empty) tree, we have $\mathsf{SW}(C_i) = \frac{2(|C_i|-1)}{|C_i|}$.

Given a game $\mathcal{G}(G)$, an *optimal* clustering $\mathbf{C}^*$ is one that maximizes the social welfare of $\mathcal{G}(G)$. A clustering $\mathbf{C}$ is *feasible* if G_{C_i} is connected, for every $i \in [n]$. Notice that an optimal configuration is always feasible. For any integer $k \geq 1$, the *k-strong price of anarchy* of a fractional hedonic game $\mathcal{G}(G)$ is defined as the worst-case ratio between the social optimum and the social welfare of a k-Strong Nash stable clustering (whenever the latter exists): $\mathsf{PoA}_k(\mathcal{G}(G)) = \max_{\mathbf{C}\in\mathsf{NSC}_k(\mathcal{G}(G))} \frac{\mathsf{SW}(\mathbf{C}^*)}{\mathsf{SW}(\mathbf{C})}$ if $\mathsf{NSC}_k(\mathcal{G}(G)) \neq \emptyset$, $\mathsf{PoA}_k(\mathcal{G}(G)) = \infty$ otherwise. Analogously, the *k-strong price of stability* of $\mathcal{G}(G)$ is defined as the best-case ratio between the social optimum and the social welfare of a k-Strong Nash stable clustering (whenever the latter exists): $\mathsf{PoS}_k(\mathcal{G}(G)) = \min_{\mathbf{C}\in\mathsf{NSC}_k(\mathcal{G}(G))} \frac{\mathsf{SW}(\mathbf{C}^*)}{\mathsf{SW}(\mathbf{C})}$ if $\mathsf{NSC}_k(\mathcal{G}(G)) \neq \emptyset$, $\mathsf{PoS}_k(\mathcal{G}(G)) = \infty$ otherwise. We simply use the terms price of anarchy and price of stability, and remove the subscript k from the notation, when $k = 1$. Note that, since for each game $\mathcal{G}(G)$ and index $k \geq 1$ it holds that $\mathsf{NSC}_k(\mathcal{G}(G)) \subseteq \mathsf{NSC}_1(\mathcal{G}(G))$, we can claim the following observation.

Observation 1. *For any graph G and index $k \geq 1$, $\mathsf{PoS}(\mathcal{G}(G)) \leq \mathsf{PoA}_k(\mathcal{G}(G))$.*

We conclude the section by showing a useful lemma relating the social welfare of an optimal solution of $\mathcal{G}(G)$ to the cardinality of a min vertex cover of G. To this aim, given $G = (V, E)$, let VC be a minimum vertex cover of G and $\overline{VC} = V \setminus VC$.

Lemma 1. *For any game $\mathcal{G}(G)$, $\frac{\mathsf{SW}(\mathbf{C}^*)}{|VC|} < 2$.*

Proof. Let C_i^* be an non-empty cluster of an optimal clustering $\mathbf{C}^*$. We partition the nodes of C_i^* in two sets $X_i^{VC} = C_i^* \cap VC$ and $X_i^{\overline{VC}} = C_i^* \cap \overline{VC}$. We distinguish between two cases:

i) $X_i^{VC} = \emptyset$; it follows that $C_i^* \subseteq \overline{VC}$. Therefore, since $\overline{VC}$ is an independent set, it follows that $\mathsf{SW}(C_i^*) = 0$.

ii) $X_i^{VC} \neq \emptyset$; in this case the total number of edges in C_i^* is at most $|X_i^{VC}| \cdot |X_i^{\overline{VC}}| + \frac{1}{2}|X_i^{VC}|^2$.

By Property 1 it follows that the contribution to the optimal social welfare of cluster C_i^* is such that

$$\begin{aligned}\mathsf{SW}(C_i^*) &\leq 2\frac{|X_i^{VC}| \cdot |X_i^{\overline{VC}}| + \frac{1}{2}|X_i^{VC}|^2}{|X_i^{VC}| + |X_i^{\overline{VC}}|} \\ &= 2|X_i^{VC}| \frac{|X_i^{\overline{VC}}| + \frac{1}{2}|X_i^{VC}|}{|X_i^{VC}| + |X_i^{\overline{VC}}|}.\end{aligned}$$

Dividing by $|X_i^{VC}|$ we obtain $\frac{\mathsf{SW}(C_i^*)}{|X_i^{VC}|} \leq 2\frac{|X_i^{\overline{VC}}| + \frac{1}{2}|X_i^{VC}|}{|X_i^{VC}| + |X_i^{\overline{VC}}|} < 2$.

By summing over all the non-empty clusters C_i^* of an optimal clustering $\mathbf{C}^*$ the theorem follows. □

3. GENERAL GRAPHS

In this section, we consider games played on unrestricted graph topologies. We start by giving a lower bound of 2 on the price of stability.

Theorem 1. *For any $\epsilon > 0$, there exists a graph G_ϵ such that $\mathsf{PoS}(\mathcal{G}(G_\epsilon)) > 2 - \epsilon$.*

Proof. For any positive integer h, define the graph $G_h = (V_h, E_h)$ as follows: $V_h = X_h \cup Y_h$, with $X_h =$

$\{x_1,\ldots,x_{h+2}\}$, $Y_h = \{y_1,\ldots,y_h\}$, and $E_h = \{\{x_i,x_j\} : i,j \in [h+2], i \neq j\} \cup \{\{x_i,y_i\} : i \in [h]\}$. Intuitively, G_h has $2h+2$ nodes, where the $h+2$ nodes in X_h form a clique and each of the h nodes in Y_h is a *leaf node*. For each $i \in [h]$, x_i is the *partner node* of leaf node y_i. Finally, nodes x_{h+1} and x_{h+2} are the *special nodes*.

Our aim is to show that, for each value of h, $\mathsf{NSC}(\mathcal{G}(G_h)) = \{\widehat{\mathbf{C}}\}$, where $\widehat{\mathbf{C}}$ denotes a clustering in which all players belong to the same cluster. We prove this claim by showing a sequence of properties that have to be satisfied by any Nash stable clustering of $\mathcal{G}(G_h)$. The first of these properties is quite intuitive and states that, in any Nash stable clustering of $\mathcal{G}(G_h)$, each leaf node has to be in the same cluster of its partner node.

Property 2. *For any $\mathbf{C} \in \mathsf{NSC}(\mathcal{G}(G_h))$ and $i \in [h]$, $C(x_i) = C(y_i)$.*

We continue by showing that, in any Nash stable clustering of $\mathcal{G}(G_h)$, the two special nodes have to be in the same cluster.

Property 3. *For any $\mathbf{C} \in \mathsf{NSC}(\mathcal{G}(G_h))$, $C(x_{h+1}) = C(x_{h+2})$.*

Property 3 can be proved as follows. Assume, for the sake of contradiction, that there exists a Nash stable clustering $\mathbf{C}$ such that $C(x_{h+1}) \neq C(x_{h+2})$. By Properties 1 and 2, it follows that $p_{x_{h+1}}(\mathbf{C}) = \frac{|C(x_{h+1})|-1}{2|C(x_{h+1})|} < \frac{1}{2}$. Let $\mathbf{C}'$ be the clustering obtained from $\mathbf{C}$ when node x_{h+1} deviates to cluster $C(x_{h+2})$. Again, by Properties 1 and 2, it follows that $p_{x_{h+1}}(\mathbf{C}') = \frac{|C(x_{h+2})|+1}{2(|C(x_{h+2})|+1)} = \frac{1}{2}$, thus contradicting the hypothesis that $\mathbf{C}$ is Nash stable.

We can now proceed to show that $\widehat{\mathbf{C}}$ is the unique Nash stable clustering of $\mathcal{G}(G_h)$. Assume, for the sake of contradiction, that there exists another Nash stable clustering $\mathbf{C} \neq \widehat{\mathbf{C}}$ for $\mathcal{G}(G_h)$. By Properties 2 and 3, there must exist an index $i \in [h]$, such that $C(x_i) \neq C(x_{h+1})$. By Properties 1 and 2 and by the fact that $C(x_i)$ does not contain any special node, it follows that $p_{x_i}(\mathbf{C}) = \frac{|C(x_i)|}{2|C(x_i)|} = \frac{1}{2}$. Let $\mathbf{C}'$ be the clustering obtained from $\mathbf{C}$ when node x_i deviates to cluster $C(x_{h+1})$. Again, by Properties 1 and 2 and by the fact that $C(x_{h+1})$ contains both special nodes (because of Property 3), it follows that $p_{x_i}(\mathbf{C}') = \frac{|C(x_{h+1})|+2}{2(|C(x_{h+1})|+1)} > \frac{1}{2}$, thus contradicting the hypothesis that $\mathbf{C}$ is Nash stable. Hence, $\widehat{\mathbf{C}}$ is the unique Nash stable clustering of $\mathcal{G}(G_h)$

Clearly, $\mathsf{SW}(\widehat{\mathbf{C}}) = \frac{(h+1)(h+2)+2h}{2(h+1)} < \frac{h+4}{2}$. Moreover, $\mathsf{SW}^*(\mathcal{G}(G_h)) \geq h+1$, since G_h admits a perfect matching. Hence, we obtain $\mathsf{PoS}(\mathcal{G}(G_h)) \geq 2 - \Theta(1/h)$ for any positive integer h. By taking the limit for h going to infinity, the claim follows. □

Determining upper bounds on the price of stability in this setting is a quite challenging task. Anyway, we can show a constant upper bound on the price of stability for all those games admitting 2-Strong Nash stable clusterings.

Theorem 2. *For any fractional hedonic game $\mathcal{G}(G)$ such that $\mathsf{NSC}_2(\mathcal{G}(G)) \neq \emptyset$, $\mathsf{PoS}(\mathcal{G}(G)) \leq 4$.*

Proof. We show that, under the hypothesis of the theorem, $\mathsf{PoA}_2(\mathcal{G}(G)) \leq 4$ which, by Observation 1, yields the claim. To this aim, fix a 2-Strong Nash stable clustering $\mathbf{C}$ and let $V^- = \{u \in V : p_u(\mathbf{C}) < \frac{1}{2}\}$ be the set of agents getting a payoff strictly smaller than $\frac{1}{2}$ in $\mathbf{C}$ and $V^+ = V \setminus V^-$. We show that V^- is an independent set of G. Assume, by way of contradiction, that there exists an edge $\{u,v\} \in E$ such that $u,v \in V^-$. In this case, u and v can jointly deviate to a new cluster and obtain both a payoff of $\frac{1}{2}$, thus contradicting the fact that $\mathbf{C} \in \mathsf{NSC}_2(\mathcal{G}(G))$. Hence, we get that V^+ is a vertex cover of G. By using Lemma 1, we obtain $\mathsf{SW}(\mathbf{C}^*) < 2|V^+|$ which, together with $\mathsf{SW}(\mathbf{C}) \geq \frac{|V^+|}{2}$, yields the claim. □

However, 2-Strong Nash stable clusterings are not always guaranteed to exist, as stated by the following theorem whose proof is omitted due to lack of space.

Theorem 3. *There exists a graph G such that $\mathsf{NSC}_2(\mathcal{G}(G)) = \emptyset$.*

Because of these restrictive results, in the following two subsections, we concentrate on games played on some specific graph topologies.

4. TRIANGLE-FREE GRAPHS

In this section, we focus on games played on triangle-free graphs and provide an upper bound of 4 on the price of stability.

For an integer $k \geq 2$, a star graph (from now on, simply, a *star*) of order k is a tree with k nodes and $k-1$ leaves. Given a star S, we denote with $\ell(S)$ the set of its leaves and with $c(S)$ its center, that is, its unique non-leaf node; so that the order of S is equal to $|\ell(S)|+1$. A *star clustering* $\mathbf{C} = (C_1,\ldots,C_n)$ for $\mathcal{G}(G)$ is a clustering such that C_i is a star for each $i \in [n]$. An *optimal star clustering* for $\mathcal{G}(G)$ is a star clustering for $\mathcal{G}(G)$ of maximum social welfare.

We give an upper bound on the price of stability by showing that an optimal star clustering for $\mathcal{G}(G)$ is Nash stable whenever G is triangle-free and then evaluating its performance with respect to the social optimum $\mathbf{C}^*$. Next lemma establishes a fundamental property possessed by an optimal star clustering.

Lemma 2. *Let $\mathbf{C}$ be a star clustering for $\mathcal{G}(G)$. For any node $u \in V$ and $k \in [n]$ such that $|C(u)| > |C_k| + 1$ and $(\mathbf{C},u,k)$ is a star clustering for G, it holds that $\mathsf{SW}(\mathbf{C},u,k) > \mathsf{SW}(\mathbf{C})$.*

Proof. Let $\mathbf{C}$, u and k be defined as in the claim. We have

$$\begin{aligned}
&\mathsf{SW}(\mathbf{C},u,k) - \mathsf{SW}(\mathbf{C})\\
&= \frac{2(|C(u)|-2)}{|C(u)|-1} + \frac{2|C_k|}{|C_k|+1} - \frac{2(|C(u)|-1)}{|C(u)|} - \frac{2(|C_k|-1)}{|C_k|}\\
&= 2\left(\frac{|C_k|}{|C_k|+1} - \frac{|C_k|-1}{|C_k|} - \frac{|C(u)|-1}{|C(u)|} + \frac{|C(u)|-2}{|C(u)|-1}\right)\\
&= \frac{2}{|C_k|(|C_k|+1)} - \frac{2}{|C(u)|(|C(u)|-1)}
\end{aligned}$$

and the last quantity is strictly positive whenever

$$|C_k|(|C_k|+1) < |C(u)|(|C(u)|-1). \tag{1}$$

Since, inequality (1) is implied by the assumption $|C(u)| > |C_k| + 1$, the claim follows. □

We can now proceed to show that an optimal star clustering for $\mathcal{G}(G)$ is Nash stable.

LEMMA 3. *Let G be a triangle-free graph, then any optimal star clustering for $\mathcal{G}(G)$ is Nash stable.*

PROOF. Fix a triangle-free graph G and let $\mathbf{C}$ be an optimal star clustering for $\mathcal{G}(G)$. Assume, for the sake of contradiction, that there exists a node $u \in V$ who can perform an improving deviation by migrating from $C(u)$ to C_i, that is, such that $p_u(\mathbf{C}) < p_u(\mathbf{C}, u, i)$.

Assume first that $u = c(C(u))$, so that $p_u(\mathbf{C}) = \frac{|C(u)|-1}{|C(u)|}$. Two cases may occur:
1) $\{u, c(C_i)\} \in E$. Since G is triangle-free, $E_u(C_i \cup \{u\}) = \{u, c(C_i)\}$ which implies

$$p_u(\mathbf{C}, u, i) = \frac{1}{|C_i|+1} < \frac{|C(u)|-1}{|C(u)|} = p_u(\mathbf{C}),$$

where the inequality follows from $|C_i|, |C(u)| \geq 2$. This contradicts the assumption that u can perform an improving deviation by migrating to C_i.
2) $\{u, c(C_i)\} \notin E$. Let $1 \leq j \leq |C_i| - 1$ be the number of nodes in C_i which are adjacent to u in G, so that $p_u(\mathbf{C}, u, i) = \frac{j}{|C_i|+1}$. Note that it cannot be $j = 0$, because otherwise, since $\{u, c(C_i)\} \notin E$, it would follow $p_u(\mathbf{C}, u, i) = 0$, thus immediately contradicting $p_u(\mathbf{C}) < p_u(\mathbf{C}, u, i)$. If $|C(u)| \geq |C_i|$, then

$$\begin{aligned} p_u(\mathbf{C}, u, i) &= \frac{j}{|C_i|+1} \leq \frac{|C_i|-1}{|C_i|+1} \\ &\leq \frac{|C(u)|-1}{|C(u)|+1} < \frac{|C(u)|-1}{|C(u)|} \\ &= p_u(\mathbf{C}), \end{aligned}$$

while, if $|C(u)| = |C_i| - 1$, then

$$\begin{aligned} p_u(\mathbf{C}, u, i) &= \frac{j}{|C_i|+1} \leq \frac{|C_i|-1}{|C_i|+1} \\ &= \frac{|C(u)|}{|C(u)|+2} \leq \frac{|C(u)|-1}{|C(u)|} \\ &= p_u(\mathbf{C}), \end{aligned}$$

where the last inequality follows from $|C(u)| \geq 2$. In both cases, we obtain a contradiction to the assumption that u can perform an improving deviation by migrating to C_i. For the leftover case of $|C(u)| < |C_j| - 1$, let z be a node of C_i which is adjacent to u (such a node always exists since $j \geq 1$). We claim that $(\mathbf{C}, z, \gamma(\mathbf{C}, u))$ is a star clustering for $\mathcal{G}(G)$. In fact, since C_i is a star and $z \neq c(C_i)$, then $C_i \setminus \{z\}$ is also a star. Moreover, since $C(u)$ is a star, then $C(u) \cup \{z\}$ cannot be a star only if there exists an edge (z, t), besides edge (u, z), for some $t \in C(u)$. But, if such an edge existed, then, the existence of edge (u, t) (remember $u = c(C(u))$) would contradict the hypothesis that G is triangle-free. By applying Lemma 2 with $x = z$ and $k = \gamma(\mathbf{C}, u)$, we obtain a contradiction to the fact that $\mathbf{C}$ is an optimal star clustering for $\mathcal{G}(G)$.

Assume now that $u \in \ell(C(u))$, so that $p_u(\mathbf{C}) = \frac{1}{|C(u)|}$. Again, the same two cases may occur:
1) $\{u, c(C_i)\} \in E$. In the case in which $|C(u)| \leq |C_i| + 1$, since G is triangle-free, it follows that $E_u(C_i \cup \{u\}) = \{u, c(C_i)\}$ which implies

$$p_u(\mathbf{C}, u, i) = \frac{1}{|C_i|+1} \leq \frac{1}{|C(u)|} = p_u(\mathbf{C}).$$

This contradicts the assumption that u can perform an improving deviation by migrating to C_i. In the leftover case in which $|C(u)| > |C_i|+1$, we claim that $(\mathbf{C}, u, i)$ is a star clustering for $\mathcal{G}(G)$. In fact, since $C(u)$ is a star, then $C(u) \setminus \{u\}$ is also a star. Moreover, since C_i is a star, then $C_i \cup \{u\}$ cannot be a star only if there exists an edge (u, z), besides edge $\{u, c(C_i)\}$, for some $z \in C_i$. But, if such an edge existed, then, the existence of edge $\{c(C_i), z\}$ would contradict the hypothesis that G is triangle-free. By applying Lemma 2 with $x = u$ and $k = i$, we obtain a contradiction to the fact that $\mathbf{C}$ is an optimal clustering for $\mathcal{G}(G)$.
2) $\{u, c(C_i)\} \notin E$. Let $1 \leq j \leq |C_i| - 1$ be the number of nodes of C_i which are adjacent to u in G. Again, it cannot be $j = 0$, because otherwise, since $\{u, c(C_i)\} \notin E$, it would follow $p_u(\mathbf{C}, u, i) = 0$, thus immediately contradicting $p_u(\mathbf{C}) < p_u(\mathbf{C}, u, i)$. Denote with z any node of C_i which is adjacent to u.
In the case in which $|C(u)| = 2$ and $|C_i| \leq 3$, we get

$$p_u(\mathbf{C}, u, i) = \frac{j}{|C_i|+1} \leq \frac{|C_i|-1}{|C_i|+1} \leq \frac{1}{2} = p_u(\mathbf{C}),$$

which contradicts the assumption that u can perform an improving deviation by migrating to C_i.
In the case in which $|C(u)| = 2$ and $|C_i| \geq 4$, then it follows that $(\mathbf{C}, z, \gamma(\mathbf{C}, u))$ is a star clustering for $\mathcal{G}(G)$. By applying Lemma 2 with $x = z$ and $k = \gamma(\mathbf{C}, u)$, we obtain a contradiction to the fact that $\mathbf{C}$ is an optimal star clustering for $\mathcal{G}(G)$.
In the case in which $|C(u)| = 3$ and $|C_i| = 2$, then $p_u(\mathbf{C}, u, i) = \frac{1}{3} = p_u(\mathbf{C})$ contradicts the assumption that u can perform an improving deviation by migrating to C_i.
In the case in which $|C(u)| \geq 4$ and $|C_i| = 2$, then it follows that $(\mathbf{C}, u, i)$ is a star clustering for $\mathcal{G}(G)$. By applying Lemma 2 with $x = u$ and $k = i$, we obtain a contradiction to the fact that $\mathbf{C}$ is an optimal star clustering for $\mathcal{G}(G)$.
Hence, it only remains to consider the case in which $|C(u)| \geq 3$ and $|C_i| \geq 3$. Let us define the star clustering for $\mathcal{G}(G)$

$$\mathbf{C}' = \mathbf{C} \setminus \{C(u), C_i\} \cup \{C(u) \setminus \{u\}\} \cup \{C_i \setminus \{z\}\} \cup \{u, z\}$$

obtained from $\mathbf{C}$ by placing both u and z in an empty cluster. We have

$$\begin{aligned} &\mathsf{SW}(\mathbf{C}') - \mathsf{SW}(\mathbf{C}) \\ =\ & \frac{2(|C(u)|-2)}{|C(u)|-1} + \frac{2(|C_i|-2)}{|C_i|-1} + 1 \\ & - \frac{2(|C(u)|-1)}{|C(u)|} - \frac{2(|C_i|-1)}{|C_i|} \\ =\ & 1 - \frac{2}{|C(u)|(|C(u)|-1)} - \frac{2}{|C_i|(|C_i|-1)} \\ \geq\ & 1 - \frac{1}{3} - \frac{1}{3} > 0 \end{aligned}$$

where the second last inequality follows from $|C(u)| \geq 3$ and $|C_i| \geq 3$. Again, we obtain a contradiction to the fact that $\mathbf{C}$ is an optimal clustering for $\mathcal{G}(G)$ and the proof is complete. □

We are now ready to prove our upper bound on the price of stability.

THEOREM 4. *For any triangle-free graph G,* $\mathsf{PoS}(\mathcal{G}(G)) \leq 4$.

PROOF. Fix a triangle-free graph G and let $\mathbf{C}$ be an optimal star clustering for $\mathcal{G}(G)$. Denote with k the number of non-empty stars in $\mathbf{C}$ and with $k_2 \leq k$ the number of stars in $\mathbf{C}$ of order 2. Clearly, we have $\mathsf{SW}(\mathbf{C}) \geq k$. We claim that the cardinality of a minimum vertex cover VC for G is at most $k + k_2$. To this aim, let V^* be the set of $k + k_2 = (k - k_2) + 2k_2$ nodes of G which are either centers of stars or leaves of stars of order 2 in $\mathbf{C}$. We show that V^* is a vertex cover for G.

Assume, by way of contradiction, that there exists an edge $\{u, v\} \in E$ such that $u \notin V^*$ and $v \notin V^*$. By definition of V^*, it must be $u \in \ell(C(u))$, $v \in \ell(C(v))$ and the order of both $C(u)$ and $C(v)$ is at least 3. Let $\mathbf{C}'$ be the star clustering for $\mathcal{G}(G)$ obtained from $\mathbf{C}$ by removing u from $C(u)$, v from $C(v)$ and creating the new star $\{u, v\}$. We get $\mathsf{SW}(\mathbf{C}') > \mathsf{SW}(\mathbf{C})$ (note that the situation is the same of the last case in the proof of Lemma 3), thus contradicting the optimality of $\mathbf{C}$.

By combining $|VC| \leq |V^*|$, Lemma 1 and $k_2 \leq k$, we obtain

$$\mathsf{SW}(\mathbf{C}^*) < 2|VC| \leq 2(k + k_2) \leq 4k,$$

which, together with $\mathsf{SW}(\mathbf{C}) \geq k$, yields the claim. □

5. BIPARTITE GRAPHS

In this section, we focus on games played on bipartite graphs and provide an upper bound of $6(3 - 2\sqrt{2}) \approx 1.0294$ and a lower bound of 1.003 on the price of stability.

We assume, throughout this section, that we are given a fractional hedonic game $\mathcal{G}(G)$, with G being a bipartite graph. We will denote with $C^* \in \mathbf{C}^*$ a generic, but fixed, cluster in $\mathbf{C}^*$ and with $V^* = \{v_1^*, \ldots, v_p^*\}$ a minimum vertex cover for C^*.

DEFINITION 1. *A fractional assignment of leaves (to stars centered at V^*) is a function $f : C^* \setminus V^* \times [p] \to \mathbb{R}_{\geq 0}$ such that*

1. $\sum_{i \in [p]} f(u, i) = 1$ *for each* $u \in C^* \setminus V^*$,
2. $\sum_{u \in C^* \setminus V^*} f(u, i) > 0$ *for each* $i \in [p]$.

We denote with $\mathcal{F}(V^*)$ the set of all fractional assignments of leaves. The following lemma has been proved in [6].

LEMMA 4 ([6]). *For any minimum vertex cover V^* of C^*, $\mathcal{F}(V^*) \neq \emptyset$.*

DEFINITION 2. *The fractional star clustering (of C^* centered at V^*) induced by f is a collection of p stars $\mathsf{S}^f = (S_1^f, \ldots, S_p^f)$ such that, for each $i \in [p]$, it holds that $c(S_i^f) = v_i^*$ and $\ell(S_i^f) = \{u \in C^* \setminus V^* : f(u, i) > 0\}$, where $f(u, i)$ measures the fractional portion of u which is meant to belong to S_i^f. Denote $x_i = \sum_{u \in \ell(S_i^f)} f(u, i)$.*

We observe that, for each $i \in [p]$, S_i^f is indeed a star since, by the definition of vertex cover, the set of nodes $C^* \setminus V^*$ is an independent set of C^* and, by property 2 in the definition of fractional assignments of leaves, the order of S_i^f is at least two. Hence, because of Lemma 4, it follows that the set of fractional star clusterings induced by the set of fractional assignments of leaves $\mathcal{F}(V^*)$ is non-empty. The social welfare of a fractional star clustering S^f is defined as $\mathsf{SW}(\mathsf{S}^f) = \sum_{i \in [p]} \mathsf{SW}(S_i^f)$, where $\mathsf{SW}(S_i^f) = \frac{2x_i}{x_i + 1}$.

Fix a fractional star clustering of maximum social welfare S^{f^*}. Let $H = \{x_i : i \in [p]\}$ and $h = |H|$. We partition the stars of S^{f^*} into h sets $A_1, \ldots, A_h$ in such a way that $S_i^{f^*} \in A_j$ if and only if x_i is the jth highest value in H. For each $i \in [h]$, define $L_i = \bigcup_{S_j^{f^*} \in A_i} \ell(S_j^{f^*})$ as the set of leaves of all the stars belonging to A_i and $K_i = \bigcup_{S_j^{f^*} \in A_i} \{v_j^*\}$ as the set of centers of all the stars belonging to A_i and denote with $l_i = |L_i|$ and with $k_i = |K_i|$. Observe that, by definition, the sets K_is are pairwise disjoint, while it is possible that two different sets L_i and L_j share some nodes. Anyway, we will show in the sequel that this is not possible.

First of all, we show that the partition $(A_1, \ldots, A_h)$ satisfies the following property.

LEMMA 5. *Fix an edge $\{u, v\} \in E(C^*)$ with $u \in L_i$ for some $i \in [h - 1]$. Then, $v \notin \bigcup_{j \in [h] \setminus [i]} K_j$.*

PROOF. Assume, by way of contradiction, that there exists an edge $\{u, v\} \in E(C^*)$ such that $u \in L_i$ and $v \in K_j$ with $j > i$. Let $S_q^{f^*}$ be a star such that $S_q^{f^*} \in A_i$ and $u \in \ell(S_q^{f^*})$ and let $S_r^{f^*}$ be the star such that $S_r^{f^*} \in A_j$ and $u = c(S_r^{f^*})$. Hence, the function f' obtained from f^* by moving an arbitrarily small probability mass $\epsilon > 0$ from $f^*(u, q)$ to $f^*(u, r)$ belongs to $\mathcal{F}(V^*)$. We obtain

$$\begin{aligned} & \mathsf{SW}(\mathsf{C}^{f^*}) - \mathsf{SW}(\mathsf{C}^{f'}) \\ = \; & \frac{2x_i}{x_i + 1} + \frac{2x_j}{x_j + 1} - \frac{2(x_i - \epsilon)}{x_i + 1 - \epsilon} - \frac{2(x_j + \epsilon)}{x_j + 1 + \epsilon} \\ = \; & 2\epsilon \left(\frac{1}{(x_i + 1)(x_i + 1 - \epsilon)} - \frac{1}{(x_j + 1)(x_j + 1 + \epsilon)} \right) \\ < \; & 0, \end{aligned}$$

where the last inequality comes from $x_i > x_j$ and the arbitrariness of ϵ. We have derived $\mathsf{SW}(\mathsf{C}^{f^*}) < \mathsf{SW}(\mathsf{C}^{f'})$ thus contradicting the optimality of C^{f^*}. □

As a consequence of Lemma 5, we obtain that the sets L_is are pairwise disjoint.

LEMMA 6. *For each $i, j \in [h]$ with $i \neq j$, $L_i \cap L_j = \emptyset$.*

PROOF. Assume, by way of contradiction, that there exist two indices $i, j \in [h]$, with $i < j$, and a node u such that $u \in L_i \cap L_j$. Let $S_r^{f^*}$ be the star such that $S_r^{f^*} \in A_j$ and denote as $v = c(S_r^{f^*})$. Clearly, by definition, it holds that $\{u, v\} \in E(C^*)$. But, since $u \in L_i$, $v \in K_j$ and $i < j$, we derive a contradiction to Lemma 5. □

Now we can exploit Lemma 6 to achieve the following additional property satisfied by the partition $(A_1, \ldots, A_h)$.

LEMMA 7. *For each $i \in [h]$, $x_i = \frac{l_i}{k_i}$.*

PROOF. Fix an index $i \in [h]$. By construction, for each $S_j^{f^*} \in A_i$, it holds that $\sum_{u \in \ell(S_j^{f^*})} f^*(u, j) = x_i$. Hence, by summing over all stars belonging to A_i, we obtain

$$\sum_{S_j^{f^*} \in A_i} \sum_{u \in \ell(S_j^{f^*})} f^*(u, j) = k_i x_i. \tag{2}$$

Moreover, because of Lemma 6 and property 1 of fractional assignments of leaves, it also holds that

$$\sum_{S_j^{f^*} \in A_i} \sum_{u \in \ell(S_j^{f^*})} f^*(u, j) = l_i. \tag{3}$$

By combining equations (2) and (3), we obtain the claim. □

We now show how to suitably round S^{f^*} so as to obtain a star clustering for C^* of high social value.

LEMMA 8. *For each $i \in [h]$, there exists star clustering $\mathsf{S} = (S_1, \ldots, S_{k_i})$ centered at K_i of the set of nodes $K_i \cup L_i$ such that $\sum_{j \in [k_i]} \mathsf{SW}(S_j) = 2\left(\frac{z_i \lfloor x_i \rfloor}{\lfloor x_i \rfloor + 1} + \frac{(k_i - z_i)\lceil x_i \rceil}{\lceil x_i \rceil + 1}\right)$ with $z_i \lfloor x_i \rfloor + (k_i - z_i)\lceil x_i \rceil = l_i$ for some $0 \le z_i \le k_i$.*

PROOF. Fix a set of fractional stars A_i. Let us create a graph $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ as follows: $\widetilde{V} = K_i \cup L_i \cup \{s, t\}$ and $\widetilde{E} = E(K_i \cup L_i) \cup \{\{s, u\} : u \in K_i\} \cup \{\{u, t\} : u \in L_i\}$. Consider now the max flow problem defined on $\widetilde{G}$ by setting to 1 the capacity of each edge in $E(K_i \cup L_i) \cup \{\{u, t\} : u \in L_i\}$ and to $\lceil x_i \rceil$ the capacity of all the remaining edges. Note that any integral flow f for $\widetilde{G}$ induces a star clustering $\mathsf{S}(f)$ centered at K_i of the set of nodes $K_i \cup L_i$. We denote with $\mathsf{S}(f)^-$ the set of stars of $\mathsf{S}(f)$ having order strictly smaller than $\lfloor x_i \rfloor + 1$ and with $o(\mathsf{S}(f)^-)$ the sum of the orders of all the stars in $\mathsf{S}(f)^-$. By the max-flow min-cut theorem and by the definitions of S^{f^*} and A_i, it follows that there exists an integral flow of value at least $k_i x_i$, which, by Lemma 7, coincides with l_i. In particular, we claim that there exists an integral flow of value at least l_i inducing a star clustering S such that each $S \in \mathsf{S}$ has order equal to either $\lfloor x_i \rfloor + 1$ or $\lceil x_i \rceil + 1$. Assume that this is not the case and let $\widetilde{f}$ be the integral flow of value at least l_i that first minimizes the value $|\mathsf{S}(f)^-|$ and then, in case of ties, maximizes the value $o(\mathsf{S}(\widetilde{f})^-)$. By assumption, we have $\mathsf{S}(\widetilde{f})^- \neq \emptyset$. Clearly, by the capacity constraints, there cannot be a star of order strictly greater than $\lceil x_i \rceil + 1$ in $\mathsf{S}(\widetilde{f})$. Thus, since $\widetilde{f}$ has value of at least $k_i x_i$ and $\mathsf{S}(\widetilde{f})$ contains k_i stars, there must be an edge connecting the center of a star $S \in \mathsf{S}(\widetilde{f})^-$ to the leaf u of a star $S' \in \mathsf{S}(\widetilde{f})$ of order $\lceil x_i \rceil + 1$. By moving all the flow routed on edge $\{c(S'), u\}$ to edge $\{c(S), u\}$, we obtain an integral flow f' of the same value of $\widetilde{f}$ and such that $\mathsf{S}(f')^- \subset \mathsf{S}(\widetilde{f})^-$ if the order of S is equal to $\lfloor x_i \rfloor$ or $\mathsf{S}(f')^- = \mathsf{S}(\widetilde{f})^-$ and $o(\mathsf{S}(f')^-) > o(\mathsf{S}(\widetilde{f})^-)$ otherwise. Since in both cases we get a contradiction on the definition of $\widetilde{f}$, it must be $\mathsf{S}(\widetilde{f})^- = \emptyset$. Now let z_i, with $0 \le z_i \le k_i$, be the number of stars in S having order equal to $\lfloor x_i \rfloor + 1$. It follows that $z_i \lfloor x_i \rfloor + (k_i - z_i)\lceil x_i \rceil = l_i$ and $\sum_{j \in [k_i]} \mathsf{SW}(S_j) = 2\left(\frac{z_i \lfloor x_i \rfloor}{\lfloor x_i \rfloor + 1} + \frac{(k_i - z_i)\lceil x_i \rceil}{\lceil x_i \rceil + 1}\right)$. □

By the arbitrariness of C^*, we can conclude that there exists a star clustering S^* such that $\mathsf{SW}(\mathsf{S}^*) = \sum_{i \in [h]} 2\left(\frac{z_i \lfloor x_i \rfloor}{\lfloor x_i \rfloor + 1} + \frac{(k_i - z_i)\lceil x_i \rceil}{\lceil x_i \rceil + 1}\right)$ with $z_i \lfloor x_i \rfloor + (k_i - z_i)\lceil x_i \rceil = l_i$ for some $0 \le z_i \le k_i$. Anyway, S^* may not be a Nash stable clustering. We now show how to obtain a Nash stable clustering from S^* without worsening its social welfare.

LEMMA 9. *There exists a Nash stable clustering $\mathbf{C}$ such that $\mathsf{SW}(\mathbf{C}) \ge \mathsf{SW}(\mathsf{S}^*)$.*

PROOF. Assume that S^* is not Nash stable, otherwise we are done. We obtain $\mathbf{C}$ by manipulating S^* as follows. Whenever there exist two stars $S, S' \in \mathsf{S}^*$ such that $|\ell(S)| > |\ell(S')| + 1$ and $\{c(S'), u\} \in E$ for some $u \in \ell(S)$, remove u from S and add it to S'. It is easy to see that the total social welfare increases. Similarly, whenever there exist two stars $S, S' \in \mathsf{S}^*$ such that $|\ell(S)| \ge 2$, $|\ell(S')| \ge 2$ and $\{u, v\} \in E$ for some $u \in \ell(S)$ and $v \in \ell(S')$, remove u from S, v from S' and create a new star formed by edge $\{u, v\}$. Again, the total social welfare increases. Finally, whenever there exist two stars $S, S' \in \mathsf{S}^*$ such that $|\ell(S)| = 1$, $|\ell(S')| \ge 3$ and $\{u, v\} \in E$ for some $u \in \ell(S)$ and $v \in \ell(S')$, remove v from S' and add it to S. Again, the total social welfare increases. Denote with $\mathbf{C}$ the star clustering obtained at the end of this process. Clearly, $\mathsf{SW}(\mathbf{C}) \ge \mathsf{SW}(\mathsf{S}^*)$. We now show that $\mathbf{C}$ is Nash stable.

Let $u \in C(u) \in \mathbf{C}$ be an agent who possesses an improving deviation in $\mathbf{C}$ by migrating to a star S and let $i = |\ell(C(u))|$ and $j = |\ell(S)|$. Assume $u = c(C(u))$. From one hand, it holds that $p_u(\mathbf{C}) = \frac{i}{i+1}$. From the other hand, since G is triangle-free, it holds that $p_u(\mathbf{C}, u, S) \le \frac{j}{j+2} \le \frac{i+1}{i+3}$, where the last inequality follows from the construction of $\mathbf{C}$. It follows that such an improving deviation is not possible. So, assume $u \in \ell(C(u))$. From one hand, it holds that $p_u(\mathbf{C}) = \frac{1}{i+1}$. If $\{u, v\} \in E$ for some $v \in \ell(S)$, then, by the construction of $\mathbf{C}$, it must be $i = 1 \wedge j = 1$, or $i = 1 \wedge j = 2$, or $i = 2 \wedge j = 1$. It is easy to see that, in all of these three cases, such an improving deviation is not possible and this concludes the proof. □

After having lower bounded the social welfare of a best possible Nash stable clustering, we now exploit the properties of the partition $(A_1, \ldots, A_h)$ to obtain an upper bound on the social welfare of $\mathbf{C}^*$. First, we show an upper bound on the number of edges in C^*, for any $C^* \in \mathbf{C}^*$.

For each $i \in [h]$, define $k_{\le i} = \sum_{j \in [i]} k_j$.

LEMMA 10. $|E(C^*)| \le \sum_{i \in [h]} l_i k_{\le i}$.

PROOF. First note that, because of the fact that $C^* \setminus V^*$ forms an independent set of C^*, there cannot be an edge in $E(C^*)$ connecting two nodes belonging to $\bigcup_{i \in [h]} L_i$. Moreover, by Lemma 5, we also know that, for any $i \in [h-1]$, there cannot be an edge in $E(C^*)$ connecting a node in L_i to a node in $\bigcup_{j \in [h] \setminus [i]} K_j$. Hence, each edge $\{u, v\} \in E(C^*)$ can be of one of the following two types:

1. $u, v \in \bigcup_{i \in [h]} K_i$,
2. $u \in K_i$ and $v \in L_j$ for some $i, j \in [h]$ with $j \ge i$.

Let us denote with E_1 (resp. E_2) the set of edges of type 1 (resp. 2). Clearly, by the above observations, we have $|E(C^*)| = |E_1| + |E_2|$.

Consider now an edge $\{u, v\} \in E_1$. Let $S_q^{f^*}$ be the star such that $c(S_q^{f^*}) = u$ and $S_r^{f^*}$ be the star such that $c(S_r^{f^*}) = v$ and assume $S_q^{f^*} \in A_i$ and $S_r^{f^*} \in A_j$ with $i \le j$. Since G is triangle-free, the existence of edge $\{u, v\} \in E(C^*)$ implies the non-existence of the $|\ell(S_r^{f^*})| \ge 1$ edges of type 2 which can be obtained by connecting u to each node in $\ell(S_r^{f^*})$. By repeating this reasoning for all the edges in E_1, we obtain that $|E_1| + |E_2|$ is upper bounded by the maximum number of edges which can potentially belong to E_2 when assuming $E_1 = \emptyset$. Hence, by the definition of E_2, we obtain $|E(C^*)| \le |E_1| + |E_2| \le \sum_{i \in [h]} l_i k_{\le i}$. □

In order to achieve our desired upper bound, we need the following technical lemma.

LEMMA 11. *Given that $\frac{l_i}{k_i} \ge \frac{l_{i+1}}{k_{i+1}}$ for any $i \in [h-1]$,*

$$\frac{\sum_{i \in [h]} l_i k_{\le i}}{\sum_{i \in [h]} (k_i + l_i)} \le \sum_{i \in [h]} \frac{k_i l_i}{k_i + l_i}.$$

PROOF. We prove the claim by induction on h.

For $h = 1$, the base of the induction is trivially verified.

As to the induction step, by assuming true the claim for $h = n$, we prove that it also holds for $h = n + 1$.

$$\begin{aligned}\frac{\sum_{i\in[n+1]} l_i k_{\leq i}}{\sum_{i\in[n+1]}(k_i+l_i)} &= \frac{\sum_{i\in[n]} l_i k_{\leq i}}{\sum_{i\in[n+1]}(k_i+l_i)} + \frac{l_{n+1}k_{\leq n+1}}{\sum_{i\in[n+1]}(k_i+l_i)} \\ &\leq \frac{\sum_{i\in[n]} l_i k_{\leq i}}{\sum_{i\in[n]}(k_i+l_i)} + \frac{l_{n+1}k_{\leq n+1}}{\sum_{i\in[n+1]}(k_i+l_i)} \\ &\leq \sum_{i\in[n]} \frac{k_i l_i}{k_i+l_i} + \frac{k_{n+1}l_{n+1}}{k_{n+1}+l_{n+1}} \qquad (4) \\ &= \sum_{i\in[n+1]} \frac{k_i l_i}{k_i+l_i}.\end{aligned}$$

Notice that, by the induction hypothesis, $\frac{\sum_{i\in[n]} l_i k_{\leq i}}{\sum_{i\in[n]}(k_i+l_i)} \leq \sum_{i\in[n]} \frac{k_i l_i}{k_i+l_i}$; therefore, in order to prove inequality (4), it remains to show that $\frac{l_{n+1}k_{\leq n+1}}{\sum_{i\in[n+1]}(k_i+l_i)} \leq \frac{k_{n+1}l_{n+1}}{k_{n+1}+l_{n+1}}$, that is equivalent to

$$\frac{k_{\leq n+1}}{\sum_{i\in[n+1]}(k_i+l_i)} \leq \frac{k_{n+1}}{k_{n+1}+l_{n+1}}.$$

Since $\frac{l_i}{k_i} \geq \frac{l_{i+1}}{k_{i+1}}$ for any $i \in [n]$, it holds that $\frac{k_i}{l_i+k_i} \leq \frac{k_{i+1}}{k_{i+1}+l_{i+1}}$ for any $i \in [n]$. Thus, it holds that $\frac{k_{n+1}}{k_{n+1}+l_{n+1}} \geq \frac{k_{i+1}}{k_{i+1}+l_{i+1}}$ for any $i \in [n]$. Let $\alpha = \frac{k_{n+1}}{k_{n+1}+l_{n+1}}$; it holds that

$$\begin{aligned}\frac{k_{\leq n+1}}{\sum_{i\in[n+1]}(k_i+l_i)} &= \frac{\sum_{i\in[n+1]} k_i}{\sum_{i\in[n+1]}(k_i+l_i)} \\ &\leq \frac{\sum_{i\in[n+1]} \alpha(k_i+l_i)}{\sum_{i\in[n+1]}(k_i+l_i)} \\ &= \alpha = \frac{k_{n+1}}{k_{n+1}+l_{n+1}}.\end{aligned}$$

□

COROLLARY 1. *For each* $C^* \in \mathbf{C}^*$, $\mathsf{SW}(C^*) \leq \frac{2l_i}{x_i+1}$.

PROOF. Fix a cluster $C^* \in \mathbf{C}^*$. By Lemmas 10 and 11, it follows that $\mathsf{SW}(C^*) \leq \sum_{i\in[h]} \frac{k_i l_i}{k_i+l_i}$. The claim follows by dividing both the numerator and the denominator of each term in the summation by k_i and by applying Lemma 7. □

We can now conclude this section by showing the following upper bound on the price of stability of games played on bipartite graphs.

THEOREM 5. *For any bipartite graph* G, $\mathsf{PoS}(\mathcal{G}(G)) \leq 6(3-2\sqrt{2}) \approx 1.0294$.

PROOF. Fix a bipartite graph G. By Lemma 9 and Corollary 1, it follows that

$$\mathsf{PoS}(\mathcal{G}(G)) \leq \frac{\sum_{i\in[h]} \frac{l_i}{x_i+1}}{\sum_{i\in[h]}\left(\frac{z_i\lfloor x_i\rfloor}{\lfloor x_i\rfloor+1} + \frac{(k_i-z_i)\lceil x_i\rceil}{\lceil x_i\rceil+1}\right)},$$

where $z_i\lfloor x_i\rfloor + (k_i - z_i)\lceil x_i\rceil = l_i$ for some integer $0 \leq z_i \leq k_i$. Hence, we obtain

$$\mathsf{PoS}(\mathcal{G}(G)) \leq \frac{\sum_{i\in[h]}\left(\frac{z_i\lfloor x_i\rfloor}{x_i+1} + \frac{(k_i-z_i)\lceil x_i\rceil}{x_i+1}\right)}{\sum_{i\in[h]}\left(\frac{z_i\lfloor x_i\rfloor}{\lfloor x_i\rfloor+1} + \frac{(k_i-z_i)\lceil x_i\rceil}{\lceil x_i\rceil+1}\right)}, \qquad (5)$$

where $z_i\lfloor x_i\rfloor + (k_i - z_i)\lceil x_i\rceil = l_i$ for some integer $0 \leq z_i \leq k_i$. Note that the contribution of each term of the summation is maximized when x_i is not an integer. So assume that, for each $i \in [h]$, $\lfloor x_i\rfloor = \alpha_i$ and $\lceil x_i\rceil = \alpha_i + 1$ for some integer $\alpha_i \geq 1$. From $z_i\lfloor x_i\rfloor + (k_i - z_i)\lceil x_i\rceil = l_i = k_i x_i$, we obtain $z_i = k_i(\alpha_i + 1 - x_i)$. By using this equality in (5), we get

$$\mathsf{PoS}(\mathcal{G}(G)) \leq \frac{\sum_{i\in[h]} \frac{\alpha_i(\alpha_i+1-x_i)+(\alpha_i+1)(x_i-\alpha_i)}{x_i+1}}{\sum_{i\in[h]}\left(\frac{\alpha_i(\alpha_i+1-x_i)}{\alpha_i+1} + \frac{(\alpha_i+1)(x_i-\alpha_i)}{\alpha_i+2}\right)}, \qquad (6)$$

where, for each $i \in [h]$, α_i is a positive integer and x_i is a rational number such that $\alpha_i < x_i < \alpha_i + 1$. By using a standard averaging argument in (6), we obtain

$$\begin{aligned}\mathsf{PoS}(\mathcal{G}(G)) &\leq \max_{i\in[h]} \frac{\frac{\alpha_i(\alpha_i+1-x_i)+(\alpha_i+1)(x_i-\alpha_i)}{x_i+1}}{\frac{\alpha_i(\alpha_i+1-x_i)}{\alpha_i+1} + \frac{(\alpha_i+1)(x_i-\alpha_i)}{\alpha_i+2}} \\ &= \max_{i\in[h]} \frac{x_i(\alpha_i+1)(\alpha_i+2)}{(x_i+1)(x_i+\alpha_i^2+\alpha_i)}.\end{aligned}$$

The last quantity is maximized for $x_i = \sqrt{\alpha_i(\alpha_i+1)}$ and $\alpha_i = 1$ which yields the claim. □

In the following theorem, whose proof is omitted due to lack of space, we give a new lower bound on the price of stability for games played on bipartite graphs improving the previous one presented in [6].

THEOREM 6. *There exists a bipartite graph* G *such that* $\mathsf{PoS}(\mathcal{G}(G)) > 1.003$.

6. CONCLUSIONS

We have investigated the quality of Nash stable outcomes in fractional hedonic games defined on undirected and unweighted graphs. It would be worth closing the subtle gap between the lower and the upper bound on the price of stability in bipartite graphs, thus getting its exact value. Providing suitable bounds for more general classes of graphs and better understanding the structure of equilibria for unrestricted topologies are interesting research directions.

Acknowledgements

This work was partially supported by PRIN 2010–2011 research project ARS TechnoMedia: "Algorithmics for Social Technological Networks" and by COST Action IC1205 on Computational Social Choice.

REFERENCES

[1] H. Aziz, F. Brandt, P. Harrenstein. Fractional hedonic games. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 5–12, 2014.

[2] H. Aziz, F. Brandt, H. G. Seedig. Stable partitions in additively separable hedonic games. In *Proceedings of the Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 183–190, 2011.

[3] E. Anshelevich, A. Dasgupta, J. M. Kleinberg, E. Tardos, T. Wexler, T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. *SIAM J. Comput* 38(4):1602–1623, 2008.

[4] K. R. Apt, A. Witzel. A generic approach to coalition formation. *International Game Theory Review*, 11(3):347–267, 2009.

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

[6] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, L. Moscardelli. Nash Stability in Fractional Hedonic Games. In *Proceedings of the 10th Conference on Web and Internet Economics (WINE 2014)*, LNCS 8877:486–491, 2014.

[7] F. Bloch, E. Diamantoudi. Noncooperative formation of coalitions in hedonic games. *International Journal of Game Theory*, 40(2):262–280, 2010.

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

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

[10] M. Charikar, V. Guruswami, A. Wirth. Clustering with qualitative information. *Journal of Computer and System Sciences*, 71(3): 360–383, 2005.

[11] E. D. Demaine, D. Emanuel, A. Fiat, N. Immorlica. Correlation clustering in general weighted graphs. *Theoretical Computer Science*, 361(2-3):172–187, 2006.

[12] J. H. Dréze, J. Greenberg. Hedonic coalitions: optimality and stability. *Econometrica*, 48(4):987–1003, 1980.

[13] E. Elkind, M. Wooldridge. Hedonic coalition nets. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 417–424, 2009.

[14] M. Feldman, L. Lewin-Eytan, J. (S.) Naor. Hedonic clustering games. In *Proceedings of the 24th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 267–276, 2012.

[15] M. Gairing, R. Savani. Computing stable outcomes in hedonic games. In *Proceedings of the 3rd International Symposium on Algorithmic Game Theory (SAGT)*, LNCS 6386:174–185, 2010.

[16] J. R. Marden, Adam Wierman. Overcoming the Limitations of Utility Design for Multiagent Systems. *IEEE Transactions on Automatic and Control*, 58(6):1402–1415, 2013.

[17] M. Olsen. Nash stability in additively separable hedonic games and community structures. *Theory of Computing Systems*, 45(4):917–925, 2009.

[18] M. Olsen. On defining and computing communities. In *Proceedings of the Conferences in Research and Practice in Information Technology*, 97–102, 2012.