Price of Pareto Optimality in Hedonic Games*

Edith Elkind
University of Oxford
Oxford, UK
elkind@cs.ox.ac.uk

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

Michele Flammini
Gran Sasso Science Institute
L'Aquila, Italy
michele.flammini@gssi.it

July 14, 2020

Abstract

The Price of Anarchy measures the welfare loss caused by selfish behavior: it is defined as the ratio of the social welfare in a socially optimal outcome and in a worst Nash equilibrium. Similar measures can be derived for other classes of stable outcomes. We observe that Pareto optimality can be seen as a notion of stability: an outcome is Pareto optimal if and only if it does not admit a deviation by the grand coalition that makes all players weakly better off and some players strictly better off. Motivated by this observation, we introduce the concept of Price of Pareto Optimality: this is an analogue of the Price of Anarchy, with the worst Nash equilibrium replaced with the worst Pareto optimal outcome. We then study this concept in the context of hedonic games, and provide lower and upper bounds on the Price of Pareto Optimality in three classes of hedonic games: additively separable hedonic games, fractional hedonic games, and modified fractional hedonic games.

1 Introduction

The prisoners' dilemma and the tragedy of commons [Osborne and Rubinstein, 1994] are two prominent examples where selfishness causes significant loss of social welfare. In game theory, the outcome of interaction among selfish agents is usually modeled as a Nash equilibrium, i.e., a collection of strategies (one for each player) such that no player wants to change her strategy given the other players' strategies. Consequently, a standard measure of disutility caused by selfish behavior is the *Price of Anarchy* [Koutsoupias and Papadimitriou, 1999]: this is the ratio of the social welfare in a socially optimal outcome of the game and in a worst (social welfare-minimizing) Nash equilibrium of the game. Good upper and lower bounds on the Price of Anarchy have been obtained for many classes of games [see, e.g., Roughgarden and Tardos, 2007]; researchers have also considered the related concept of *Price of Stability* [Correa *et al.*, 2004; Anshelevich *et al.*, 2008], which compares socially optimal outcomes and *best* Nash equilibria.

Importantly, the concept of the Price of Anarchy is defined for a specific notion of stability, namely, Nash equilibrium. Nevertheless, its analogues can be defined for other game-theoretic solution concepts: e.g., Strong Price of Anarchy [Andelman *et al.*, 2007] and Enforcement Value [Ashlagi *et al.*, 2008; Brandt *et al.*, 2009] measure the worst-case welfare loss in, respectively, strong Nash equilibria and correlated equilibria (we refer the reader to the textbook of Osborne and Rubinstein [1994] for the definitions of these concepts; we will briefly discuss strong Nash equilibria in Section 6.2). Indeed, one can extend these concepts beyond normal-form games and explore the worst-case efficiency loss caused by strategic behavior in other types of games.

In this paper, we are interested in exploring Price of Anarchy-like measures in hedonic games. These are games where players form coalitions, and each player has preferences over coalitions that she can be a part of [Drèze and Greenberg, 1980; Banerjee *et al.*, 2001; Bogomolnaia and Jackson, 2002]. While the standard model assumes that players' preferences over coalitions are ordinal, there are several prominent classes of hedonic games where players assign cardinal utilities to coalitions (e.g., a player may assign utilities to individual players, and lift them to coalitions

*A preliminary version of this work has been published in AAAI 2016 [Elkind *et al.*, 2016].

by computing the sum or average of its utilities for players in a coalition). In such settings it is desirable to have a measure of welfare loss caused by stability considerations.

This agenda was recently pursued by Bilò *et al.* [2018] and Kaklamanis *et al.* [2016], who have analyzed the analogues of the Price of Anarchy and the Price of Stability for the concept of Nash stability in the context of *fractional hedonic games*—a simple, but expressive class of hedonic games that was proposed by Aziz *et al.* [2014] (see also the extended version by Aziz *et al.* [2019]). Kaklamanis *et al.* [2016] and Monaco *et al.* [2018] extended this analysis to a variant of this model known as modified fractional hedonic games. A similar study for a different subclass of hedonic games, called social distance games [Brânzei and Larson, 2011], has been performed by Balliu *et al.* [2017a] and Kaklamanis *et al.* [2018]. In a similar spirit, Feldman *et al.* [2015] studied the Price of Anarchy and the Price of Stability in hedonic clustering games, where players are located in a metric space and their utility for a coalition is determined by their distance from a coalition center or from other players in their coalition, and Feldman and Friedler [2015] provided bounds on the Price of Anarchy and the Strong Price of Anarchy for another class of clustering scenarios, which can also be viewed as hedonic games.

Now, Nash stability is a well-known notion of stability for hedonic games, and can be seen as the closest analogue of Nash equilibrium for such games. However, in contrast to normal-form games, where Nash equilibrium is clearly the most prominent solution concept, there are several notions of stability that are commonly studied for hedonic games. Indeed, Nash stability focuses on individual deviations and assumes that any player can join any existing coalition, without asking permission of the coalition members. Given that hedonic games are intended to model group formation, we may consider modifying the notion of a permissible deviation along two dimensions: first, we can allow for group deviations, and second, we can allow (some of) the non-deviators to veto the deviators' moves. These two modifications have the opposite effect: the former enriches the set of actions available to the deviators, while the latter shrinks it. By combining these ideas and their variants, one arrives at the well-known notions of individual stability, contractual individual stability, core, strict core, and several others (see an overview by Aziz and Savani [2015]; Sung and Dimitrov [2007] propose a somewhat different classification).

The classic notion of Pareto optimality has a natural interpretation within this framework. Indeed, according to the standard definition, an outcome is Pareto optimal if there is no other outcome that makes all players weakly better off and some players strictly better off. In the language of deviations and vetoes, this can be restated as follows: an outcome is Pareto optimal if there is no group of players that can deviate (possibly by forming several pairwise disjoint coalitions) so that all of the deviators are weakly better off, some of them are strictly better off, and no non-deviating player is negatively affected by the deviation (and therefore does not want to veto it). Indeed, Pareto optimality is recognized as a valid notion of stability for hedonic games [Morrill, 2010; Aziz and Savani, 2015]. We remark that it can be viewed as a refinement of *contractual strict core* [Sung and Dimitrov, 2007]: the latter is defined similarly, the only difference being that the deviating players should form a single coalition.

In this paper, we introduce and study the *Price of Pareto Optimality (*PPO*)*: this is the ratio of the social welfare in a social welfare-maximizing outcome of the game and the social welfare in a worst Pareto optimal outcome. This concept is a direct analogue of the Price of Anarchy—the only difference is that we maximize over all Pareto optimal outcomes rather than all Nash equilibria. (Note that defining an analogue of the Price of Stability with respect to Pareto optimal outcomes is meaningless: every social welfare-maximizing outcome is Pareto optimal, and therefore the respective quantity would always be 1). While viewing Pareto optimality as a notion of stability is motivated by the analysis of solution concepts in hedonic games, and our technical results pertain to hedonic games, the definition of the PPO applies equally well to arbitrary non-cooperative games. This concept has the following intuitive interpretation. Consider a society where all decisions are made consensually, and therefore the status quo can only be changed if the change does not harm any of the members of the society. The Price of Pareto Optimality is exactly the worst-case loss of total welfare that such a society may experience because of its principles.

While similar measures can be defined for other notions of stability in hedonic games, we believe that the PPO is particularly appealing, because every hedonic game admits a Pareto optimal outcome; in contrast, many well-known classes of hedonic games (including the ones considered in this paper) contain games with no Nash stable outcomes. Thus, the PPO is immune to an important critique of the Price of Anarchy, namely, that it is not clear how to interpret bounds on welfare loss in welfare-pessimal Nash stable outcomes: even if such bounds are not too bad, when a Nash stable outcome does not exist, players may cycle through outcomes with arbitrarily bad social welfare. Indeed, Pareto optimality appears to be the most demanding solution concept for hedonic games that has this property: the set of individually stable or core stable outcomes may be empty, and, while the contractual strict core is always non-empty, the argument above shows that the set of Pareto optimal outcomes is a subset of the contractual strict core.

Our technical contribution in this paper is the study of the PPO in three classes of succinctly representable hedonic games: additively separable hedonic games (ASHGs), fractional hedonic games (FHGs), and a variant of fractional hedonic games, which we call modified fractional hedonic games (mFHGs). In each of these classes, every player assigns a numerical utility to every other player and these utilities are then lifted to coalitions: in ASHGs, the utility of a player i for a coalition C containing i is the sum of her utilities for other players in C, in FHGs her utility for C is the ratio between the sum of her utilities for other players in C and the size of C, and in mFHGs her utility for C is her average utility for other members of C. Formally, if we denote the sum of i's utilities for members of C by $w_i(C)$ (we assume that each player's utility for herself is 0), then her utility for C can be expressed as, respectively, $w_i(C)$ in ASHGs, $\frac{w_i(C)}{|C|}$ in FHGs, and $\frac{w_i(C)}{|C|-1}$ in mFHGs.

We focus on these classes of games as they capture a broad range of coalition formation scenarios. Indeed, it is very natural to measure the value of a coalition by the total utility or the average utility of its members (the difference between FHGs and mFHGs reflects the fact that, when computing the average utility, the player has to decide whether to count herself when calculating the coalition size). They have also been extensively studied in hedonic games literature. In particular, the work of Aziz *et al.* [2011], Aziz *et al.* [2019] and Olsen [2012] provides a discussion of applications of, respectively, additively separable hedonic games, fractional hedonic games and modified fractional hedonic games; for further results, see, e.g., the papers by Aziz *et al.* [2015], Kaklamanis *et al.* [2016], Flammini *et al.* [2017], Bilò *et al.* [2018], Monaco *et al.* [2018], and Bullinger [2020].

There are also several technical reasons to study the PPO in these games. First, unlike in general hedonic games, in these games players assign numerical values to coalitions, which means that the notion of PPO is well-defined in this setting. Further, for fractional hedonic games and modified fractional hedonic games there are bounds in the literature on the analogues of the Price of Anarchy and Price of Stability for the concept of Nash stability [Bilò *et al.*, 2018; Kaklamanis *et al.*, 2016; Monaco *et al.*, 2018], which enables us to directly compare the quality of Pareto optimal outcomes and that of Nash stable outcomes. Finally, the analysis of the PPO for these classes of games presents an interesting technical challenge, and our results provide new insights into the structure of Pareto optimal outcomes in these games. We provide a summary of our technical results, together with a discussion of their significance, in Section 6. However, we believe that the best way to view our results is as a proof of concept, showing that the PPO is a reasonable measure, which can also be investigated in other scenarios (including, but not limited to, other classes of hedonic games). In fact, the preliminary version of the current work has been followed by the study of the PPO in another class of hedonic games [Balliu *et al.*, 2017b], and very recently our result for modified fractional hedonic games has been strengthened by Bullinger [2020].

2 Preliminaries

We consider games with a finite set of players $N = \{1, \ldots, n\}$. A *coalition* is a non-empty subset of N. The set of all players N is called the *grand coalition*, and a coalition of size 1 is called a *singleton coalition*. Given a player $i \in N$, let $\mathcal{N}_i = \{S \subseteq N : i \in S\}$ be the set of all coalitions containing i. For the purposes of this paper, it will be convenient to define a *hedonic game* as a pair $(N, (v_i)_{i\in N})$, where $v_i : \mathcal{N}_i \to \mathbb{R}$ is the *utility function* of player i (traditionally, a hedonic game is defined by endowing each player i with a weak order on $\mathcal{N}_i$; in contrast, in our definition we assume that we are given cardinal representations of these orders). We assume that $v_i(\{i\}) = 0$ for all $i \in N$. For every pair of coalitions $S, S' \in \mathcal{N}_i$, we say that i *strictly prefers* S to S' if $v_i(S) > v_i(S')$; if $v_i(S) = v_i(S')$, we say that i is *indifferent* between S and S'. The *value* of a coalition S is defined as $V(S) = \sum_{i\in S} v_i(S)$. A *coalition structure* (also called a *partition* or an *outcome*) is a partition $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ of N into pairwise disjoint coalitions. We denote by $\mathcal{P}(i)$ the coalition of $\mathcal{P}$ that includes player i; the value $v_i(\mathcal{P}(i))$, also denoted by $v_i(\mathcal{P})$, is the *utility of i in $\mathcal{P}$*.

The *social welfare of a partition* $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ is defined as

$$SW(\mathcal{P}) = \sum_{k=1}^{m} V(P_k) = \sum_{i\in N} v_i(\mathcal{P}(i)).$$

A partition $\mathcal{P}$ is *optimal* if $SW(\mathcal{P}) \geq SW(\mathcal{P}')$ for every other partition $\mathcal{P}'$. We say that $\mathcal{P}$ *Pareto dominates* another partition $\mathcal{P}'$ if $v_i(\mathcal{P}) \geq v_i(\mathcal{P}')$ for every $i \in N$ and $v_i(\mathcal{P}) > v_i(\mathcal{P}')$ for some $i \in N$. A partition $\mathcal{P}$ is *Pareto optimal* if there is no partition $\mathcal{P}'$ that Pareto dominates $\mathcal{P}$. In other words, a Pareto optimal partition is an outcome that does not permit a deviation by the grand coalition that makes all players weakly better off and some players strictly better off. Note that an optimal partition is necessarily Pareto optimal.

Let $\mathbb{P}$ be the set of all Pareto optimal partitions and let $\mathcal{P}^*$ be an optimal partition. Clearly $SW(\mathcal{P}^*) \geq 0$, as the players can form the partition that consists of n singletons.

Definition 2.1. Given a hedonic game $\Gamma = (N, (v_i)_{i \in N})$, the *Price of Pareto Optimality* (PPO) of Γ is defined as

$$\mathsf{PPO}(\Gamma) = \max_{\mathcal{P} \in \mathbb{P}} \frac{SW(\mathcal{P}^*)}{SW(\mathcal{P})}$$

if $SW(\mathcal{P}) > 0$ for all $\mathcal{P} \in \mathbb{P}$, $\mathsf{PPO}(\Gamma) = 1$ if $SW(\mathcal{P}) = 0$ for all $\mathcal{P} \in \mathbb{P}$, and $\mathsf{PPO}(\Gamma) = +\infty$ in all other cases. Given a class of games $\mathcal{G}$, we define the *Price of Pareto Optimality in* $\mathcal{G}$ as $\mathsf{PPO}(\mathcal{G}) = \sup_{\Gamma \in \mathcal{G}} \mathsf{PPO}(\Gamma)$.

We will consider several classes of hedonic games defined on graphs. Let $G = (N, E, w)$ be a weighted directed graph, where N is the node set, E is the edge set, and $w : E \to \mathbb{R}$ is a real-valued edge weight function. We denote a generic edge of G by (i, j) and denote its weight by $w_{i,j}$. We say that G is *unweighted* and write $G = (N, E)$ if $w_{i,j} = 1$ for every $(i, j) \in E$; otherwise we say that G is *weighted*. We say that G is *symmetric* if it holds that (a) $(i, j) \in E$ if and only if $(j, i) \in E$ and (b) $w_{i,j} = w_{j,i}$; for symmetric graphs, we will treat the pair of directed edges connecting i and j as a single undirected edge $\{i, j\}$. The *degree* $\delta_G(i)$ of a node $i \in N$ in a symmetric unweighted graph G is the number of nodes $j \in N$ with $\{i, j\} \in E$. We let $\Delta_G = \max_{i \in N} \delta_G(i)$. The subgraph of G induced by a subset $S \subseteq N$ is denoted by $G_S = (S, E_S)$, where E_S is the subset of all edges in E that have both of their endpoints in S. For readability, when the graph G is clear from the context we denote the degree of a node i in a subgraph G_S of G by $\delta_S(i)$ (instead of $\delta_{G_S}(i)$). A *tree* is a symmetric unweighted graph with at least two nodes that is acyclic and connected. In a tree, a node with degree one is called a *leaf*; a non-leaf node is called an *internal* node. A tree with no internal nodes and only two leaves (i.e., two nodes connected by an edge) is called a 1-*star*. A tree with one internal node and $d \geq 2$ leaves is called a d-*star*. A d-star with $d \geq 2$ is also called a *nondegenerate star*. The only internal node of a nondegenerate star is also referred to as its *center*. We use the term *star* to refer to both 1-stars and nondegenerate stars; thus, in particular, a singleton node is not a star. A tree G is called a (d, e)-*superstar*, where $d, e \geq 2$, if it has a node of degree d (called the *center*) that is adjacent to d internal nodes, and each of these nodes is adjacent to $e - 1$ leaves. Note that a (d, e)-superstar has diameter 4 and admits a vertex cover of size d (which consists of all internal nodes other than the center). A symmetric unweighted graph $G = (N, E)$ is a *clique* if it holds that $\{i, j\} \in E$ for each pair $i, j \in N$; throughout this paper, we only use this term to refer to graphs with at least three nodes. A clique with three nodes is also called *triangle*. We emphasize that, according to our definitions, a singleton node is neither a star nor a clique, and no star graph is a clique.

A *rooted tree* is a tree in which there is a distinguished node called the *root*. The root serves as a reference point to classify the nodes of the tree. Let $G = (N, E)$ be a tree rooted at $r \in N$. The *level* of a node $i \in N$ is the length (number of edges) of the unique path connecting i to r. Notice that the root is at level 0. For every node $i \in N \setminus \{r\}$, the nodes other than i along the unique path connecting i to r are called the *ancestors of* i. A node $j \in N$ is called a *descendant* of node $i \in N$ if i is an ancestor of j; the set of all descendants of i is denoted by $\mathsf{Descendants}_r(G, i)$. If $\{i, j\} \in E$ and $j \in \mathsf{Descendants}_r(G, i)$ then j is called a *child* of i (with respect to r). The set of all children of i is denoted by $\mathsf{Children}_r(G, i)$.

In what follows, we will consider several classes of hedonic games where each player assigns numerical values to all other players, and the utility that player i derives from being in a coalition S is computed based on the values i assigns to the other members of S. Any such game can be associated with a complete weighted directed graph whose set of nodes is the set of all players, and the weight of the edge from player i to player j is the value that i assigns to j. If two players assign value 0 to each other, we omit the respective edge from the graph, so that the resulting graph is no longer complete. We say that a hedonic game is *unweighted* or *symmetric* if its respective graph has these properties, and use other conventional graph-theoretic terminology when speaking about players and coalitions.

3 Additively Separable Hedonic Games

In this section, we consider a well-studied class of hedonic games known as *additively separable hedonic games*. The analysis of the PPO for this class of games turns out to be fairly straightforward, and can be seen as a warm-up for the more sophisticated analysis in subsequent sections.

An *additively separable hedonic game (ASHG)* is defined by a weighted directed graph $G = (N, E, w)$. In this game, the set of players corresponds to the set of nodes, and the utility of player i from a coalition $S \in \mathcal{N}_i$ is given by $v_i(S) = \sum_{(i,j)\in E: j\in S} w_{i,j}$. We denote the additively separable hedonic game that corresponds to a graph G by $\mathcal{H}(G)$.

Our first observation is that if we allow the edge weights to be negative, the Price of Pareto Optimality may be $+\infty$, even if the game is symmetric; this is due to the fact that the the social welfare of a Pareto optimal partition could be negative.

Example 3.1. *Let $N = \{1, 2, 3\}$, $w_{1,2} = w_{2,1} = w_{1,3} = w_{3,1} = 1$, $w_{2,3} = w_{3,2} = -3$. Then the grand coalition is Pareto optimal, as any deviation will lower the utility of player 1. However, its social welfare is negative.*

The game in Example 3.1 is symmetric. However, it contains a cycle. We will now show that Pareto optimal partitions with negative social welfare may exist even in the absence of (directed) cycles as long as we allow asymmetric weights.

Example 3.2. *Let $N = \{1, 2, 3\}$, $w_{1,2} = w_{1,3} = 1$, $w_{2,3} = -3$. Then the grand coalition is Pareto optimal, as any deviation will lower the utility of player 1. However, its social welfare is negative.*

In contrast, if the game is symmetric and acyclic, then it has a unique Pareto optimal partition. Since every social welfare-maximizing partition is Pareto optimal, this implies that the Price of Pareto Optimality is 1 in this case.

Proposition 3.3. *For every additively separable hedonic game $\mathcal{H}(G)$ where $G = (V, E, w)$ is symmetric and every connected component of G is acyclic it holds that every Pareto optimal partition in $\mathcal{H}(G)$ maximizes the social welfare, and hence $\mathsf{PPO}(\mathcal{H}(G)) = 1$.*

Proof. We obtain an optimal partition by removing all negative-weight edges from G, and placing nodes in each connected component in a coalition of their own (note that the notion of a negative-weight edge is well-defined, since the graph is symmetric). Indeed, as the graph is acyclic, the social welfare of this partition is $2\sum_{(i,j)\in E: w_{i,j}>0} w_{i,j}$, which is an upper bound of the social welfare of any partition of N.

We will now argue that for any Pareto optimal partition it holds that if $w_{i,j} > 0$ then i and j belong to the same part of the partition and if $w_{i,j} < 0$ then i and j belong to different parts. this would imply that the social welfare in a Pareto optimal partition equals the social welfare in an optimal partition. Indeed, consider a Pareto optimal partition $\mathcal{P}$, and suppose that $w_{i,j} > 0$, but $\mathcal{P}(i) \neq \mathcal{P}(j)$. Then the deviation where $\mathcal{P}(i)$ and $\mathcal{P}(j)$ merge increases the utility of i and j and does not affect other players. The argument for the case where $w_{i,j} < 0$, but $\mathcal{P}(i) = \mathcal{P}(j)$ is symmetric: splitting this coalition along the edge $\{i, j\}$ increases the utility of i and j and does not affect other players. □

Now, suppose that all weights are non-negative. In this case, every Pareto optimal partition maximizes the social welfare. This holds even if the game is not symmetric, and irrespective of the graph topology.

Proposition 3.4. *For every additively separable hedonic game $\mathcal{H}(G)$ where all weights are non-negative it holds that every Pareto optimal partition in $\mathcal{H}(G)$ maximizes the social welfare, and hence $\mathsf{PPO}(\mathcal{H}(G)) = 1$.*

Proof. If all weights are non-negative, for any pair of players i, j with $w_{i,j} + w_{j,i} > 0$ it holds that in every Pareto optimal outcome i and j belong to the same coalition: if this is not the case, merging i's coalition with j's coalition offers a Pareto improvement. Thus, in every Pareto optimal outcome the utility of each player is maximized and hence the Price of Pareto Optimality is 1. □

The proof of Proposition 3.4 also shows that if all weights are strictly positive, the grand coalition is the unique Pareto optimal outcome; however, in the presence of zero weights there may be several Pareto optimal outcomes.

4 Fractional Hedonic Games

In this section we consider fractional hedonic games. In these games, a player's utility is its average value for the members of its coalition (including itself). Formally, given a weighted directed graph $G = (N, E, w)$, we define a *fractional hedonic game (FHG)* $\mathcal{F}(G)$, where the set of players is N and the utility of player i from a coalition $S \in \mathcal{N}_i$ is given by

$$v_i(S) = \frac{1}{|S|} \sum_{(i,j) \in E : j \in S} w_{i,j}.$$

To illustrate this definition, we will now compute the social welfare of two types of coalitions that will play an important role in our analysis in this section.

Example 4.1. *Suppose that G is a d-star and the players form the grand coalition. If $d \geq 2$, the utility of the center is $\frac{d}{d+1}$ and the utility of each of the d leaves is $\frac{1}{d+1}$, so the social welfare is $\frac{2d}{d+1}$. For $d = 1$ the utility of each player is $\frac{1}{2}$, so the social welfare can be written as $\frac{2d}{d+1}$ in this case as well.*

Now, suppose that G is a (d, e)-superstar and the players form the grand coalition. The utility of the center is $\frac{d}{de+1}$, the utility of each of the d internal nodes other than the center is $\frac{e}{de+1}$, and the utility of each of the $d(e-1)$ leaves is $\frac{1}{de+1}$. Thus, the social welfare is $\frac{2de}{de+1}$.

We first observe that for weighted games the PPO can be unbounded, even if all weights are positive, the game is symmetric, and the underlying graph is a tree.

Proposition 4.2. *For any $M > 1$ there is a symmetric weighted fractional hedonic game $\mathcal{F}(G)$ where $G = (N, E, w)$ is a tree and all weights are positive such that $\mathsf{PPO}(\mathcal{F}(G)) = M$.*

Proof. Let $N = \{1, 2, 3, 4\}$, $E = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3)\}$, $w_{1,2} = w_{2,1} = w_{3,4} = w_{4,3} = 1$, $w_{2,3} = w_{3,2} = 2M$. It is immediate that $\{\{1, 2\}, \{3, 4\}\}$ is Pareto optimal: player 1 would be worse off in any coalition other than $\{1, 2\}$, and, symmetrically, player 4 would be worse off in any coalition other than $\{3, 4\}$. On the other hand, the only partition maximizing the social welfare is $\{\{1\}, \{2, 3\}, \{4\}\}$. $\square$

Therefore from now on we will focus on symmetric unweighted graphs. We collect a few useful observations about games on such graphs in the following proposition.

Proposition 4.3. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$, and let $\mathcal{P} = \{P_1, \ldots, P_m\}$ be a Pareto optimal partition for $\mathcal{F}(G)$. Then*

(a) every coalition in $\mathcal{P}$ is connected,

(b) the set of players in singleton coalitions of $\mathcal{P}$ forms an independent set in G,

(c) if $E \neq \emptyset$, then $\mathcal{P}$ contains at least one non-singleton coalition.

Proof. We prove each statement separately.

(a) Assume for the sake of contradiction that $\mathcal{P}$ contains a coalition P_1 that is not connected. We have $|P_1| \geq 2$. P_1 can be partitioned into two non-empty components S and S' that are not connected to each other, i.e., $P_1 = S \cup S'$, where $S \cap S' = \emptyset$, $|S| \geq 1$, $|S'| \geq 1$ and there is no edge in E connecting a node in S to a node in S'. We prove the claim by showing that the partition $\mathcal{S} = \{S, S', P_2, \ldots, P_m\}$ Pareto dominates $\mathcal{P}$. We first note that $v_h(\mathcal{S}) = v_h(\mathcal{P})$ for every $h \notin P_1$. Moreover, for every player $h \in S$ we have $\mathcal{P}(h) = P_1$ and $\mathcal{S}(h) = S$. Hence, by definition, $v_h(\mathcal{P}) = v_h(P_1) = \frac{1}{|P_1|}\delta_{P_1}(h)$ and $v_h(\mathcal{S}) = v_h(S) = \frac{1}{|S|}\delta_S(h)$. Since $\delta_{P_1}(h) = \delta_S(h)$ and $|P_1| > |S|$, we get $v_h(\mathcal{P}) < v_h(\mathcal{S})$. The same argument applies to every player in S'.

(b) If $E = \emptyset$, then the claim follows trivially. Otherwise, assume for the sake of contradiction that $\mathcal{P}$ contains some singleton coalitions P_1, P_2 such that $P_1 = \{i\}$, $P_2 = \{j\}$, and $\{i, j\} \in E$. Let $S = \{i, j\}$. We prove the claim by showing that the partition $\mathcal{S} = \{S, P_3, \ldots, P_m\}$ Pareto dominates $\mathcal{P}$. We first note that $v_h(\mathcal{S}) = v_h(\mathcal{P})$ for every $h \in N \setminus \{i, j\}$. For player i we have $\mathcal{P}(i) = P_1$ and $\mathcal{S}(i) = S$. Hence, by definition, $v_i(\mathcal{P}) = v_i(P_1) = \frac{1}{|P_1|}\delta_{P_1}(i) = 0$ and $v_i(\mathcal{S}) = v_i(S) = \frac{1}{|S|}\delta_S(i) = \frac{1}{2}$. Thus, $v_i(\mathcal{P}) < v_i(\mathcal{S})$. The same argument applies to player j.

(c) If $|P| = 1$ for every $P \in \mathcal{P}$ and $E \neq \emptyset$, every edge in E connects players in two different singleton coalitions, a contradiction with part (b).

$\square$

Using Proposition 4.3, we can show that for FHGs on symmetric unweighted graphs the Price of Pareto Optimality is upper-bounded by the number of players.

Proposition 4.4. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. Then* $\mathsf{PPO}(\mathcal{F}(G)) \leq |N|$.

Proof. Let $n = |N|$. The utility of every player in a socially optimal partition is at most 1, so the social welfare in an optimal partition is at most n. Now, if $E = \emptyset$, every partition maximizes the social welfare, so there is nothing to prove. Otherwise, consider a Pareto optimal partition $\mathcal{P}$. By Proposition 4.3, $\mathcal{P}$ contains a connected non-singleton coalition. Let P be some such coalition, and let $s = |P|$. Since P is connected, the utility of every player in P is at least $\frac{1}{s}$, so the social welfare of $\mathcal{P}$ is at least $|P| \cdot \frac{1}{s} = 1$, and the bound follows. $\square$

The following theorem upper-bounds the PPO in terms of the maximum degree of the graph.

Theorem 4.5. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. Then* $\mathsf{PPO}(\mathcal{F}(G)) \leq 2\Delta_G(\Delta_G + 1)$.

Proof. Let $\Delta = \Delta_G$ and let N_i denote the set of all neighbors of a node i in G. Let $\mathcal{P}$ be a Pareto optimal partition. Note that the size of each coalition in $\mathcal{P}$ is at most $\Delta(\Delta+1)$. Indeed, if there is a coalition $P \in \mathcal{P}$ with $|P| > \Delta(\Delta+1)$, then the utility of each player in P is strictly less than $\frac{1}{\Delta+1}$. On the other hand, we can take a spanning tree of P (note that P is connected by Proposition 4.3), split it into stars, and obtain a utility of at least $\frac{1}{\Delta+1}$ for everyone in P. This means, in particular, that the utility of each player in a non-singleton coalition in $\mathcal{P}$ is at least $\frac{1}{\Delta(\Delta+1)}$.

Let $\mathcal{P}^*$ be an optimal partition. Let S denote the set of all players that form singleton coalitions in $\mathcal{P}$; by Proposition 4.3 the set S is an independent set in G. Consider an arbitrary player i in $N \setminus S$. Let $D(i) = \mathcal{P}^*(i) \cap S \cap N_i$, and let $d(i) = |D(i)|$. We define the following payment scheme: for each $i \in N \setminus S$ we pay 1 to node i to keep for itself, and also give it another $\frac{d(i)}{d(i)+1}$ units of currency, and ask it to pass on $\frac{1}{d(i)+1}$ to each of the nodes in $D(i)$. In this way, we initially give at most 2 units of payoff to each node in $N \setminus S$ and 0 to nodes in S, but the nodes in S will then receive some transfers from their "neighbors" in the optimal partition $\mathcal{P}^*$.

We will now argue that under this payment scheme each player gets at least as much utility as in the optimal partition $\mathcal{P}^*$. Consider first a player in $N \setminus S$. It gets to keep 1 unit of payoff, and in any outcome of an unweighted fractional hedonic game the utility of every player is at most 1. Now, consider a player $j \in S$. If in $\mathcal{P}^*$ player j also forms a singleton coalition, we are done. Otherwise, let $F(j) = N_j \cap \mathcal{P}^*(j)$. By Proposition 4.3 S forms an independent set in G, so all nodes in $F(j)$ belong to $N \setminus S$. Pick $r \in F(j)$ so that $d(r) \geq d(\ell)$ for all $\ell \in F(j)$. By construction, r belongs to $\mathcal{P}^*(j)$, and all $d(r)$ nodes in $D(r)$ also belong to $\mathcal{P}^*(j)$. Thus, $|\mathcal{P}^*(j)| \geq d(r) + 1$, and therefore the utility of j in $\mathcal{P}^*$ is at most $\frac{1}{d(r)+1}|F(j)|$. On the other hand, if t is some other node in $F(j)$, then it transfers $\frac{1}{d(t)+1} \geq \frac{1}{d(r)+1}$ units of payoff to j (where the inequality holds by our choice of r), so the sum of transfers received by j is at least $\frac{1}{d(r)+1}|F(j)|$, which is exactly what we wanted to prove.

To conclude, under our payment scheme we paid at most 2 to each node in $N \setminus S$, and, after the transfers, each node received at least as much as in $\mathcal{P}^*$. Therefore we have $2|N \setminus S| \geq SW(\mathcal{P}^*)$. Each node in $N \setminus S$ earns at least $\frac{1}{\Delta(\Delta+1)}$ in $\mathcal{P}$, so the social welfare in $\mathcal{P}$ is at least $\frac{1}{\Delta(\Delta+1)}|N \setminus S| \geq \frac{1}{2\Delta(\Delta+1)}SW(\mathcal{P}^*)$. Thus, $\mathsf{PPO} \leq 2\Delta(\Delta + 1)$. $\square$

For FHGs on trees, we can show a better bound on the PPO. We use the following characterization of the structure of optimal partitions.

Lemma 4.6 (Bilò *et al.* [2018]). *Let $G = (N, E)$ be a tree with $|N| \geq 2$, and let $\mathcal{P}^*$ be an optimal partition for the fractional hedonic game $\mathcal{F}(G)$. Then every coalition in $\mathcal{P}^*$ is a star.*

Lemma 4.6 enables us to prove an upper bound on the social welfare of an optimal partition.

Lemma 4.7. *Let $G = (N, E)$ be a tree with $|N| \geq 2$. If G admits a vertex cover of size s then the social welfare of any optimal partition for the fractional hedonic game $\mathcal{F}(G)$ is at most $\frac{2\Delta_G}{\Delta_G+1}s$.*

Proof. Let $\mathcal{P}^* = \{P_1^*, P_2^*, \ldots, P_m^*\}$ be an optimal partition of N into m coalitions. From Lemma 4.6 we know that each coalition P_k^* is a d_k-star with $d_k \geq 1$, hence its value is $V(P_k^*) = \frac{2d_k}{d_k+1}$. We conclude that the social welfare of $\mathcal{P}^*$ can be upper-bounded as follows:

$$SW(\mathcal{P}^*) = \sum_{k=1}^{m} \frac{2d_k}{d_k+1} \leq \frac{2\Delta_G}{\Delta_G+1} \cdot m,$$

where the inequality follows from observing that the value of a star coalition is an increasing function of its degree.

On the other hand, since each coalition in $\mathcal{P}^*$ is a star, by picking an edge from every coalition, we obtain a matching of size m in G. Hence, the size of every vertex cover for G is at least m, and our claim follows. □

Lemma 4.6 states that every coalition in an optimal partition is a star. In contrast, Pareto optimal partitions consist of singletons, stars and superstars, as stated by the following lemma. We defer its proof to Appendix A.2.

Lemma 4.8. *Let $G = (N, E)$ be a tree with $|N| \geq 2$, and let $\mathcal{P}$ be a Pareto optimal partition for the fractional hedonic game $\mathcal{F}(G)$. Then every coalition in $\mathcal{P}$ is a singleton, a star or a superstar.*

Our next proposition describes neighbors of singleton coalitions in Pareto optimal partitions. We defer its proof to Appendix A.3.

Proposition 4.9. *Let $G = (N, E)$ be a tree with $|N| \geq 2$. Let $\mathcal{P}$ be a Pareto optimal partition for the fractional hedonic game $\mathcal{F}(G)$. If $\mathcal{P}$ contains a singleton $P = \{i\}$ then every node j such that $\{i, j\} \in E$ satisfies the following conditions:*

(a) j is not in a singleton coalition in $\mathcal{P}$,

(b) j is not the center of a superstar coalition in $\mathcal{P}$,

(c) j is not the leaf of a superstar coalition in $\mathcal{P}$.

A direct consequence of Proposition 4.9 is that, in a Pareto optimal partition, a player in a singleton coalition can be adjacent only to a node of a d-star with $d \geq 1$, or to the internal nodes of a superstar (other than the center).

We are now ready to present our upper bound on the PPO of fractional hedonic games on trees.

Theorem 4.10. *Let $G = (N, E)$ be a tree with $|N| \geq 2$. Then $\mathsf{PPO}(\mathcal{F}(G)) \leq \Delta_G + 2$.*

Proof. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ be a Pareto optimal partition with m coalitions. Let Q_1 be the set of indices of the star coalitions in $\mathcal{P}$, and let Q_2 be the set of indices of the superstar coalitions in $\mathcal{P}$. Let $Q = Q_1 \cup Q_2$. Then the number of singleton coalitions is $m - |Q|$, and by Proposition 4.3 we have $|Q| \geq 1$. Note that every non-singleton coalition contributes at least 1 to the social welfare and hence $SW(\mathcal{P}) \geq |Q|$.

We will first show that for every $k \in Q$ we can pick a node $r_k \in P_k$ so that

$$V(P_k) \geq \frac{2\delta_{P_k}(r_k)}{\Delta_G + 1}. \tag{1}$$

Indeed, if $k \in Q_1$ and P_k is a nondegenerate star coalition with degree d_k and center r_k, we get

$$V(P_k) = \frac{2d_k}{d_k+1} \geq \frac{2\delta_{P_k}(r_k)}{\Delta_G+1}.$$

If P_k is a 1-star coalition, with r_k being one of its two nodes, we get

$$V(P_k) = 1 \geq \frac{2\delta_{P_k}(r_k)}{\Delta_G+1}.$$

Finally, suppose that $k \in Q_2$ and P_k is a (d_k, e_k)-superstar with center r_k. Since $e_k > 1$, we have

$$V(P_k) = \frac{2d_ke_k}{d_ke_k+1} > \frac{2d_ke_k}{d_ke_k+e_k} = \frac{2\delta_{P_k}(r_k)}{\delta_{P_k}(r_k)+1} \geq \frac{2\delta_{P_k}(r_k)}{\Delta_G+1}. \tag{2}$$

For every $k \in Q$, define the set C_k as follows: for $k \in Q_1$ let $C_k = P_k$, while for $k \in Q_2$ let $C_k = \{j \in P_k \mid \{r_k, j\} \in E\}$, i.e., C_k is the set of internal nodes of P_k other than the center. By construction, C_k is a vertex cover for P_k. Let $C' = \bigcup_{k \in Q} C_k$. Clearly, since $|Q| \geq 1$, we have $C' \neq \emptyset$. Moreover, as C_k covers P_k, every edge inside every non-singleton coalition is covered by C'. By Proposition 4.9 a player in a singleton coalition can only be adjacent to a player in C', so all edges incident with players in singleton coalitions are covered by C'. Thus, the only edges that may be left uncovered by C' are edges that connect nodes in different superstars. Since G is a tree, there are at most $|Q_2| - 1$ such edges, which can be covered by $|Q_2| - 1$ nodes. Thus, G admits a vertex cover of size $|C'| + |Q_2| - 1$, and if C is a minimum vertex cover, we have

$$\begin{aligned} |C| &\leq |C'| + |Q_2| - 1 \\ &= \Big(\sum_{k \in Q_1} (\delta_{P_k}(r_k) + 1) + \sum_{k \in Q_2} \delta_{P_k}(r_k)\Big) + |Q_2| - 1 \\ &= \Big(\sum_{k \in Q_1 \cup Q_2} \delta_{P_k}(r_k)\Big) + |Q| - 1. \end{aligned} \tag{3}$$

Now from (1) we obtain

$$\begin{aligned} SW(\mathcal{P}) = \sum_{k \in Q} V(P_k) &\geq \frac{2}{\Delta_G + 1}\Big(\sum_{k \in Q_1 \cup Q_2} \delta_{P_k}(r_k)\Big) \\ &\geq \frac{2}{\Delta_G + 1}\Big(|C| - |Q| + 1\Big) \qquad (4) \\ &\geq \frac{2}{\Delta_G + 1}\Big(|C| - SW(\mathcal{P}) + 1\Big), \qquad (5) \end{aligned}$$

where (4) follows from (3), and (5) follows from the fact that $SW(\mathcal{P}) \geq |Q|$. From (5) we obtain

$$\Big(1 + \frac{2}{\Delta_G + 1}\Big) SW(\mathcal{P}) \geq \Big(\frac{2}{\Delta_G + 1}\Big) \cdot (|C| + 1) \geq \Big(\frac{2}{\Delta_G + 1}\Big) \cdot |C|,$$

from which

$$SW(\mathcal{P}) \geq \frac{\frac{2}{\Delta_G + 1}}{1 + \frac{2}{\Delta_G + 1}} \cdot |C| = \frac{2}{\Delta_G + 3} \cdot |C|. \tag{6}$$

Let $\mathcal{P}^*$ be an optimal partition. Combining (6) and Lemma 4.7, we get

$$\frac{SW(\mathcal{P}^*)}{SW(\mathcal{P})} \leq \frac{\frac{2\Delta_G}{\Delta_G + 1} \cdot |C|}{\frac{2}{\Delta_G + 3} \cdot |C|} = \Delta_G + \frac{2\Delta_G}{\Delta_G + 1} \leq \Delta_G + 2.$$

□

The following proposition shows that the upper bound given by Theorem 4.10 is optimal up to a small additive error.

Proposition 4.11. *For every $n \in \mathbb{N}$ there exists a fractional hedonic game on a tree $G = (N, E)$ with $|N| \geq n$ for which the Price of Pareto Optimality is strictly greater than $\Delta_G - \frac{1}{\Delta_G}$.*

Proof. Given a $n \in \mathbb{N}$, let $d = n$ and let G be a (d, d)-superstar with center r (note that by definition $d \geq 2$). Consider the partition $\mathcal{P}$ where each leaf of G forms a singleton coalition and the remaining nodes form a d-star[1]. This partition is Pareto optimal: in any partition where the internal nodes do not form a d-star the center has a lower utility. The social welfare of $\mathcal{P}$ is $\frac{2d}{d+1}$.

We now give a lower bound on the social welfare of an optimal partition. Without loss of generality, assume that the neighbors of r in G are $\{1, \ldots, d\}$, and for each $k \in \{1, \ldots, d\}$ let S_k be the set consisting of k and the leaves of G that are adjacent to k.

[1] We are grateful to Panagiotis Kanellopoulos for suggesting this partition; in the preliminary version of this paper, we used another partition, which had higher social welfare

Let $R = S_1 \cup \{r\}$. Consider the partition $\mathcal{P}' = \{R, S_2, S_3, \dots, S_d\}$. The value of R is $\frac{2d}{d+1}$, and the value of every other coalition in $\mathcal{P}'$ is $\frac{2(d-1)}{d}$. Thus, the social welfare of $\mathcal{P}'$ is

$$\begin{aligned} SW(\mathcal{P}') &= \frac{2d}{d+1} + \frac{2(d-1)^2}{d} = (d-1)\left(\frac{2d}{d^2-1} + \frac{2(d-1)}{d}\right) \\ &> (d-1)\left(\frac{2}{d} + \frac{2(d-1)}{d}\right) = 2(d-1), \end{aligned} \tag{7}$$

where (7) follows from the fact that $d^2/(d^2-1) > 1$, and hence $d/(d^2-1) > 1/d$.

We can conclude that the Price of Pareto Optimality can be bounded from below as follows:

$$\frac{SW(\mathcal{P}^*)}{SW(\mathcal{P})} \geq \frac{SW(\mathcal{P}')}{SW(\mathcal{P})} > \frac{2(d-1)(d+1)}{2d} = d - \frac{1}{d} = \Delta_G - \frac{1}{\Delta_G},$$

proving our claim. □

5 Modified Fractional Hedonic Games

In fractional hedonic games, the value that a player i assigns to a coalition is averaged over all members of that coalition, including i itself. Arguably, it is more natural to compute the average value of all *other* members of the coalition. This approach gives rise to a new class of hedonic games, which has been introduced by Olsen [2012] and which we call *modified fractional hedonic games*. Formally, given a weighted directed graph $G = (N, E, w)$ we define a *modified fractional hedonic game (mFHG)* $\mathcal{MF}(G)$, where the set of players is N and the utility of player i from a coalition $S \in \mathcal{N}_i$ with $|S| \geq 2$ is given by

$$v_i(S) = \frac{1}{|S|-1} \sum_{(i,j)\in E: j \in S} w_{i,j};$$

the utility of i from the singleton coalition $\{i\}$ is assumed to be 0.

mFHGs share many properties of FHGs; for instance, the example in Proposition 4.2 can be adapted to show that for weighted graphs the PPO may be unbounded, observations in Proposition 4.3 also apply to mFHGs, and so do the upper bounds on the PPO for general symmetric unweighted graphs (Proposition 4.4 and Theorem 4.5). However, for general symmetric unweighted graphs and unweighted bipartite graphs we can obtain a much stronger upper bound on the PPO.

We start by presenting an example illustrating the differences between FHGs and mFHGs.

Example 5.1. *Just as in Example 4.1, suppose that G is a d-star with $d \geq 2$, and the players form the grand coalition. In the associated mFHG, the utility of the center is 1, and the utility of each of the d leaves is $\frac{1}{d}$, so the social welfare is 2, irrespective of d. If G is a 1-star, the social welfare is 2 as well.*

Now, suppose that G is a (d,e)-superstar with $d \geq 2, e \geq 2$, and the players form the grand coalition. The utility of the center is $\frac{d}{de} = \frac{1}{e}$, the utility of each of the internal nodes other than the center is $\frac{e}{de} = \frac{1}{d}$, and the utility of each of the leaves is $\frac{1}{de}$. Note that the grand coalition is not Pareto optimal in this setting: if one of the neighbors of the center deviates together with all of its adjacent leaves to form a separate coalition of size e, this does not affect the utility of the center and makes all other players better off.

On the other hand, a clique of any size is Pareto optimal, as every player has the highest possible utility of 1. Conversely, if a partition $\mathcal{P}$ contains a coalition P that is not a clique, but can be partitioned into cliques, then it is not Pareto optimal: replacing P by a collection of cliques that form a partition of P makes no player worse off and, at the same time, makes some player (e.g., a player with fewer than $|P|-1$ neighbors) better off.

In some of the proofs of this section, we will make use of the notion of matching, Hall's theorem and König's theorem, that we summarize in the following.

Let $G = (N, E)$ be a symmetric graph. For every subset of nodes $A \subseteq N$, let $\mathcal{D}_G(A)$ be the set of all nodes in $N \setminus A$ that are adjacent to some node in A. Given any pair of subsets of nodes $A_1, A_2 \subseteq N$, a *matching* between A_1 and A_2 in G is an injective mapping $f : A_1 \mapsto A_2$ such that $\{i, f(i)\} \in E$ for every $i \in A_1$, and if $i \in A_1 \cap A_2$ and $f(i) \in A_1 \cap A_2$ then it must hold that $f(f(i)) = i$. We say that the nodes in $A_1 \cup A_2$ are *matched* by f. Equivalently, the matching f can be specified by the set of edges $\{\{i, f(i)\} \mid i \in A_1\}$, which we denote by $\mathcal{F}(f)$. We refer to $|\mathcal{F}(f)|$ as the *size* of f. A *maximum* matching in G is a matching with maximum size, i.e., $f^* \in \arg\max_{f \in Match(G)} |\mathcal{F}(f)|$, where $Match(G)$ is the set of all matchings in G.

Theorem 5.2 (Hall's theorem [Diestel, 2005])**.** *Let $G = (N_1 \cup N_2, E)$ be a symmetric unweighted bipartite graph. There exists a matching $f : N_1 \mapsto N_2$ if and only if $|\mathcal{D}_G(A)| \geq |A|$ for every subset $A \subset N_1$.*

Theorem 5.3 (König's theorem [Diestel, 2005])**.** *Let G be a symmetric unweighted bipartite graph. The size of any maximum matching in G is equal to the size of any minimum vertex cover of G.*

We are now ready to show our first result. We will derive a bound on the optimal social welfare in mFHGs. Our proof is very similar to the proof of the bound on the optimal social welfare in FHGs by Bilò *et al.* [2018].

Proposition 5.4. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. Let $\mathcal{P} = \{P_1^*, P_2^*, \ldots, P_m^*\}$ be an optimal partition for $\mathcal{MF}(G)$. If G admits a vertex cover of size s then the social welfare of $\mathcal{P}$ is at most $2s$.*

Proof. Let C be a minimum vertex cover of G and consider some $k \in \{1, \ldots, m\}$. If $|P_k^*| = 1$, we have $V(P_k^*) = 0$. Thus, suppose that $|P_k^*| \geq 2$. Let $C_k = P_k^* \cap C$ and let $I_k = P_k^* \setminus C$; note that $P_k^* = C_k \cup I_k$, $C = \bigcup_{1 \leq k \leq m} C_k$, and I_k is an independent set. We have

$$\begin{aligned} V(P_k^*) &= \frac{\sum_{i \in P_k^*} \delta_{P_k^*}(i)}{|C_k| + |I_k| - 1} \\ &\leq \frac{2\Big(|C_k||I_k| + \frac{1}{2}|C_k|(|C_k| - 1)\Big)}{|C_k| + |I_k| - 1} \\ &= 2|C_k| \frac{|I_k| + \frac{1}{2}(|C_k| - 1)}{|C_k| + |I_k| - 1} \leq 2|C_k|. \end{aligned} \tag{8}$$

The last inequality in (8) holds trivially if $|C_k| = 0$ and for $|C_k| \geq 1$ it follows from observing that $|I_k| + \frac{1}{2}(|C_k| - 1) \leq |C_k| + |I_k| - 1$ whenever $|C_k| \geq 1$.

We can now bound the social welfare of $\mathcal{P}$:

$$SW(\mathcal{P}) = \sum_{k=1}^{m} V(P_k^*) \leq \sum_{k=1}^{m} 2|C_k| = 2|C|.$$

$\square$

A crucial difference between FHGs and mFHGs is the structure of Pareto optimal outcomes: it turns out that in mFHGs Pareto optimal outcomes consist of singletons, stars and cliques. The proof of Theorem 5.5 requires a substantial amount of groundwork, provided by Claims 1–14; we conclude the proof on p. 20.

Theorem 5.5. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. Let $\mathcal{P}$ be a Pareto optimal partition for the modified fractional hedonic game $\mathcal{MF}(G)$. Then every coalition in $\mathcal{P}$ is a singleton, a star, or a clique.*

Proof. Assume for the sake of contradiction that $\mathcal{P}$ contains a coalition P that is not a singleton, a star, or a clique. We will show that we can obtain a new partition $\mathcal{S}$ that Pareto dominates $\mathcal{P}$ by decomposing P into a set of stars and triangles.

Let $G_P = (P, E_P)$ be the subgraph of G induced by P. By Proposition 4.3, the graph G_P is connected. Let C be a minimum vertex cover of G_P. Let $I = P \setminus C$; note that I is an independent set in G_P. Consider a function $f : E_P \mapsto \mathbb{R}^+ \cup \{0\}$ such that

(a) $f(i,j) = 0$ if $\{i,j\} \notin I \times C$,

(b) $\sum_{j \in C} f(i,j) = 1$ for each $i \in I$.

We refer to f as a *fractional assignment* of nodes in I to nodes in C. Given a fractional assignment f, for every node $j \in C$ we set

$$L_j^f = \{i \in I \mid f(i,j) > 0\} \quad \text{and} \quad x_j^f = \sum_{i \in I} f(i,j).$$

The quantity x_j^f is called the *value* of node j with respect to f.

Our first objective is to define a fractional assignment f^* which evenly distributes the fractions of the nodes in I to the nodes in C. As we show in the sequel, this property guarantees that f^* can be converted into an integral assignment which can be interpreted as a partition of G_P into starts and triangles. To conclude the proof, we will argue that in the combined partition the utility of every node in P is at least as high as its utility in $\mathcal{P}$, and there is at least one node whose utility strictly increases.

Let $\phi : \mathbb{R}^+ \cup \{0\} \mapsto \mathbb{R}^+ \cup \{0\}$ be an arbitrary positive strictly increasing and strictly concave function of its argument; for instance, we can take $\phi(x) = \frac{x}{x+1}$. Given a fractional assignment f, we define $\xi(f)$ as

$$\xi(f) = \sum_{j \in C} \phi(x_j^f),$$

We define f^* to be a fractional assignment that maximizes the function ξ over the polytope defined by the (linear) constraints (a) and (b).

We first use f^* to partition the set C into $h+1$ pairwise disjoint subsets $C_0, C_1, \ldots, C_h$ as follows. Let H be a set of real values defined as $H = \{x_j^{f^*} \geq 1 \mid j \in C\}$, and set $h = |H|$. The set C_0 consists of all the players $j \in C$ such that $x_j^{f^*} < 1$, for each $k \in \{1, \ldots, h\}$ we have $x_j^{f^*} = x_\ell^{f^*}$ for all $j, \ell \in C_k$ and for each k, k' with $0 \leq k < k' \leq h$ and all $j \in C_k$, $\ell \in C_{k'}$ we have $x_j^{f^*} < x_\ell^{f^*}$. Now, let us divide the set I into $h+1$ sets $I_0, I_1, \ldots, I_h$ defined as

$$I_k = \bigcup_{j \in C_k} L_j^{f^*} \qquad \text{for every } k \in \{0, 1, \ldots, h\}.$$

Further, for every $k \in \{1, \ldots, h\}$, set

$$C^{\geq k} = \bigcup_{k \leq t \leq h} C_t, \qquad I^{\geq k} = \bigcup_{k \leq t \leq h} I_t.$$

In Claim 1 and Claim 2 we establish important properties of these sets. Specifically, in Claim 1 we show that no edge can connect a node in I_k to a node in $C_{k'}$, with $k' < k$; in Claim 2 we show that if there is an edge connecting a node i in I_k to a node j in $C_{k'}$, with $k' > k$, then f^* does not assign any portion of i to j, i.e., $f^*(i,j) = 0$. An immediate consequence of these two claims is that the sets I_k are pairwise disjoint, i.e., for every $k, k' \in \{0, 1, 2, \ldots, h\}$ with $k \neq k'$ we have $I_k \cap I_{k'} = \emptyset$. The subdivision of the two sets C and I induced by f^* is illustrated in Figure 1.

Claim 1. *For every pair of integers k, k' with $0 \leq k' < k \leq h$ and for every $i \in I_k$ and $\ell \in C_{k'}$ we have $\{i, \ell\} \notin E$.*

Proof. Fix k, k' with $0 \leq k' < k \leq h$, and assume for the sake of contradiction that $\{i, \ell\} \in E$ for some $i \in I_k$, $\ell \in C_{k'}$. Since $i \in I_k$, there exists a node $j \in C_k$ such that $f^*(i,j) > 0$. Moreover, since $k > k'$, it holds that $x_j^{f^*} > x_\ell^{f^*}$. Set $\epsilon = \min\left\{f^*(i,j), \frac{x_j^{f^*} - x_\ell^{f^*}}{2}\right\}$ and define a new fractional assignment $\bar{f}$ as follows:

$$\bar{f}(i',t) = \begin{cases} f^*(i',t) - \epsilon & \text{if } i' = i, t = j \\ f^*(i',t) + \epsilon & \text{if } i' = i, t = \ell \\ f^*(i',t) & \text{otherwise.} \end{cases}$$

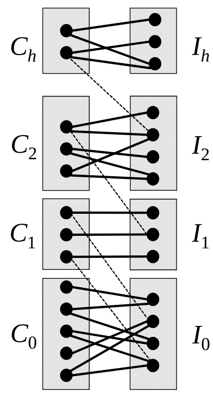

Figure 1: A fractional assignment. An edge $\{i, j\}$ is shown as a solid line if it may be the case that $f^*(i, j) > 0$; an edge $\{i, j\}$ such that $f^*(i, j)$ is necessarily 0 is shown as a dotted line. Edges between nodes in C are not shown.

We have $x_j^{\bar{f}} = x_j^{f^*} - \epsilon$, $x_\ell^{\bar{f}} = x_\ell^{f^*} + \epsilon$ and $x_t^{\bar{f}} = x_t^{f^*}$ for every $t \notin \{j, \ell\}$. Note that we defined ϵ so that $\bar{f}(i, j) \geq 0$, $\bar{f}(i, \ell) \geq 0$ and $x_j^{\bar{f}} \geq x_\ell^{\bar{f}}$. We denote by D_j the decrease in j's contribution to the value of ξ, and by D_ℓ the increase in ℓ's contribution:

$$D_j = \phi(x_j^{f^*}) - \phi(x_j^{\bar{f}}), \qquad D_\ell = \phi(x_\ell^{\bar{f}}) - \phi(x_\ell^{f^*}).$$

We have $\xi(\bar{f}) = \xi(f^*) - D_j + D_\ell$. Since ϕ is a positive strictly increasing and strictly concave function of its argument, we obtain $D_j < D_\ell$ and hence $\xi(\bar{f}) > \xi(f^*)$, which contradicts the optimality of f^*. □

Claim 2. *For every pair of integers k, k' with $0 \leq k < k' \leq h$ and for every $i \in I_k$ and $\ell \in C_{k'}$ we have $f^*(i, \ell) = 0$.*

Proof. Fix k, k' with $0 \leq k < k' \leq h$ and assume for the sake of contradiction that $f^*(i, \ell) > 0$ for some $i \in I_k$, $\ell \in C_{k'}$. Since $i \in I_k$, there exists a node $j \in C_k$ such that $f^*(i, j) > 0$. Moreover, since $k' > k$, it holds that $x_\ell^{f^*} > x_j^{f^*}$. Set $\epsilon = \min\left\{f^*(i, \ell), \frac{x_\ell^{f^*} - x_j^{f^*}}{2}\right\}$ and define a new fractional assignment $\bar{f}$ as follows:

$$\bar{f}(i', t) = \begin{cases} f^*(i', t) + \epsilon & \text{if } i' = i, t = j \\ f^*(i', t) - \epsilon & \text{if } i' = i, t = \ell \\ f^*(i', t) & \text{otherwise.} \end{cases}$$

We have $x_\ell^{\bar{f}} = x_\ell^{f^*} - \epsilon$, $x_j^{\bar{f}} = x_j^{f^*} + \epsilon$ and $x_t^{\bar{f}} = x_t^{f^*}$ for every $t \notin \{j, \ell\}$. Note that we have defined ϵ so that $\bar{f}(i, \ell) \geq 0$, $\bar{f}(i, j) \geq 0$ and $x_\ell^{\bar{f}} \geq x_j^{\bar{f}}$. We denote by D_j the increase in j's contribution to the value of ξ, and by D_ℓ the decrease of ℓ's contribution:

$$D_j = \phi(x_j^{\bar{f}}) - \phi(x_j^{f^*}), \qquad D_\ell = \phi(x_\ell^{f^*}) - \phi(x_\ell^{\bar{f}}).$$

We have $\xi(\bar{f}) = \xi(f^*) + D_j - D_\ell$. Since ϕ is a positive strictly increasing and strictly concave function of its argument, we obtain $D_\ell < D_j$ and hence $\xi(\bar{f}) > \xi(f^*)$, which contradicts the optimality of f^*. □

Together, Claims 1 and 2 imply that the sets I_k, $k = 0, 1, \ldots, h$, are pairwise disjoint. In the remainder of the proof, we will first show how to partition each set $C_k \cup I_k$ with $k \in \{1, \ldots, h\}$ into a collection of stars with centers in C_k so that each player in $C^{\geq 1} \cup I^{\geq 1}$ weakly prefers the new partition to $\mathcal{P}$, and, if $C^{\geq 1} \cup I^{\geq 1} \neq \emptyset$, then at least one player in $C^{\geq 1} \cup I^{\geq 1}$ strictly prefers the new partition to $\mathcal{P}$. Then, we show how to partition $C_0 \cup I_0$ into a collection of stars and triangles so that each player in $C_0 \cup I_0$ weakly prefers the new partition to $\mathcal{P}$ and, if $C^{\geq 1} \cup I^{\geq 1} = \emptyset$, at least one player in $C_0 \cup I_0$ strictly prefers the new partition to $\mathcal{P}$. This enables us to conclude that the new partition Pareto dominates $\mathcal{P}$.

Since the sets I_k are pairwise disjoint, it follows that for every $k \in \{0, 1, 2, \dots, h\}$ it holds that the nodes in I_k distribute their unit of value among the nodes in C_k:

$$|I_k| = \sum_{i \in I_k} \sum_{j \in C} f^*(i,j) = \sum_{j \in C_k} x_j^{f^*}.$$

Since $x_j^{f^*} = x_\ell^{f^*}$ for each $k \geq 1$ and each $j, \ell \in C_k$, from the previous identity we obtain

$$x_j^{f^*} = \frac{|I_k|}{|C_k|} \qquad \text{for each } k \geq 1 \text{ and each } j \in C_k. \tag{9}$$

Note that for each $k \geq 1$ we have $x_j^{f^*} \geq 1$ for all $j \in C_k$, so equation (9) implies that $|C_k| \leq |I_k|$ for all $k \geq 1$. Similarly, since for each k, k' with $0 \leq k < k' \leq h$ and all $j \in C_k$, $\ell \in C_{k'}$ we have $x_j^{f^*} < x_\ell^{f^*}$, equation (9) implies that $\frac{|I_k|}{|C_k|} < \frac{|I_{k'}|}{|C_{k'}|}$ for each k, k' with $0 \leq k < k' \leq h$.

Claim 3. *For every $k \in \{1, 2, \dots, h\}$ there exists a partition of the set I_k into $|C_k|$ sets $(I_k^j)_{j \in C_k}$ such that for each $j \in C_k$ we have $\{i, j\} \in E$ for each $i \in I_k^j$ and $|I_k^j| \in \left\{ \left\lfloor \frac{|I_k|}{|C_k|} \right\rfloor, \left\lceil \frac{|I_k|}{|C_k|} \right\rceil \right\}$.*

Proof. We prove the claim by defining a instance of the network flow problem, and using an integral flow in this network to obtain the desired partition.

Fix a $k \in \{1, 2, \dots, h\}$ and let $x = \frac{|I_k|}{|C_k|}$. We construct an s-t flow network as follows. The set of nodes is $\{s, t, u\} \cup C_k \cup I_k$, where s, t are the source and the sink, respectively. The set of arcs is

$$\{(s,j) \mid j \in C_k\} \cup \{(s,u)\} \cup \{(u,j) \mid j \in C_k\} \cup \{(j,i) \mid j \in C_k, i \in I_k, \{j,i\} \in E\} \cup \{(i,t) \mid i \in I_k\}.$$

The arc capacities are defined as follows. We set $c(s,j) = \lfloor x \rfloor$ for every $j \in C_k$, and $c(s,u) = |I_k| - \lfloor x \rfloor |C_k|$. All other arcs have capacity 1. Note that all capacities are integer and $c(s,u) = 0$ if x is an integer.

We define a flow φ^* as follows.

- $\varphi^*(s,j) = c(s,j)$ for every $j \in C_k$,
- $\varphi^*(s,u) = c(s,u)$,
- $\varphi^*(u,j) = x - \lfloor x \rfloor$ for every $j \in C_k$,
- $\varphi^*(i,t) = c(i,t)$ for each $i \in I_k$,
- $\varphi^*(j,i) = f^*(i,j) \leq 1$ for every $\{j,i\} \in E$ such that $j \in C_k$ and $i \in I_k$.

It follows from (9) that φ^* is a feasible flow from the source s to the sink t with value $|I_k|$. It is a maximum flow, as its value is equal to the sum of the capacities of the edges leaving s:

$$c(s,u) + \sum_{j \in C_k} c(s,j) = |I_k| - \lfloor x \rfloor |C_k| + \sum_{j \in C_k} \lfloor x \rfloor = |I_k|.$$

Since all capacities are integer, by the integrality theorem it follows that there exists an integral flow φ of value $|I_k|$. Set $I_k^j = \{i \in I_k \mid \varphi(j,i) = 1\}$. By construction, we have $\{i,j\} \in E$ for each $i \in I_k^j$, and it remains to show that $|I_k^j| \in \{\lfloor x \rfloor, \lceil x \rceil\}$.

Observe that φ, in order to reach value $|I_k|$, must necessarily route $c(s,j)$ units of flow along (s,j) for every $j \in C_k$ and $c(s,u)$ units of flow along (s,u). Moreover, exactly $c(s,u)$ of the arcs leaving the node u must carry one unit of flow. Therefore there are exactly $c(s,u)$ nodes $j \in C_k$ that have outgoing flow $c(s,j)+c(u,j) = \lfloor x \rfloor + 1$; for the remaining nodes $j \in C_k$ the flow is $c(s,j) = \lfloor x \rfloor$. Since φ is integral, the flow along each arc (j,i) with $j \in C_k$ and $i \in I_k$ is either 0 or 1. To establish our claim, it remains to note that if x is not an integer then $\lfloor x \rfloor + 1 = \lceil x \rceil$, and otherwise $c(s,u) = 0$. $\square$

Now, for each $k \in \{1, \dots, h\}$ and each $j \in C_k$, set $S_k^j = \{j\} \cup I_k^j$. By Claim 3 each coalition S_k^j is a star with center j. We add all such coalitions to $\mathcal{S}$. We will now argue that all players in these coalitions weakly prefer $\mathcal{S}$ to $\mathcal{P}$ and, moreover, some of them strictly prefer $\mathcal{S}$ to $\mathcal{P}$.

For each $k \in \{1, \dots, h\}$ and each $j \in C_k$ player j's utility in $\mathcal{S}$ is 1, hence at least as large as its utility in $\mathcal{P}$. Now, consider a player $i \in S_k^j \setminus \{j\}$. Its utility in $\mathcal{S}$ is at least $\frac{1}{\lceil \frac{|I_k|}{|C_k|} \rceil}$. On the other hand, by Claim 1, the utility of i in $\mathcal{P}$ is at most $\frac{|C^{\geq k}|}{|P|-1}$. We have

$$\frac{|C^{\geq k}|}{|P|-1} \leq \frac{|C^{\geq k}|}{|C^{\geq k}|+|I^{\geq k}|-1} = \frac{1}{1+\frac{|I^{\geq k}|}{|C^{\geq k}|}-\frac{1}{|C^{\geq k}|}} \leq \frac{1}{1+\frac{\sum_{t=k}^{h}|I_t|}{\sum_{t=k}^{h}|C_t|}-\frac{1}{|C_k|}} \leq \frac{1}{1+\frac{|I_k|}{|C_k|}-\frac{1}{|C_k|}} \leq \frac{1}{\left\lceil \frac{|I_k|}{|C_k|} \right\rceil}, \quad (10)$$

where we use the fact that $\frac{|I_{k'}|}{|C_{k'}|} \geq \frac{|I_k|}{|C_k|}$ for every $k' > k$, and observe that for every pair of positive integers x, y it holds that $1 + \frac{x}{y} - \frac{1}{y} \geq \lceil \frac{x}{y} \rceil$. Thus, player j weakly prefers $\mathcal{S}$ to $\mathcal{P}$.

Now, if $k > 1$ or if $C_0 \cup I_0 \neq \emptyset$, the first inequality in (10) is strict, so at least one of the players in I_k strictly prefers $\mathcal{S}$ to $\mathcal{P}$. It remains to identify a strictly improving player in the case where $C_0 = I_0 = \emptyset$ and $h = 1$, i.e., $C = C_1$, $I = I_1$. Note first that we have $|C| > 1$, since otherwise P would be a star. There exists a player $j \in C$ such that $|I_1^j| = \left\lfloor \frac{|I|}{|C|} \right\rfloor$. In $\mathcal{P}$ the utility of each player in I_1^j is at most $\frac{|C|}{|C|+|I|-1}$, whereas its utility in $\mathcal{S}$ is

$$\frac{1}{\left\lfloor \frac{|I|}{|C|} \right\rfloor} \geq \frac{|C|}{|I|} > \frac{|C|}{|C|+|I|-1},$$

where the last inequality uses the fact that $|C| > 1$. Thus, every player in I_1^j strictly prefers $\mathcal{S}$ to $\mathcal{P}$.

In the remainder of the proof we deal with the nodes in C_0 and I_0. Let $Z = C_0 \cup I_0$. We will show how to partition Z into stars and triangles so that the utility of every node in Z is at least as high as in $\mathcal{P}$. This is not yet sufficient to establish that $\mathcal{P}$ is not Pareto optimal: while we have argued that if $C^{\geq 1} \cup I^{\geq 1} \neq \emptyset$ there is at least one player in $C^{\geq 1} \cup I^{\geq 1}$ that strictly prefers $\mathcal{S}$ to $\mathcal{P}$, we cannot rule out the possibility that $C^{\geq 1} \cup I^{\geq 1}$ is empty. Thus, to complete the proof, we will show that if $C^{\geq 1} \cup I^{\geq 1} = \emptyset$, i.e., if $P = Z$, then for some node in Z its utility in $\mathcal{S}$ is strictly higher than its utility in $\mathcal{P}$.

We describe the partitioning of Z by a sequence of claims. Recall that, given a matching f with domain A, we denote by $\mathcal{F}(f)$ the set of edges specified by f, i.e., $\mathcal{F}(f) = \{\{i, f(i)\} \mid i \in A\}$.

Claim 4. *There exists a matching $C_0 \mapsto Z$.*

Proof. Let us consider the subgraph $G_{C_0} = (C_0, E_{C_0})$ induced by the set C_0. Let $M \subset E_{C_0}$ be a maximum-size collection of edges such that $e \cap e' = \emptyset$ for every pair $e, e' \in M$ with $e \neq e'$. Let $X \subseteq C_0$ be the set of nodes that are not incident with any edge in M. Note that if there is a matching $C_0 \mapsto C_0$ then X is empty. X is an independent set in G_{C_0}, and hence the subgraph $G_{X \cup I_0}$ induced by $X \cup I_0$ is bipartite. For every $Y \subseteq X$ we have $\mathcal{D}_{G_{X\cup I_0}}(Y) \geq Y$: otherwise, $(C \setminus Y) \cup \mathcal{D}_{G_{X\cup I_0}}(Y)$ would be a vertex cover for G_P that is strictly smaller than C, a contradiction with our choice of C. Hence, by Theorem 5.2 there exists a matching $f : X \mapsto I_0$. Therefore we can define a matching $g : C_0 \mapsto Z$ such that $\mathcal{F}(g) = M \cup \mathcal{F}(f)$. $\square$

Let us denote by $\mathcal{G}$ the set of matchings $C_0 \mapsto Z$. Given a matching $g \in \mathcal{G}$, we partition the set $\mathcal{F}(g)$ into two sets, $Q_{CC}(g)$ and $Q_{CI}(g)$, defined as follows:

$$Q_{CC}(g) = \{\{i, g(i)\} \in E \mid i \in C_0, g(i) \in C_0\}, \qquad Q_{CI}(g) = \{\{i, g(i)\} \in E \mid i \in C_0, g(i) \in I_0\}.$$

We denote by $C_0(g)$ the subset of nodes in C_0 that are incident with edges in $Q_{CI}(g)$, and by $I_0(g)$ the subset of nodes in I_0 that are incident with edges in $Q_{CI}(g)$. Note that I_0 may contain nodes that are not incident with any of the edges of $\mathcal{F}(g)$. We denote the set of all such nodes by $\mathit{Free}(g)$, i.e., $\mathit{Free}(g) = I_0 \setminus I_0(g)$. Finally, let us denote by $\mathcal{G}^*$ the subset of matchings in $\mathcal{G}$ that maximize the number of nodes in C_0 matched to nodes in I_0, i.e., $\mathcal{G}^* = \{g' \in \mathcal{G} \mid g' \in \arg\max_{g\in\mathcal{G}} |Q_{CI}(g)|\}$ (or, equivalently, $\mathcal{G}^* = \{g' \in \mathcal{G} \mid g' \in \arg\min_{g\in\mathcal{G}} |\mathit{Free}(g)|\}$).

Claim 5. *For every matching $g \in \mathcal{G}^*$, every $\{i,j\} \in Q_{CC}(g)$, and every pair of distinct nodes $i', j' \in \mathit{Free}(g)$ it is not the case that $\{i,i'\} \in E$ and $\{j,j'\} \in E$.*

Proof. Fix a matching $g \in \mathcal{G}^*$ and assume for the sake of contradiction that there is an edge $\{i,j\} \in Q_{CC}(g)$ and two nodes $i', j' \in \mathit{Free}(g)$, $i' \neq j'$, such that $\{i,i'\} \in E$ and $\{j,j'\} \in E$. Define a new mapping g' so that $\mathcal{F}(g') = (\mathcal{F}(g) \setminus \{\{i,j\}\}) \cup \{\{i,i'\},\{j,j'\}\}$. Then g' matches more nodes in I_0 than g does, a contradiction with the assumption that g is in $\mathcal{G}^*$. □

Claim 6. *There exist a matching $g^* \in \mathcal{G}^*$ and an injective function $\gamma : \mathit{Free}(g^*) \mapsto Q_{CC}(g^*)$ such that for every $i \in \mathit{Free}(g^*)$ there is an edge in E connecting i to an endpoint of the edge $\gamma(i)$.*

Proof. Recall that for every $j \in C_0$ we have $x_j^{f^*} < 1$ and, by Claims 1 and 2, the nodes in I_0 distribute their unit of value among the nodes in C_0 only. This implies that for every subset $H \subseteq I_0$ we have $|\mathcal{D}_{G_Z}(H)| \geq |H|$, as otherwise at least one node j in $\mathcal{D}_{G_Z}(H)$ would have $x_j^{f^*} > 1$. Hence, by Theorem 5.2 there is a matching $p : I_0 \mapsto C_0$ in the subgraph G_Z. For every $H \subseteq I_0$, let $p_H : H \mapsto C_0$ be the restriction of p to the domain H, i.e., $p_H(i) = p(i)$ for every $i \in H$; denote the image of p_H by $\mathit{Im}(p_H)$. For every matching $g \in \mathcal{G}^*$, let $T(g) = \mathit{Im}(p_{\mathit{Free}(g)}) \cap C_0(g)$. In order to construct g^*, we start from an arbitrary matching $g^0 \in \mathcal{G}^*$ and produce a finite sequence of matchings $g^0, g^1, g^2, \ldots, g^\tau$ in $\mathcal{G}^*$, with the property that $T(g^\tau) = \emptyset$ and $|T(g^{k+1})| < |T(g^k)|$ for every $k \in \{0,1,\ldots,\tau-1\}$. Note that for every $i \in \mathit{Free}(g^\tau)$, the node $p(i)$ is an endpoint of an edge in $Q_{CC}(g^\tau)$. Consider a function $\gamma : \mathit{Free}(g^\tau) \mapsto Q_{CC}(g^\tau)$ that maps a player $i \in \mathit{Free}(g^\tau)$ to the edge in $Q_{CC}(g^\tau)$ with endpoint $p(i)$. Since p is a matching in G_Z, for every $i \in \mathit{Free}(g^\tau)$ we have $\{i, p(i)\} \in E$. To see that γ is injective, we note that for every pair $i, j \in \mathit{Free}(g^\tau)$ with $i \neq j$ it holds that $p(i) \neq p(j)$ (because p is a matching) and $\{p(i), p(j)\} \notin Q_{CC}(g^\tau)$ (by Claim 5). Hence, the claim follows by setting $g^* = g^\tau$.

It remains to show how to obtain the sequence of matchings $g^0, g^1, g^2, \ldots, g^\tau$. Initially, we set $k = 0$. We then repeat the following step until $T(g^k) = \emptyset$, in which case we set $\tau = k$ and terminate. Given a matching g^k with $|T(g^k)| > 0$, we build the matching g^{k+1} in the following way. There exists a node $u \in \mathit{Free}(g^k)$ that has been matched by p to a node in $C_0(g^k)$. Let $i_1 = p(u)$ and $j_1 = g^k(i_1)$. Note that, by definition, $\{u, i_1\} \in \mathcal{F}(p)$ and $\{i_1, j_1\} \in Q_{CI}(g^k) \subseteq \mathcal{F}(g^k)$. Since $Q_{CI}(g^k)$ is finite and p and g^k are matchings, there must exist a finite sequence of edges of $Q_{CI}(g^k)$, $(\{i_1,j_1\}, \ldots, \{i_q,j_q\})$, such that $i_{r+1} = p(j_r)$ for every $r \in \{1,\ldots,q-1\}$ and $p(j_q) = v$ is an endpoint of an edge in $Q_{CC}(g^k)$. Let

$$X = \{\{i_1,j_1\},\ldots,\{i_q,j_q\}\}, \qquad Y = \{\{u,i_1\},\{j_1,i_2\},\ldots,\{j_r,i_{r+1}\},\ldots,\{j_q,v\}\};$$

we have $X \subseteq Q_{CI}(g^k) \subseteq \mathcal{F}(g^k)$ and $Y \subseteq \mathcal{F}(p)$. Define a new matching g^{k+1} so that $\mathcal{F}(g^{k+1}) = (\mathcal{F}(g^k) \setminus X) \cup (Y \setminus \{\{j_q, v\}\})$ (see Figure 2). Then $C_0(g^{k+1}) = C_0(g^k)$ and hence g^{k+1} belongs to $\mathcal{G}^*$. On the other hand, $\mathit{Free}(g^{k+1}) = (\mathit{Free}(g^k) \setminus \{u\}) \cup \{j_q\}$. Since $p(j_q) = v$, and v is an endpoint of an edge in $Q_{CC}(g^{k+1})$, we have $T(g^{k+1}) = T(g^k) \setminus \{u\}$ and hence $|T(g^{k+1})| = |T(g^k)| - 1$. We can now increment k and start the next iteration. □

In the remainder of the proof, we fix a matching g^* that satisfies the conditions of Claim 6, and use the following notation (see Figure 3).

$$Q_{CC}(g^*) = \{\{u_1,s_1\},\{u_2,s_2\},\ldots,\{u_t,s_t\}\}, \qquad Q_{CI}(g^*) = \{\{x_1,y_1\},\{x_2,y_2\},\ldots,\{x_q,y_q\}\},$$

where $s_i = g^*(u_i)$ and $u_i = g^*(s_i)$ for every $i \in \{1,2,\ldots,t\}$, and $y_i = g^*(x_i)$ for every $i \in \{1,2,\ldots,q\}$. Then we have

$$C_0(g^*) = \{x_1,x_2,\ldots,x_q\} \subseteq C_0, \quad I_0(g^*) = \{y_1,y_2,\ldots,y_q\} \subseteq I_0.$$

Also, let

$$\mathit{Free}(g^*) = \{w_1,w_2,\ldots,w_\ell\}.$$

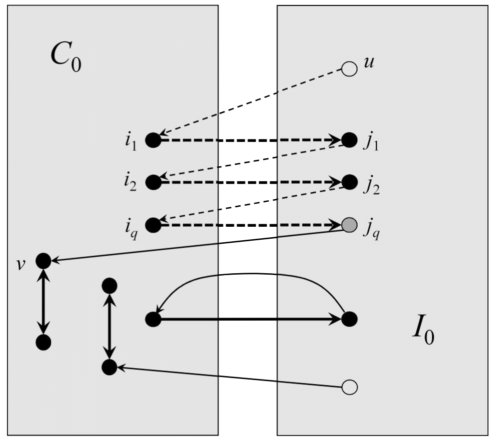

Figure 2: Proof of Claim 6. The thick edges (both solid and dashed) belong to the matching g^k; the thin edges (both solid and dashed) belong to the matching p. The matching g^{k+1} is obtained from the matching g^k by replacing the thick dashed edges with the thin dashed edges.

Let $Im(\gamma)$ be the image of γ, i.e., $Im(\gamma) = \{\gamma(i) \in Q_{CC}(g^*) \mid i \in \mathit{Free}(g^*)\}$. Without loss of generality we assume that $Im(\gamma)$ is given by the first ℓ elements of $Q_{CC}(g^*)$, i.e., $Im(\gamma) = \{\{u_1, s_1\}, \{u_2, s_2\}, \ldots, \{u_\ell, s_\ell\}\}$, where $\{u_i, s_i\} = \gamma(w_i)$ and $\{w_i, u_i\} \in E$ for every $i \in \{1, 2, \ldots, \ell\}$. Finally, let us denote by $S(g^*)$ the set $\{s_1, s_2, \ldots, s_\ell\}$, and let $\overline{Q}_{CC}(g^*) = Q_{CC}(g^*) \setminus Im(\gamma)$. We are ready to state some properties of the nodes in Z.

Claim 7. *For every node $w_i \in \mathit{Free}(g^*)$ it holds that there is no edge in E connecting w_i to a node in $S(g^*) \setminus \{s_i\}$.*

Proof. Assume for the sake of contradiction that there is a node $w_i \in \mathit{Free}(g^*)$ such that $\{w_i, s_j\} \in E$ for some $j \in \{1, 2, \ldots, \ell\}$ and $i \neq j$. We know that $\{u_j, s_j\} \in \mathcal{F}(g^*)$. Since $\{w_i, s_j\} \in E$, we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_j, s_j\}\}) \cup \{\{w_i, s_j\}, \{w_j, u_j\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$. $\square$

Recall that $C = C^{\geq 1} \cup C_0$ and $I = I^{\geq 1} \cup I_0$. The proofs of Claims 8 and 11 use the fact that for each $k \geq 1$ and each node $j \in C_k$ we have $x_j^{f^*} = \frac{|I_k|}{|C_k|}$ (see equation (9)) and, on the other hand, $x_j^{f^*} \geq 1$, which means that $|I_k| \geq |C_k|$ and, consequently, $|I^{\geq 1}| \geq |C^{\geq 1}|$.

Claim 8. *Suppose that $|\mathit{Free}(g^*)| \geq 2$. If the utility of a node $w_i \in \mathit{Free}(g^*)$ in $\mathcal{P}$ is at least $1/2$ then w_i forms a triangle with $\{u_i, s_i\}$ or with an edge in $\overline{Q}_{CC}(g^*)$.*

Proof. Consider a player $w_i \in \mathit{Free}(g^*)$ whose utility in $\mathcal{P}$ is at least $1/2$. In the coalition P player w_i is connected to at least $\frac{1}{2}|P \setminus \{w_i\}|$ nodes. Since w_i belongs to $I_0 \subseteq I$, it can only be connected to nodes in C. This implies that, since $|C^{\geq 1}| \leq |I^{\geq 1}|$, even if w_i is connected to all the nodes in $C^{\geq 1}$, it must be connected to at least half of the nodes in $Z \setminus \{w_i\}$. In particular, since $|C_0| = |Q_{CI}(g^*)| + 2|Q_{CC}(g^*)|$ and $|I_0 \setminus \{w_i\}| = |Q_{CI}(g^*)| + |\mathit{Free}(g^*)| - 1$, this implies that w_i must be connected to at least

$$|Q_{CI}(g^*)| + |Q_{CC}(g^*)| + \left\lceil \frac{1}{2} \left(|\mathit{Free}(g^*)| - 1\right) \right\rceil > |Q_{CI}(g^*)| + |Q_{CC}(g^*)|$$

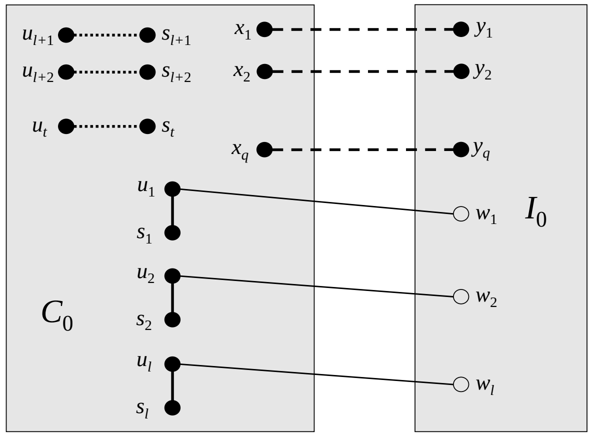

Figure 3: The nodes in $\mathit{Free}(g^*)$ are white; all other nodes are black. Thick dashed lines indicate edges in $Q_{CI}(g^*)$. Thick solid lines indicate edges in $Im(\gamma)$. Thick dotted lines indicate edges in $\overline{Q}_{CC}(g^*)$. Each node $w_i \in \mathit{Free}(g^*)$ is connected to an endpoint of the edge $\gamma(w_i)$ by a thin solid line.

nodes in Z. Within Z, w_i can be connected only to nodes in C_0, and C_0 consists of $|Q_{CI}(g^*)|$ nodes in $C_0(g^*)$ and the endpoints of the edges in $Q_{CC}(g^*)$. This implies that w_i must form a triangle with some edge in $Q_{CC}(g^*)$. By Claim 7 the node w_i cannot be connected to any node in $S(g^*) \setminus \{s_i\}$, i.e., w_i cannot form a triangle with any edge in $Q_{CC}(g^*) \setminus \overline{Q}_{CC}(g^*)$ other than $\{u_i, s_i\}$. Hence, if w_i does not form a triangle with $\{u_i, s_i\}$, it has to form a triangle with an edge in $\overline{Q}_{CC}(g^*)$. ☐

The argument in the proof of Claim 8 goes through for the case $|\mathit{Free}(g^*)| = 1$ as long as we assume that in P the utility of the unique node in $\mathit{Free}(g^*)$ is strictly greater than $1/2$.

Claim 9. *Suppose that $|\mathit{Free}(g^*)| = 1$. If the utility of the unique node $w_1 \in \mathit{Free}(g^*)$ in $\mathcal{P}$ is strictly greater than $1/2$ then w_1 forms a triangle with $\{u_1, s_1\}$ or with an edge in $\overline{Q}_{CC}(g^*)$.*

Claim 10. *The set $S(g^*)$ is an independent set in G_P.*

Proof. Assume for the sake of contradiction that there is an edge $\{s_i, s_j\} \in E$ for some $i, j \in \{1, 2, \dots, \ell\}$ such that $i \neq j$. We know that $\{u_i, s_i\} \in \mathcal{F}(g^*)$ and $\{u_j, s_j\} \in \mathcal{F}(g^*)$. Since $\{s_i, s_j\} \in E$, we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_i, s_i\}, \{u_j, s_j\}\}) \cup \{\{s_i, s_j\}, \{w_i, u_i\}, \{w_j, u_j\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$. ☐

Claim 11. *Suppose that $|\mathit{Free}(g^*)| \geq 2$. If the utility of a node $s_i \in S(g^*)$ in $\mathcal{P}$ is at least $1/2$ then s_i forms a triangle with $\{w_i, u_i\}$ or with an edge in $\overline{Q}_{CC}(g^*) \cup Q_{CI}(g^*)$.*

Proof. Consider a player $s_i \in S(g^*)$ whose utility in $\mathcal{P}$ is at least $1/2$. In the coalition P player s_i is connected to at least $\frac{1}{2}|P \setminus \{s_i\}|$ nodes. Note that s_i belongs to the subset C_0 of the minimum vertex cover C, so by Claim 1 it cannot be connected to nodes in $I^{\geq 1}$. This implies that, since $|C^{\geq 1}| \leq |I^{\geq 1}|$, even if s_i is connected to all the nodes in $C^{\geq 1}$, it must be connected to at least half of the nodes in $Z \setminus \{s_i\}$. In particular, since $|C_0 \setminus \{s_i\}| = |Q_{CI}(g^*)| + 2|Q_{CC}(g^*)| - 1$ and $|I_0| = |Q_{CI}(g^*)| + |\mathit{Free}(g^*)|$, this implies that s_i must be connected to at least

$$|Q_{CI}(g^*)| + |Q_{CC}(g^*)| + \left\lceil \frac{1}{2}\left(|\mathit{Free}(g^*)| - 1\right) \right\rceil > |Q_{CI}(g^*)| + |Q_{CC}(g^*)|$$

nodes in Z. By Claim 7 the node s_i is not connected to any node in $\mathit{Free}(g^*) \setminus \{w_i\}$. Therefore, within Z, s_i can be connected only to w_i, the endpoints of the edges in $Q_{CI}(g^*)$ and the endpoints of the edges in $Q_{CC}(g^*)$. This implies that s_i must form a triangle with $\{w_i, u_i\}$, some edge in $Q_{CI}(g^*)$ or some edge in $Q_{CC}(g^*)$. By Claim 10 the node s_i is not connected to any node in $S(g^*)$, i.e., s_i cannot form a triangle with any edge in $Q_{CC}(g^*) \setminus \overline{Q}_{CC}(g^*)$. Hence, if s_i does not form a triangle with $\{u_i, w_i\}$ then it forms a triangle with an edge in $\overline{Q}_{CC}(g^*) \cup Q_{CI}(g^*)$. $\square$

The argument in the proof of Claim 11 goes through for the case $|\mathit{Free}(g^*)| = 1$ as long as we assume that in P the utility of the unique node in $S(g^*)$ is strictly greater than $1/2$.

Claim 12. *Suppose that $|\mathit{Free}(g^*)| = 1$. If the utility of the unique node $s_1 \in S(g^*)$ in $\mathcal{P}$ is strictly greater than $1/2$ then s_1 forms a triangle with $\{w_1, u_1\}$ or with an edge in $\overline{Q}_{CC}(g^*) \cup Q_{CI}(g^*)$.*

Claim 13. *For every edge $\{u_i, s_i\} \in \overline{Q}_{CC}(g^*)$ there is at most one value $h \in \{1, \ldots, \ell\}$ such that w_h or s_h forms a triangle with $\{u_i, s_i\}$.*

Proof. Assume for the sake of contradiction that there is an edge $\{u_i, s_i\} \in \overline{Q}_{CC}(g^*)$ forming a triangle with $w_k \in \mathit{Free}(g^*)$ and with $w_h \in \mathit{Free}(g^*)$, where $k \neq h$. We know that the edge $\{u_i, s_i\}$ belongs to $\mathcal{F}(g^*)$. By our assumption, E contains edges $\{w_k, u_i\}, \{w_k, s_i\}, \{w_h, u_i\}, \{w_h, s_i\}$, so we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_i, s_i\}\}) \cup \{\{w_k, u_i\}, \{w_h, s_i\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$.

Now, let us assume that there is an edge $\{u_i, s_i\} \in \overline{Q}_{CC}(g^*)$ forming a triangle with $w_k \in \mathit{Free}(g^*)$ and with $s_h \in S(g^*)$, where $k \neq h$. We know that the two edges $\{u_i, s_i\}, \{u_h, s_h\}$ belong to $\mathcal{F}(g^*)$. By our assumption, E contains edges $\{w_k, u_i\}, \{w_k, s_i\}, \{s_h, u_i\}, \{s_h, s_i\}$, so we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_i, s_i\}, \{u_h, s_h\}\}) \cup \{\{s_h, u_i\}, \{w_k, s_i\}, \{w_h, u_h\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$.

Finally, let us assume that there is an edge $\{u_i, s_i\} \in \overline{Q}_{CC}(g^*)$ forming a triangle with $s_k \in S(g^*)$ and with $s_h \in S(g^*)$, where $k \neq h$. We know that all the edges $\{u_i, s_i\}, \{u_k, s_k\}, \{u_h, s_h\}$ belong to $\mathcal{F}(g^*)$. By our assumption, E contains edges $\{s_k, u_i\}, \{s_k, s_i\}, \{s_h, u_i\}, \{s_h, s_i\}$, so we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_i, s_i\}, \{u_k, s_k\}, \{u_h, s_h\}\}) \cup \{\{s_k, s_i\}, \{s_h, u_i\}, \{w_k, u_k\}, \{w_h, u_h\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$. $\square$

Claim 14. *For every edge $\{x_i, y_i\} \in Q_{CI}(g^*)$ there is at most one node in $S(g^*)$ that forms a triangle with $\{x_i, y_i\}$.*

Proof. Assume for the sake of contradiction that there is an edge $\{x_i, y_i\} \in Q_{CI}(g^*)$ forming a triangle with $s_k \in S(g^*)$ and with $s_h \in S(g^*)$, where $k \neq h$. We know that the two edges $\{u_k, s_k\}, \{u_h, s_h\}$ belong to $\mathcal{F}(g^*)$. By our assumption, E contains edges $\{x_i, s_k\}, \{x_i, s_h\}, \{y_i, s_k\}, \{y_i, s_h\}$, so we can define a new function g' such that $\mathcal{F}(g') = (\mathcal{F}(g^*) \setminus \{\{u_k, s_k\}, \{u_h, s_h\}\}) \cup \{\{x_i, s_h\}, \{y_i, s_k\}, \{w_h, u_h\}, \{w_k, u_k\}\}$. Then g' matches more nodes in I_0 than g^* does, a contradiction with the fact that g^* is in $\mathcal{G}^*$. $\square$

We can now conclude the proof of Theorem 5.5 by describing how to partition the set Z into stars and triangles, so that the players in these coalitions (strictly) prefer them to their coalitions in $\mathcal{P}$. We distinguish among three cases: $|\mathit{Free}(g^*)| = 0$, $|\mathit{Free}(g^*)| \geq 2$ and $|\mathit{Free}(g^*)| = 1$. For readability, given a player $z \in Z$, we denote her utility in $\mathcal{P}$ by $v(z)$.

$|\boldsymbol{Free(g^*)}| = \mathbf{0}$ For every edge $\{u_j, s_j\}$ in $Q_{CC}(g^*)$ and for every edge $\{x_2, y_2\}$ in $Q_{CI}(g^*)$ we construct a 1-star coalition and add it to $\mathcal{S}$. In the resulting partition the utility of each node in Z is 1. Thus, each player weakly prefers $\mathcal{S}$ to $\mathcal{P}$.

As argued after Claim 3, it remains to show that if $Z = P$ then some player in Z strictly prefers $\mathcal{S}$ to $\mathcal{P}$. To see this, it suffices to observe that it cannot be the case that $v(z) = 1$ for all $z \in P$, as this could only happen if P was a clique, and we assume that this is not the case.

$|\boldsymbol{Free(g^*)}| \geq \mathbf{2}$ Let us consider the tuples $\langle w_i, u_i, s_i \rangle$ for every $i \in \{1, 2, \ldots, \ell\}$. If $v(w_i) < 1/2$, $v(s_i) < 1/2$, we construct the 2-star coalition $\{w_i, u_i, s_i\}$ with center u_i and add it to $\mathcal{S}$. In this case the new utility of u_i is 1, hence at least as high as its utility in $\mathcal{P}$, while the utilities of w_i and s_i are $1/2$, which is strictly higher that their utility in $\mathcal{P}$. On the other hand, if $v(w_i) \geq 1/2$ or $v(s_i) \geq 1/2$, we construct a triangle coalition and possibly a 1-star coalition to add to $\mathcal{S}$ in the following way.

If $v(w_i) \geq 1/2$ then we construct a triangle coalition induced by w_i and the endpoints of an edge in $\overline{Q}_{CC}(g^*) \cup \{\{u_i, s_i\}\}$ with which w_i forms a triangle and whose existence is guaranteed by Claim 8. Furthermore, if $\{u_i, s_i\}$ has not been employed in a triangle coalition with w_i, we construct a 1-star coalition $\{u_i, s_i\}$.

Otherwise, we have $v(w_i) < 1/2$ and $v(s_i) \geq 1/2$. We then construct a triangle coalition induced by s_i and the endpoints of an edge in $\overline{Q}_{CC}(g^*) \cup Q_{CI}(g^*) \cup \{\{w_i, u_i\}\}$ with which s_i forms a triangle and whose existence is guaranteed by Claim 11. Furthermore, if $\{w_i, u_i\}$ has not been employed in a triangle coalition with s_i, we construct a 1-star coalition $\{w_i, u_i\}$.

Note that at every step $i \in \{1, 2, \ldots, \ell\}$ of the procedure described in the previous two paragraphs, we are able to construct a triangle when needed because Claims 13 and 14 guarantee that there is no conflict in picking edges from $\overline{Q}_{CC}(g^*)$ or $Q_{CI}(g^*)$.

Finally, for every edge in $Q_{CI}(g^*)$ and $\overline{Q}_{CC}(g^*)$ that has not been used to construct a triangle in the previous steps, we construct a 1-star coalition and add it to $\mathcal{S}$.

In the resulting partition $\mathcal{S}$, every player in Z whose utility in $\mathcal{P}$ was strictly less than $1/2$ achieves utility at least $1/2$; for every other player in Z its utility in $\mathcal{S}$ is 1. Thus, this partitioning of Z into stars and triangles does not lower the utility of any node in Z.

Now, suppose that $Z = P$; we will argue that in this case some player strictly prefers $\mathcal{S}$ to $\mathcal{P}$. If some player in P has utility $1/2$ in $\mathcal{S}$ then this player strictly prefers $\mathcal{S}$ to $\mathcal{P}$. Otherwise, every player in P obtains utility 1 in $\mathcal{S}$. In this case at least one player strictly prefers $\mathcal{S}$ to $\mathcal{P}$ since we assume that P is not a clique.

$|\boldsymbol{Free}(g^*)| = 1$ Let us consider the tuple $\langle w_1, u_1, s_1 \rangle$. If $v(w_1) \leq 1/2$, $v(s_1) \leq 1/2$, we construct the 2-star coalition $\{w_1, u_1, s_1\}$ with center u_1 and add it to $\mathcal{S}$. In this case the new utility of u_1 is 1, hence at least as high as its utility in $\mathcal{P}$, while the utilities of w_1 and s_1 are $1/2$, which is at least as high as their utility in $\mathcal{P}$. On the other hand, if $v(w_1) > 1/2$ or $v(s_1) > 1/2$, we construct a triangle coalition and possibly a 1-star coalition to add to $\mathcal{S}$ in the following way.

If $v(w_1) > 1/2$ then we construct a triangle coalition induced by w_1 and the endpoints of an edge in $\overline{Q}_{CC}(g^*) \cup \{\{u_1, s_1\}\}$ with which w_1 forms a triangle and whose existence is guaranteed by Claim 9. Furthermore, if $\{u_1, s_1\}$ has not been employed in a triangle coalition with w_1, we construct a 1-star coalition $\{u_1, s_1\}$.

Otherwise, $v(w_1) \leq 1/2$ and $v(s_1) > 1/2$. In this case we construct a triangle coalition induced by s_1 and the endpoints of an edge in $Q'_{CC}(g^*) \cup Q_{CI}(g^*) \cup \{\{w_1, u_1\}\}$ with which s_1 forms a triangle and whose existence is guaranteed by Claim 12. Furthermore, if $\{w_1, u_1\}$ has not been employed in a triangle coalition with s_1, we construct a 1-star coalition $\{w_1, u_1\}$.

Finally, for every edge in $Q_{CI}(g^*)$ and $\overline{Q}_{CC}(g^*)$ that has not been used to construct a triangle in the previous steps, we construct a 1-star coalition and add it to $\mathcal{S}$.

In the resulting partition $\mathcal{S}$ the utility of each node in $Z \setminus \{w_1, s_1\}$ is 1; the utility of w_1 and s_1 is 1 unless $v(w_1) \leq 1/2$, $v(s_1) \leq 1/2$, in which case the utility of each of these nodes in $\mathcal{S}$ is equal to $1/2$. Thus, each player weakly prefers $\mathcal{S}$ to $\mathcal{P}$.

Now, suppose that $Z = P$; we will argue that in this case some player strictly prefers $\mathcal{S}$ to $\mathcal{P}$. If in $\mathcal{S}$ the utility of every player in P is 1, at least one player strictly prefers $\mathcal{S}$ to $\mathcal{P}$ since we assume that P is not a clique. Moreover, if $v(w_1) < 1/2$ then w_1 strictly prefers $\mathcal{S}$ to $\mathcal{P}$, if $v(s_1) < 1/2$ then s_1 strictly prefers $\mathcal{S}$ to $\mathcal{P}$, and if $v(z) < 1$ for some $z \in P \setminus \{w_1, s_1\}$ then z strictly prefers $\mathcal{S}$ to $\mathcal{P}$. Thus, it remains to consider the case $v(s_1) = 1/2$, $v(w_1) = 1/2$, $v(z) = 1$ for all $z \in P \setminus \{w_1, s_1\}$. If $v(z) = 1$ for all $z \in P \setminus \{w_1, s_1\}$ then s_1 is adjacent in G to all nodes in $P \setminus \{w_1, s_1\}$. Therefore, $v(s_1) = 1/2$ implies $\frac{|P|-2}{|P|-1} = \frac{1}{2}$ and hence $|P| = 3$. But this is only possible if P is a star with center u_1 and leaves s_1 and w_1, a contradiction with our assumption that P is not a star.

□

The next proposition identifies additional structural properties of Pareto optimal partitions. Given a Pareto optimal partition $\mathcal{P}$ of N, we say that a node $i \in N$ is a *critical node* in $\mathcal{P}$ if it either forms a singleton coalition or is a leaf of a nondegenerate star.

Proposition 5.6. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. If $\mathcal{P}$ is a Pareto optimal partition for the modified fractional hedonic game $\mathcal{MF}(G)$ and i and j are critical nodes in $\mathcal{P}$ then $\{i, j\} \notin E$.*

Proof. Let i and j be two critical nodes in $\mathcal{P}$. Assume for the sake of contradiction that $\{i, j\} \in E$. By definition of a critical node we have $\mathcal{P}(i) \neq \mathcal{P}(j)$. Let $R = \{i, j\}$, $P_1 = \mathcal{P}(i) \setminus \{i\}$ and $P_2 = \mathcal{P}(j) \setminus \{j\}$. Note that if $\mathcal{P}(i)$ is a singleton coalition, then $P_1 = \emptyset$ and otherwise $|P_1| \geq 2$; similarly, if $\mathcal{P}(j)$ is a singleton coalition, then $P_2 = \emptyset$ and otherwise $|P_2| \geq 2$. Consider the partition $\mathcal{S} = (\mathcal{P} \setminus \{\mathcal{P}(i), \mathcal{P}(j)\}) \cup \{R, P_1, P_2\}$. We claim that $\mathcal{S}$ Pareto dominates $\mathcal{P}$. Indeed, for every $z \notin \mathcal{P}(i) \cup \mathcal{P}(j)$ we have $v_z(\mathcal{P}) = v_z(\mathcal{S})$. Now, consider an $x \in \{i, j\}$. If $\mathcal{P}(x)$ is a singleton coalition, then we have $0 = v_x(\mathcal{P}) < v_x(\mathcal{S}) = 1$. On the other hand, if $\mathcal{P}(x)$ is a nondegenerate star with center r, then trivially $v_r(\mathcal{P}) = v_r(\mathcal{S}) = 1$, $v_x(\mathcal{P}) \leq 1/2 < v_x(\mathcal{S}) = 1$, while for every $z \in \mathcal{P}(x) \setminus \{r, x\}$ we have $v_z(\mathcal{P}) < v_z(\mathcal{S})$. □

We remark that neither Theorem 5.5 nor Proposition 5.6 fully characterize Pareto optimal partitions: there are partitions that consist of singletons, stars and cliques, and such that no two critical nodes are connected, but that are nevertheless not Pareto optimal. For instance, consider a symmetric unweighted graph that consists of a triangle with nodes a, b, c and a node d that is connected to a by an edge. Then partition $\mathcal{P} = \{\{a, b, c\}, \{d\}\}$ consists of a triangle coalition and a singleton coalition, and no edge of the graph connects two nodes that are critical with respect to $\mathcal{P}$, but nevertheless $\mathcal{P}$ is Pareto dominated by $\{\{a, d\}, \{b, c\}\}$.

Another important observation is that Proposition 4.3 directly extends to modified fractional hedonic games; we omit the proof, as it is identical to the proof of Proposition 4.3.

Proposition 5.7. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$, and let $\mathcal{P}$ be a Pareto optimal partition for $\mathcal{MF}(G)$. Then*

(a) every coalition in $\mathcal{P}$ is connected,

(b) the set of players in singleton coalitions of $\mathcal{P}$ forms an independent set in G,

(c) if $E \neq \emptyset$, then $\mathcal{P}$ contains at least one non-singleton coalition.

We can now establish the following upper bound on the Price of Pareto Optimality.

Theorem 5.8. *Let $G = (N, E)$ be a symmetric unweighted graph with $|N| \geq 2$. Then $\mathsf{PPO}(\mathcal{MF}(G)) \leq 2$.*

Proof. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ be a Pareto optimal partition with m coalitions. By Theorem 5.5 $\mathcal{P}$ consists of singletons, stars and cliques. Let Q_1 be the set of indices of the nondegenerate stars in $\mathcal{P}$, let Q_2 be the set of indices of the 1-stars in $\mathcal{P}$, and let Q_3 be the set of indices of the cliques in $\mathcal{P}$. Let $Q = Q_1 \cup Q_2 \cup Q_3$. By Proposition 5.7 we have $|Q| \geq 1$. Note that the number of singleton coalitions is $m - |Q|$.

We will now define a vertex cover C for G and relate its size to the social welfare of $\mathcal{P}$. For every $k \in Q_1$, let r_k be the center of the nondegenerate star P_k. For every $k \in Q$, define the set C_k as follows: if $k \in Q_1$ then $C_k = \{r_k\}$, if $k \in Q_2 \cup Q_3$ then $C_k = P_k$. Let

$$C_{Q_1} = \bigcup_{k \in Q_1} C_k, \qquad C_{Q_2} = \bigcup_{k \in Q_2} C_k, \qquad C_{Q_3} = \bigcup_{k \in Q_3} C_k, \qquad C = C'_{Q_1} \bigcup C'_{Q_2} \bigcup C'_{Q_3}.$$

Since $|Q| \geq 1$, the set C is non-empty. Observe that N can be partitioned into two sets: C and the set of critical nodes. By Proposition 5.6 there is no edge between any pair of critical nodes, so C is a cover for G. Using observations in Example 5.1, we get

$$SW(\mathcal{P}) = \sum_{k \in Q_1} V(P_k) + \sum_{k \in Q_2} V(P_k) + \sum_{k \in Q_3} V(P_k) = 2|Q_1| + 2|Q_2| + \sum_{k \in Q_3} |P_k| \geq |C_{Q_1}| + |C_{Q_2}| + |C_{Q_3}| = |C|.$$

Combining this bound and Proposition 5.4, we obtain $\mathsf{PPO}(\mathcal{MF}(G)) \leq 2$. □

For symmetric unweighted bipartite graphs, we can show a stronger result: every Pareto optimal partition maximizes the social welfare.[2]

[2] After the publication of the preliminary version of this work, Bullinger [2020] extended this result to general symmetric unweighted graphs.

Theorem 5.9. *Let $G = (N, E)$ be an symmetric unweighted bipartite graph with $|N| \geq 2$. Then* $\mathsf{PPO}(\mathcal{MF}(G)) \leq 1$.

Proof. Let $\mathcal{P} = \{P_1, P_2, \dots, P_m\}$ be a Pareto optimal partition with m coalitions. Since G is bipartite, Theorem 5.5 implies that $\mathcal{P}$ consists of singletons and stars. We can assume without loss of generality that the coalitions in $\mathcal{P}$ are ordered so that $P_1, \dots, P_\ell$ are nondegenerate stars, $P_{\ell+1}, \dots, P_r$ are 1-stars, and all remaining coalitions are singletons, for some $0 \leq \ell \leq r \leq m$. Since the value of every star coalition is 2, we have $SW(\mathcal{P}) = 2r$. We will now argue that G admits a vertex cover C of size r; from this, together with Proposition 5.4, claiming that the social welfare of the optimal partition is at most $2r$, we obtain $\mathsf{PPO}(\mathcal{MF}(G)) \leq 1$.

We construct the vertex cover C as follows. First, we include the centers of the ℓ nondegenerate stars. This covers all edges within these stars. Moreover, by Proposition 5.6 we know that there is no edge between any pair of critical nodes (recall that a critical node either forms a singleton coalition or is the leaf of a nondegenerate star). Thus, all remaining edges to cover are incident with nodes that belong to 1-stars. We will show that it is possible to select a set C' of $r - \ell$ nodes, one for each 1-star, to cover these remaining edges[3]. Finally, by adding C' to the centers of all nondegenerate stars, we obtain a vertex cover C of size r.

We now show how to construct C'. For every $j = \ell + 1, \dots, r$, let $P_j = \{x_j, y_j\}$, where the nodes in $X = \{x_{\ell+1}, \dots, x_r\}$ belong to one side of the bipartition of G, while the nodes in $Y = \{y_{\ell+1}, \dots, y_r\}$ belong to its other side. Let Z be the set of critical nodes with respect to $\mathcal{P}$. Consider the graph H induced by the set of nodes $X \cup Y \cup Z$. Let us consider the matching $M : X \mapsto Y$ which maps x_j to y_j for every $j = \ell + 1, \dots, r$. We claim that M is a maximum matching in H. As H is a bipartite graph, by Theorem 5.3 it follows that it has a vertex cover C' of size $|\mathcal{F}(M)| = r - \ell$.

To conclude the proof, it remains to show that M is a maximum matching in H. To this aim, assume for the sake of contradiction that H admits another matching M' with $|\mathcal{F}(M')| > |\mathcal{F}(M)|$. Then these exists an augmenting path in H, i.e., a sequence of nodes $z_1, \dots, z_t$, where t is an odd integer, such that $\{z_i, z_{i+1}\} \in \mathcal{F}(M') \setminus \mathcal{F}(M)$ if i is odd, $\{z_i, z_{i+1}\} \in \mathcal{F}(M) \setminus \mathcal{F}(M')$ if i is even, and neither z_1 nor z_t is matched by M. It follows that z_1 and z_t are critical nodes. Moreover, they cannot belong to the same nondegenerate star, since t is odd, and G does not have odd-length cycles. Given the augmenting path, we can modify $\mathcal{P}$ as follows. First, we remove coalitions $\{z_2, z_3\}, \{z_4, z_5\}, \dots, \{z_{t-2}, z_{t-1}\}$ from $\mathcal{P}$. Then if z_1 is a leaf of a nondegenerate star P, we replace P with $P \setminus \{z_1\}$, and if $\{z_1\}$ is a singleton coalition in $\mathcal{P}$, we remove $\{z_1\}$ from $\mathcal{P}$. Similarly, if z_t is a leaf of a nondegenerate star P', we replace P' with $P' \setminus \{z_t\}$, and if $\{z_t\}$ is a singleton coalition in $\mathcal{P}$, we remove $\{z_t\}$ from $\mathcal{P}$. Finally, we add coalitions $\{z_1, z_2\}, \{z_3, z_4\}, \dots, \{z_{t-1}, z_t\}$. In the resulting partition, the utility of every player in $\{z_1, \dots, z_t\}$ is 1. Hence, each of the players $z_2, \dots, z_{i-1}$ has the same utility as in $\mathcal{P}$ and z_1 and z_t have strictly higher utility than in $\mathcal{P}$. Furthermore, if in $\mathcal{P}$ player z_1 or z_t is a leaf of a nondegenerate star, the utility of other leaves in that coalition increases, whereas the utility of the center remains the same (recall that z_1 and z_t cannot belong to the same nondegenerate star). We conclude that the new partition Pareto dominates $\mathcal{P}$, a contradiction with the assumption that $\mathcal{P}$ is Pareto optimal. Thus, it follows that M is a maximum matching in H. $\square$

6 Conclusion

We have introduced the notion of Price of Pareto Optimality (PPO) and obtained upper and lower bounds on the PPO in three classes of hedonic games. We provide a summary of our results in Table 1.

6.1 Summary of Results

For ASHGs, the key feature of the game is whether players may have negative utilities for each other: if this is not the case, every Pareto optimal outcome maximizes the social welfare, so $\mathsf{PPO} = 1$. In the presence of negative utilities, it is still the case that $\mathsf{PPO} = 1$ if the game is symmetric and the underlying social graph is acyclic, but if either of these conditions is violated, the Price of Pareto Optimality may be $+\infty$.

For FHGs and mFHGs the Price of Pareto Optimality may be unbounded even if the game is symmetric, all players have positive utilities for each other and the underlying social network is acyclic. Thus, for these classes of games we focus on scenarios where the underlying social graph is symmetric and unweighted; under these assumptions in both

[3]We are grateful to the anonymous AI Journal reviewer who simplified our proof of this fact considerably.

	ASHG	FHG	mFHG
General graphs	$+\infty$ (Ex. 3.1)	$+\infty$ (Prop. 4.2)	$+\infty$ (Prop. 4.2)
Symmetric acyclic graphs, non-negative weights	1 (Prop. 3.3, Prop. 3.4)	$+\infty$ (Prop. 4.2)	$+\infty$ (Prop. 4.2)
Symmetric unweighted graphs	1 (Prop. 3.4)	$\geq \Delta_G - \frac{1}{\Delta_G}$ (Prop. 4.11), $\leq \min\{n, 2\Delta_G(\Delta_G+1)\}$ (Prop. 4.4, Thm. 4.5)	≤ 2 (Thm. 5.8)
Bipartite symmetric unweighted graphs	1 (Prop. 3.4)	$\geq \Delta_G - \frac{1}{\Delta_G}$ (Prop. 4.11), $\leq \min\{n, 2\Delta_G(\Delta_G+1)\}$ (Prop. 4.4, Thm. 4.5)	1 (Thm. 5.9)
Acyclic symmetric unweighted graphs	1 (Prop. 3.4)	$\geq \Delta_G - \frac{1}{\Delta_G}$ (Prop. 4.11), $\leq \Delta_G + 2$ (Thm. 4.10)	1 (Thm. 5.9)

Table 1: Summary of the results on the Price of Pareto Optimality. In the second row, the upper bound holds for graphs that are symmetric and acyclic (Prop. 3.3) *or* have non-negative weights (Prop. 3.4), whereas the lower bounds hold for graphs that are symmetric, acyclic *and* have non-negative weights (Prop. 4.2).

classes of games the Price of Pareto Optimality is trivially upper-bounded by the number of players n. It turns out that in this setting mFHGs are much more well-behaved than FHGs: for mFHGs, we show that $\mathsf{PPO} \leq 2$ for arbitrary graphs, whereas for FHGs PPO can be essentially as large as Δ_G, i.e., the maximum degree of the graph, even if the graph is acyclic; on the positive side, for FHGs we obtain $\mathsf{PPO} \leq O(\Delta_G^2)$. This difference between FHGs and mFHGs is striking, giving the similarity between their definitions, and suggests that mFHGs deserve further attention; indeed, after the conference version of our paper was published, several authors explored this class of games in more detail [Kaklamanis *et al.*, 2016; Monaco *et al.*, 2018, 2019; Bullinger, 2020].

Both for FHGs and for mFHGs we obtain improved upper bounds on PPO for special classes of graphs: for FHGs on acyclic graphs we improve the bound from $O(\Delta_G^2)$ to $\Delta_G + 2$ (and show that this bound is tight up to a small additive constant), and for mFHGs on bipartite graphs we improve the bound from 2 to 1, thereby showing that if the social network is bipartite, any Pareto optimal outcome maximizes the social welfare. These results are interesting both because they identify the features of the social network which may lead to low social welfare in Pareto optimal outcomes, and because real-life networks often have additional structure. In particular, an acyclic network models a hierarchical organizational structure (see, e.g., Demange [2004]), and a bipartite network models a setting where players belong to two types and only care about the presence of the agents of the other type (as in the Bakers and Millers game described by Aziz *et al.* [2014]).

6.2 Pareto Optimality and Other Notions of Stability

We will now discuss the relationship between Pareto optimality and other notions of stability. First, we compare Pareto optimality to two other concepts of stability that are based on resilience to group deviations, and then we compare our results for the Price of Pareto Optimality to known results on the Price of Anarchy and the Price of Stability in ASHGs, FHGs and mFHGs.

6.2.1 Pareto optimal outcomes vs. (super-)strong equilibria

Pareto optimal outcomes are similar in spirit to super-strong Nash equilibria [Rozenfeld, 2007], i.e., outcomes that do not admit group deviations that make all of the deviating players weakly better off and some of the deviating players strictly better off (some authors use the term *strong equilibrium* to refer to this solution concept, but the standard notion of strong equilibrium [Aumann, 1959] only requires resistance to deviations that make all deviators strictly better off). Following the standard terminology, we will refer to the ratio between the maximum social welfare and the social welfare in the worst super-strong Nash equilibrium as the *super-strong Price of Anarchy*.

In the context of hedonic games, strong equilibria can be mapped to partitions in the core, and super-strong equilibria can be mapped to partitions in the strict core [Aziz and Savani, 2015]: briefly, a partition $\mathcal{P} = \{P_1, \dots, P_m\}$ is in the core (respectively, strict core) if there is no group of players D such that each player $i \in D$ strictly prefers D to $\mathcal{P}(i)$ (respectively, each player $i \in D$ weakly prefers D to $\mathcal{P}(i)$ and some player $i \in D$ strictly prefers D to $\mathcal{P}(i)$).

It is immediate that every super-strong Nash equilibrium (respectively, strict core outcome) is Pareto optimal, as otherwise the grand coalition would have a beneficial deviation. However, the converse is not true: a Pareto optimal outcome may admit a group deviation that is beneficial to the deviating players, but negatively affects some of the non-deviators. Indeed, while all games have Pareto optimal outcomes, there are many interesting games with no strong Nash equilibria (respectively, with an empty core).

Consider, for instance, the classic Prisoners' Dilemma game: in this game, there are two players who can each choose to cooperate (C) or defect (D); each player gets utility 2 if both of them defect, utility 1 if she cooperates and the other player defects, utility 5 if she defects and the other player cooperates, and utility 3 if both of them cooperate. It can be verified that this game has no strong Nash equilibria and hence no super-strong Nash equilibria; however, the outcome where both players cooperate is Pareto optimal, and so are the outcomes where one player cooperates and the other player defects.

Since every super-strong Nash equilibrium/strict core outcome is Pareto optimal, the social welfare in the worst Pareto optimal outcome is less than or equal to the social welfare in the worst super-strong Nash equilibrium/strict core outcome, i.e., the Price of Pareto Optimality is at least as high as the super-strong Price of Anarchy, and can be much higher. However, this comparison is somewhat unfair, as it does not take into account that the set of super-strong equilibria (respectively, the strict core) may be empty.

The relationship between Pareto optimality and strong equilibria/core outcomes is more complicated. There are games where a strong equilibrium (respectively, an outcome in the core) is not Pareto optimal. Consider, for instance, an additively separable hedonic game with $N = \{1, 2, 3\}$ and $w_{1,2} = 1$, $w_{2,1} = 100$, $w_{2,3} = -100$, $w_{1,3} = w_{3,1} = w_{3,2} = 0$. The only Pareto optimal partition in this game is $\{\{1, 2\}, \{3\}\}$; this is also the partition that maximizes the social welfare. However, the grand coalition $\{1, 2, 3\}$ admits no deviation that makes all deviating players strictly better off: to see this, note that players 1 and 3 obtain their maximum possible utility in the grand coalition, and player 2 cannot profitably deviate on her own. Thus, in this game the Price of Pareto Optimality is 1, whereas there is an outcome in the core with very low social welfare.

Conversely, a Pareto optimal outcome may have a much lower social welfare than every core outcome. Indeed, consider, for instance, an unweighted symmetric FHG where the underlying graph is a $(d, 2)$-superstar, $d > 2$. In this game the grand coalition is Pareto optimal, and its social welfare is $\frac{4d}{2d+1}$ (see Example 4.1). On the other hand, a partition is in the core of this game if and only if it consists of d connected coalitions of size 2 and an isolated node (which may be the center node or one of the leaves); in any such partition the social welfare is d.

To summarize these observations, the Price of Pareto Optimality is at least as high as the super-strong Price of Anarchy, and can be higher or lower than the Strong Price of Anarchy.

6.2.2 Price of Pareto Optimality vs. Price of Anarchy/Price of Stability in hedonic games

It is also interesting to compare the Price of Pareto Optimality to the Price of Anarchy/Price of Stability in hedonic games, and in particular in ASHGs, FHGs and mFHGs.

For ASHGs, we were unable to find published results on the Price of Anarchy or Price of Stability. However, we observe that it is easy to construct an ASHG on a symmetric unweighted graph in which the Price of Anarchy is linear in the number of players n (recall that the Price of Pareto Optimality is 1 in this case). Specifically, consider a graph on the node set $\{a_1, \dots, a_s, b_1, \dots, b_s\}$, where both $\{a_1, \dots, a_s\}$ and $\{b_1, \dots, b_s\}$ form cliques and there are also edges between a_i and b_i for each $i = 1, \dots, s$. Then the partition $\{\{a_1, b_1\}, \dots, \{a_s, b_s\}\}$ is Nash stable, but its social welfare is only $2s$, whereas the optimal social welfare is $2s(s-1) + 2s = 2s^2$ (this social welfare can be attained when the players form the grand coalition). Conversely, there are ASHGs where $\mathsf{PPO} = +\infty$, but the Price of Anarchy equals 1. For instance, set $N = \{1, 2, 3\}$, $w_{1,2} = w_{2,1} = w_{2,3} = 1$, $w_{3,2} = -10$. The unique Nash stable partition in this game is $\{\{1, 2\}, \{3\}\}$ and its social welfare is 2, which is the maximum possible social welfare in this game; on the other hand, the grand coalition is Pareto optimal (as any other partition would provide lower utility to player 2), but has negative social welfare. Thus, in general ASHGs the Price of Pareto Optimality and the Price of Anarchy are incomparable. Our second example in this paragraph also shows that the Price of Pareto Optimality can be much higher than the Price of Stability, but we have not been able to construct an example where the converse is true.

For FHGs, Bilò *et al.* [2018] show that the Price of Anarchy may be linear in the number of players, even on symmetric unweighted acyclic graphs whose maximum degree is equal to 2, simply because players may be stuck in the grand coalition. In contrast, Pareto optimality allows group deviations, and therefore PPO is reasonably low in low-degree graphs, especially if these graphs are acyclic. Bilò *et al.* [2018] and Kaklamanis *et al.* [2016] also obtain upper and lower bounds on the Price of Stability in FHGs. These bounds are somewhat similar to our bounds on PPO in these games: on weighted graphs, the Price of Stability can grow linearly with the number of players, but in the unweighted case the Price of Stability is bounded by a small constant for bipartite graphs and is equal to 1 on trees and graphs of girth at least 5. Interestingly, in contrast to our results, Bilò *et al.* [2018] and Kaklamanis *et al.* [2016] do not provide upper or lower bounds on the Price of Stability in terms of the maximum degree.

For mFHGs, Monaco *et al.* [2018] observe that the lower bound on the Price of Anarchy for paths obtained by Bilò *et al.* [2018] in the context of FHGs applies to mFHGs as well. They also show a linear lower bound on the Price of Stability for weighted symmetric acyclic graphs. However, for symmetric unweighted graphs the Price of Stability is 1 [Kaklamanis *et al.*, 2016; Monaco *et al.*, 2018].

These results indicate that in many classes of games the Price of Pareto Optimality tends to be much lower than the Price of Anarchy, but behaves broadly similarly to the Price of Stability; there are also several examples where the known upper bounds on the Price of Stability are stronger than the known upper bounds on the Price of Pareto Optimality, and in particular where the Price of Stability is 1, but the Price of Pareto Optimality is greater than one. However, in general it is not the case that the Price of Stability is always lower than the Price of Pareto Optimality, as illustrated by the Prisoners' Dilemma; moreover, while Pareto optimal outcomes are guaranteed to exist, Nash stable outcomes may fail to exist.

6.3 Future Work

There are many open problems suggested by our work. For instance, it is not clear if the upper bound in Theorem 4.5 is tight; in fact, we do not have examples of fractional hedonic games on symmetric unweighted graphs whose PPO exceeds Δ_G. In particular, we do not know if the upper bound of $\Delta_G + 2$ can be extended from acyclic graphs to bipartite graphs, in line with the results for mFHGs. More broadly, we believe that the PPO is a useful measure, and it would be interesting to compute or bound it for other classes of (cooperative and non-cooperative) games; while first steps in this direction have been made by Balliu *et al.* [2017b], there is more to be done.

7 Acknowledgments

This work was partially supported by ANR-14-CE24-0007-01 *"CoCoRICo-CoDec"*, by European Research Council (ERC) under grant number 639945 (ACCORD), by MIUR PRIN 2017 Project 1080 ALGADIMAR "Algorithms, Games, and Digital Markets", and by COST Action IC1205 on Computational Social Choice. We would like to thank the anonymous AAAI and AI Journal reviewers for their valuable feedback.

References

Nir Andelman, Michal Feldman, and Yishay Mansour. Strong price of anarchy. In *SODA'07*, pages 189–198, 2007.

Elliot Anshelevich, Anirban Dasgupta, Jon M. Kleinberg, Éva Tardos, Tom Wexler, and Tim Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal of Computing*, 38(4):1602–1623, 2008.

Itai Ashlagi, Dov Monderer, and Moshe Tennenholtz. On the value of correlation. *Journal of Artificial Intelligence Research*, 33:575–613, 2008.

Robert J. Aumann. Acceptable points in general cooperative n-person games. In *Contributions to the Theory of Games IV(Annals of Mathematics Studies 40)*, pages 287–324. Princeton University Press, 1959.

Haris Aziz and Rahul Savani. Hedonic games. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 15. Cambridge University Press, 2015.

Haris Aziz, Felix Brandt, and Hans Georg Seedig. Stable partitions in additively separable hedonic games. In *AAMAS'11*, pages 183–190, 2011.

Haris Aziz, Felix Brandt, and Paul Harrenstein. Fractional hedonic games. In *AAMAS'14*, pages 5–12, 2014.

Haris Aziz, Serge Gaspers, Joachim Gudmundsson, Julián Mestre, and Hanjo Täubig. Welfare maximization in fractional hedonic games. In *IJCAI'15*, pages 461–467, 2015.

Haris Aziz, Florian Brandl, Felix Brandt, Paul Harrenstein, Martin Olsen, and Dominik Peters. Fractional hedonic games. *ACM Transactions on Economics and Computation*, 7(2):6:1–6:29, 2019.

Alkida Balliu, Michele Flammini, Giovanna Melideo, and Dennis Olivetti. Nash stability in social distance games. In *AAAI'17*, pages 342–348, 2017.

Alkida Balliu, Michele Flammini, and Dennis Olivetti. On Pareto optimality in social distance games. In *AAAI'17*, pages 349–355, 2017.

Suryapratim Banerjee, Hideo Konishi, and Tayfun Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, 2001.

Vittorio Bilò, Angelo Fanelli, Michele Flammini, Gianpiero Monaco, and Luca Moscardelli. Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. *Journal of Artificial Intelligence Research*, 62:315–371, 2018.

Anna Bogomolnaia and Matthew O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

Felix Brandt, Felix A. Fischer, Paul Harrenstein, and Yoav Shoham. Ranking games. *Artificial Intelligence*, 173(2):221–239, 2009.

Simina Brânzei and Kate Larson. Social distance games. In *IJCAI'11*, pages 91–96, 2011.

Martin Bullinger. Pareto-optimality in cardinal hedonic games. In *AAMAS'20*, pages 213–221, 2020.

José R. Correa, Andreas S. Schulz, and Nicolás E. Stier Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, 2004.

Gabrielle Demange. On group stability in hierarchies and networks. *Journal of Political Economy*, 112(4):754–778, 2004.

Reinhard Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, 2005.

Jacques H. Drèze and Joseph Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica*, 48(4):987–1003, 1980.

Edith Elkind, Angelo Fanelli, and Michele Flammini. Price of Pareto optimality in hedonic games. In *AAAI'16*, pages 475–481, 2016.

Michal Feldman and Ophir Friedler. A unified framework for strong price of anarchy in clustering games. In *ICALP'15*, pages 601–613, 2015.

Moran Feldman, Liane Lewin-Eytan, and Joseph Naor. Hedonic clustering games. *ACM Transactions on Parallel Computing*, 2(1):4:1–4:48, 2015.

Michele Flammini, Gianpiero Monaco, and Qiang Zhang. Strategyproof mechanisms for additively separable hedonic games and fractional hedonic games. In *WAOA'17*, pages 301–316, 2017.

Christos Kaklamanis, Panagiotis Kanellopoulos, and Konstantinos Papaioannou. The price of stability of simple symmetric fractional hedonic games. In *SAGT'16*, pages 220–232, 2016.

Christos Kaklamanis, Panagiotis Kanellopoulos, and Dimitris Patouchas. On the price of stability of social distance games. In *SAGT'18*, pages 125–136, 2018.

Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *STACS'99*, pages 404–413. Springer, 1999.

Gianpiero Monaco, Luca Moscardelli, and Yllka Velaj. Stable outcomes in modified fractional hedonic games. In *AAMAS'18*, pages 937–945, 2018.

Gianpiero Monaco, Luca Moscardelli, and Yllka Velaj. On the performance of stable outcomes in modified fractional hedonic games with egalitarian social welfare. In *AAMAS'19*, pages 873–881, 2019.

Thayer Morrill. The roommates problem revisited. *Journal of Economic Theory*, 145:1739–1756, 2010.

Martin Olsen. On defining and computing communities. In *CATS'12*, pages 97–102, 2012.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

Tim Roughgarden and Éva Tardos. Introduction to the inefficiency of equilibria. In N. Nisan, T. Roughgarden, É. Tardos, and V.V. Vazirani, editors, *Algorithmic game theory*. Cambridge University Press, Cambridge, 2007.

Ola Rozenfeld. Strong equilibrium in congestion games. Master's thesis, Technion-Israel Institute of Technology, Faculty of Computer Science, 2007.

Shao Chin Sung and Dinko Dimitrov. On myopic stability concepts for hedonic games. *Theory and Decision*, 62(1):31–45, 2007.

APPENDIX

A.1 Preliminaries for the Appendix

We often make use of the following technical claim.

Proposition A.1. *For every pair of positive reals $x, y \geq 1$ it holds that $\frac{x+1}{2y} \leq \frac{x}{y}$.*

Proof. We have $y \leq xy$ and hence $y + yx \leq 2xy$. Dividing both sides by $2y^2$, we get the desired inequality. $\square$

A.2 Proof of Lemma 4.8

Assume for the sake of contradiction that $\mathcal{P}$ contains a coalition P that is not a singleton, a star, or a superstar. By Proposition 4.3 coalition P is connected. Let $r \in P$ be a median of P, that is, a node such that the size of each connected component of $G_{P \setminus \{r\}}$ is at most $\frac{|P|}{2}$. Assume without loss of generality that $\mathsf{Children}_r(G_P, r) = \{1, 2, \ldots, d\}$ and for every $k \in \{1, 2 \ldots, d\}$ set $S_k = \mathsf{Descendants}_r(G_P, k) \cup \{k\}$. Since r is the median of P, it must hold that $d = \delta_P(r) \geq 2$ and $|S_k| \leq \frac{|P|}{2}$ for every $k \in \{1, 2 \ldots, d\}$.

Assume without loss of generality that for every $1 \leq k \leq k' \leq d$ it holds that $|S_k| \leq |S_{k'}|$. Suppose first that there exists an index $h \in \{1, \ldots, d-1\}$ such that $|S_1| = |S_2| = \ldots = |S_h| < |S_{h+1}|$. Let $R = S_1 \cup S_2 \cup \ldots S_h \cup \{r\}$. Note that $|R| \geq 2$ and $|S_k| \geq 2$ for every $k \geq h+1$. Let us define a new partition $\mathcal{S} = (\mathcal{P} \setminus \{P\}) \cup \{R, S_{h+1}, \ldots, S_d\}$. We will argue that $\mathcal{S}$ Pareto dominates $\mathcal{P}$, which contradicts the Pareto optimality of $\mathcal{P}$.

For every $j \notin P$ we have $v_j(\mathcal{P}) = v_j(\mathcal{S})$. Let us now consider the members of P.

For every $k \in \{h+1, \ldots, d\}$ and every $j \in S_k \setminus \{k\}$ we have $\mathcal{P}(j) = P$ and $\mathcal{S}(j) = S_k$. Since $\delta_P(j) = \delta_{S_k}(j)$ and $|P| > |S_k|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_P(j)}{|P|} < \frac{\delta_{S_k}(j)}{|S_k|} = v_j(\mathcal{S}).$$

Further, for every $k \in \{h+1, \ldots, d\}$ we have $\mathcal{P}(k) = P$ and $\mathcal{S}(k) = S_k$. Since $\delta_P(k) = \delta_{S_k}(k) + 1$ and $|P| \geq 2|S_k|$, we get

$$v_k(\mathcal{P}) = \frac{\delta_P(k)}{|P|} \leq \frac{\delta_{S_k}(k) + 1}{2|S_k|} \leq \frac{\delta_{S_k}(k)}{|S_k|} = v_k(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For every $k = 1, \ldots, h$ and every $j \in S_k$, we have $\mathcal{P}(j) = P$ and $\mathcal{S}(j) = R$. Since $\delta_P(j) = \delta_R(j)$ and $|P| > |R|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_P(j)}{|P|} < \frac{\delta_R(j)}{|R|} = v_j(\mathcal{S}).$$

Finally, for r, we have $\mathcal{P}(r) = P$ and $\mathcal{S}(r) = R$. Hence, by our choice of h,

$$v_r(\mathcal{P}) = \frac{\delta_P(r)}{|P|} = \frac{d}{|S_1| + |S_2| + \ldots + |S_d| + 1} \leq \frac{d}{h|S_h| + (d-h)|S_{h+1}| + 1},$$
$$v_r(\mathcal{S}) = \frac{\delta_R(r)}{|R|} = \frac{h}{h|S_h| + 1} = \frac{1}{|S_h| + 1/h}.$$

Since $|S_h| + 1 \leq |S_{h+1}|$ and $1 \leq h < d$, we get

$$v_r(\mathcal{P}) \leq \frac{d}{h|S_h| + (d-h)(|S_h| + 1) + 1} = \frac{1}{|S_h| + 1 - (h-1)/d} \leq \frac{1}{|S_h| + 1/h} = v_r(\mathcal{S}),$$

where the last inequality follows from the fact that $1/h \leq 1 - (h-1)/d$, for every $h < d$.

Alternatively, suppose that $|S_k| = |S_{k'}|$ for every $k, k' \in \{1, 2, \ldots, d\}$; since P is not a star, each S_k contains at least two players. Further, since all sets S_k have the same size, and P is not a superstar, there must exist a player j at distance 3 from r. Let j^* be the parent of j with respect to r; we can assume without loss of generality that j^* is in

S_1. Let $R = \mathsf{Descendants}_r(G_P, j^*) \cup \{j^*\}$ and $T = (S_1 \setminus R) \cup \{r\}$. Note that $|R| \geq 2$. We will now argue that the partition $\mathcal{S} = (\mathcal{P} \setminus \{P\}) \cup \{R, T, S_2, S_3, \ldots, S_d\}$. Pareto dominates $\mathcal{P}$, which contradicts the Pareto optimality of $\mathcal{P}$.

For every $j \notin P$ we have $v_j(\mathcal{P}) = v_j(\mathcal{S})$. Let us now consider the members of P.

For every $k = 2, \ldots, d$ and every $j \in S_k \setminus \{k\}$ we have $\mathcal{P}(j) = P$ and $\mathcal{S}(j) = S_k$. Since $\delta_P(j) = \delta_{S_k}(j)$ and $|P| > |S_k|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_P(j)}{|P|} < \frac{\delta_{S_k}(j)}{|S_k|} = v_j(\mathcal{S}).$$

Further, for every $k = 2, \ldots, d$ we have $\mathcal{P}(k) = P$ and $\mathcal{S}(k) = S_k$. Since $\delta_P(k) = \delta_{S_k}(k) + 1$ and $|P| \geq 2|S_k|$, we get

$$v_k(\mathcal{P}) = \frac{\delta_P(k)}{|P|} \leq \frac{\delta_{S_k}(k) + 1}{2|S_k|} \leq \frac{\delta_{S_k}(k)}{|S_k|} = v_k(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For every $j \in R \setminus \{j^*\}$, we have $\mathcal{P}(j) = P$ and $\mathcal{S}(j) = R$. Since $\delta_R(j) = \delta_P(j)$ and $|P| > |R|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_P(j)}{|P|} < \frac{\delta_R(j)}{|R|} = v_j(\mathcal{S}).$$

For j^*, we have $\mathcal{P}(j^*) = P$ and $\mathcal{S}(j^*) = R$. Since $\delta_P(j^*) = \delta_R(j^*) + 1$ and $|P| \geq 2|S_1| > 2|R|$, we get

$$v_{j^*}(\mathcal{P}) = \frac{\delta_P(j^*)}{|P|} < \frac{\delta_R(j^*) + 1}{2|R|} \leq \frac{\delta_R(j^*)}{|R|} = v_{j^*}(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For every $j \in S_1 \setminus (R \cup \{1\})$ we have $\mathcal{P}(j) = P$ and $\mathcal{S}(j) = T$. Since $\delta_P(j) = \delta_T(j)$ and $|P| > |T|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_P(j)}{|P|} < \frac{\delta_T(j)}{|T|} = v_j(\mathcal{S}).$$

For player 1, we have $\mathcal{P}(1) = P$ and $\mathcal{S}(1) = T$. Since $\delta_P(1) = \delta_T(1) + 1$ and $|P| \geq 2|S_1| > 2|T|$, we get

$$v_1(\mathcal{P}) = \frac{\delta_P(1)}{|P|} < \frac{\delta_T(1) + 1}{2|T|} \leq \frac{\delta_T(1)}{|T|} = v_1(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For r, we have $\mathcal{P}(r) = P$ and $\mathcal{S}(r) = T$. We get

$$v_r(\mathcal{P}) = \frac{\delta_P(r)}{|P|} = \frac{d}{|S_1| + |S_2| + \ldots + |S_d| + 1} = \frac{d}{d|S_1| + 1} = \frac{1}{|S_1| + 1/d} < \frac{\delta_T(r)}{|T|} = v_r(\mathcal{S}),$$

where the last inequality follows from the fact that $|T| < |S_1|$. $\square$

A.3 Proof of Proposition 4.9

We prove each statement separately.

(a) Assume for the sake of contradiction that j forms a singleton coalition $R = \{j\}$ in $\mathcal{P}$. Let $R' = \{i, j\}$, and let $\mathcal{S} = (\mathcal{P} \setminus \{P, R\}) \cup \{R'\}$. It is easy to see that $\mathcal{S}$ Pareto dominates $\mathcal{P}$, which contradicts the Pareto optimality of $\mathcal{P}$.

(b) Assume for the sake of contradiction that j is the center of a superstar R. Assume without loss of generality that $\mathsf{Children}_j(G_R, j) = \{1, 2, \ldots, d\}$, and set $S_k = \mathsf{Descendants}_j(G_R, k) \cup \{k\}$ for every $k \in \{1, 2, \ldots, d\}$, and $R' = (R \setminus S_1) \cup \{i\}$. We prove our claim by showing that the partition $\mathcal{S} = (\mathcal{P} \setminus \{P, R\}) \cup \{S_1, R'\}$ Pareto dominates $\mathcal{P}$.

We first note that for every $t \notin P \cup R$ we have $v_t(\mathcal{P}) = v_t(\mathcal{S})$.

For i we have $\mathcal{P}(i) = P$ and $\mathcal{S}(i) = R'$. Since $\delta_P(i) = 0$ and $\delta_{R'}(i) = 1$, we get

$$v_i(\mathcal{P}) = \frac{\delta_P(i)}{|P|} < \frac{\delta_{R'}(i)}{|R'|} = v_i(\mathcal{S}).$$

For j we have $\mathcal{P}(j) = R$ and $\mathcal{S}(j) = R'$. Since $\delta_R(j) = \delta_{R_1}(j)$ and $|R| > |R'|$, where the last inequality follows from the fact that $|S_1| > |P|$, we get

$$v_j(\mathcal{P}) = \frac{\delta_R(j)}{|R|} < \frac{\delta_{R_1}(j)}{|R'|} = v_j(\mathcal{S}).$$

For every $k \in \{2, \dots, d\}$ and every $t \in S_k$ we have $\mathcal{P}(t) = R$ and $\mathcal{S}(t) = R'$. Since $\delta_R(t) = \delta_{R'}(t)$ and $|R| > |R'|$, where the last inequality follows from the fact that $|S_1| > |P|$, we get

$$v_t(\mathcal{P}) = \frac{\delta_R(t)}{|R|} < \frac{\delta_{R'}(t)}{|R'|} = v_t(\mathcal{S}).$$

For every $t \in S_1 \setminus \{1\}$ we have $\mathcal{P}(t) = R$ and $\mathcal{S}(t) = S_1$. Since $\delta_R(t) = \delta_{S_1}(t) = 1$ and $|R| \geq 2|S_1| + 1$, we get

$$v_t(\mathcal{P}) = \frac{\delta_R(t)}{|R|} \leq \frac{\delta_{S_1}(t)}{2|S_1| + 1} < \frac{\delta_{S_1}(t)}{|S_1|} = v_t(\mathcal{S}).$$

Finally, for player 1 we have $\mathcal{P}(1) = R$ and $\mathcal{S}(1) = S_1$. Since $\delta_R(1) = \delta_{S_1}(1) + 1$ and $|R| > 2|S_1|$, we get

$$v_1(\mathcal{P}) = \frac{\delta_R(1)}{|R|} < \frac{\delta_{S_1}(1) + 1}{2|S_1|} \leq \frac{\delta_{S_1}(1)}{|S_1|} = v_1(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

(c) Assume for the sake of contradiction that j is a leaf of a superstar R with center r. Assume without loss of generality that $\mathsf{Children}_r(G_R, r) = \{1, 2, \dots, d\}$ and set $S_k = \mathsf{Descendants}_r(G_R, k) \cup \{k\}$ for every $k \in \{1, 2, \dots, d\}$. Without loss of generality, we assume that $j \in S_1$. Let $R' = \{i, j\}$ and $R_1 = (S_1 \setminus \{j\}) \cup \{r\}$. To prove our claim, we show that the partition $\mathcal{S} = (\mathcal{P} \setminus \{P, R\}) \cup \{R', R_1, S_2, S_3, \dots, S_d\}$ Pareto dominates $\mathcal{P}$.

For every $t \notin P \cup R$ we have $v_t(\mathcal{P}) = v_t(\mathcal{S})$.

For i we have $\mathcal{P}(i) = P$ and $\mathcal{S}(i) = R'$. Since $\delta_P(i) = 0$ and $\delta_{R'}(i) = 1$, we get

$$v_i(\mathcal{P}) = \frac{\delta_P(i)}{|P|} < \frac{\delta_{R'}(i)}{|R'|} = v_i(\mathcal{S}).$$

For j we have $\mathcal{P}(j) = R$ and $\mathcal{S}(j) = R'$. Since $\delta_R(j) = \delta_{R'}(j) = 1$ and $|R| > |R'| = 2$, we get

$$v_j(\mathcal{P}) = \frac{\delta_R(j)}{|R|} < \frac{\delta_{R'}(j)}{|R'|} = v_j(\mathcal{S}).$$

For every $k \in \{2, \dots, d\}$ and every $t \in S_k \setminus \{k\}$ we have $\mathcal{P}(t) = R$ and $\mathcal{S}(t) = S_k$. Since $\delta_R(t) = \delta_{S_k}(t) = 1$ and $|R| > |S_k|$, we get

$$v_t(\mathcal{P}) = \frac{\delta_R(t)}{|R|} < \frac{\delta_{S_k}(t)}{|S_k|} = v_t(\mathcal{S}).$$

For every $k \in \{2, \dots, d\}$ we have $\mathcal{P}(k) = R$ and $\mathcal{S}(k) = S_k$. Since $\delta_R(k) = \delta_{S_k}(k) + 1$ and $|R| > 2|S_k|$, we get

$$v_k(\mathcal{P}) = \frac{\delta_R(k)}{|R|} < \frac{\delta_{S_k}(k) + 1}{2|S_k|} \leq \frac{\delta_{S_k}(k)}{|S_k|} = v_k(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For every $t \in S_1 \setminus \{1, j\}$, we have $\mathcal{P}(t) = R$ and $\mathcal{S}(t) = R_1$. Since $\delta_R(t) = \delta_{R_1}(t) = 1$ and $|R| > |R_1|$, we get

$$v_t(\mathcal{P}) = \frac{\delta_R(t)}{|R|} < \frac{\delta_{R_1}(t)}{|R_1|} = v_t(\mathcal{S}).$$

For player 1 we have $\mathcal{P}(1) = R$ and $\mathcal{S}(1) = R_1$. Since $\delta_R(1) = \delta_{R_1}(1) + 1$ and $|R| \geq 2|S_1| + 1 = 2|R_1| + 1 > 2|R_1|$, we get

$$v_1(\mathcal{P}) = \frac{\delta_R(1)}{|R|} < \frac{\delta_{R_1}(1) + 1}{2|R_1|} \leq \frac{\delta_{R_1}(1)}{|R_1|} = v_1(\mathcal{S}),$$

where the second inequality follows from Proposition A.1.

For r, we have $\mathcal{P}(r) = R$, $\mathcal{S}(r) = R_1$. Since $|R| = d \cdot |S_1| + 1$, $|R_1| = |S_1|$ and $\delta_{R_1}(r) = 1$, we get

$$v_r(\mathcal{P}) = \frac{\delta_R(r)}{|R|} = \frac{d}{d \cdot |S_1| + 1} < \frac{1}{|S_1|} = \frac{\delta_{R_1}(r)}{|R_1|} = v_r(\mathcal{S}).$$

$\square$