# Welfare Maximization in Fractional Hedonic Games

**Haris Aziz** and **Serge Gaspers**
NICTA and UNSW, Sydney, Australia
{haris.aziz, serge.gaspers}@nicta.com.au

**Joachim Gudmundsson** and **Julián Mestre**
University of Sydney, Sydney, Australia
{joachim.gudmundsson, mestre}@sydney.edu.au

**Hanjo Täubig**
TU Munich, Munich, Germany
taeubig@in.tum.de

## Abstract

We consider the computational complexity of computing welfare maximizing partitions for fractional hedonic games—a natural class of coalition formation games that can be succinctly represented by a graph. For such games, welfare maximizing partitions constitute desirable ways to cluster the vertices of the graph. We present both intractability results and approximation algorithms for computing welfare maximizing partitions.

## 1 Introduction

Social network analysis is one of the pressing problems in various disciplines including sociology, economics, physics, and computer science [Brandes and Erlebach, 2005]. One of the most prominent issues within the area is that of *network clustering*: given a network, can it be partitioned into cohesive groups [Newman, 2004; Schaeffer, 2007]? The problem is also referred to as *community detection* [Olsen, 2012].

In network clustering, the predominant approaches to find useful clusters is to maximize intra-cluster density measures and to minimize inter-cluster density measures. Various clustering measures that constitute tradeoffs between the two criteria have been introduced. An important research issue is to identify the most desirable clustering measure among the plethora of clustering measures and rules. Most measures do not have a game-theoretic justification that may have some merit when modeling network clustering in distributed environments.

In recent work, there has been a push to use game-theoretic concepts to model network clustering [Papadimitriou, 2001]. *Hedonic games* is an example of such models. A hedonic game comprises a set of agents who express preferences over coalitions they are present in and outcomes are partitions of the agents into disjoint coalitions [Bogomolnaia and Jackson, 2002; Aziz and Savani, 2015]. It provides a natural framework to study coalition formation. A particularly relevant class of hedonic games that helps model network clustering is that of *fractional hedonic games (FHGs)*, which further satisfy desirable properties [Aziz *et al.*, 2014]. In FHGs, each vertex of the network can be considered as an agent. An agent $i$'s valuation $v_i(j)$ of an agent $j$ can be represented by the weight of the directed edge $(i, j)$. Agent $i$'s valuation of a coalition $S$ of agents is then the mean valuation $\sum_{j \in S} v_i(j)/|S|$ of the members of $S$. Based on the FHG, outcomes that satisfy some notion of stability or welfare are considered desirable clusterings. FHGs have a nice correspondence with clustering: each agent corresponds to a vertex; each coalition corresponds to a cluster; and a partition corresponds to a clustering. We refer to Section 6 in [Aziz *et al.*, 2014] for detailed motivation behind studying welfare maximizing partitions for network clustering. Welfare maximizing partitions of fractional hedonic games satisfy some axioms (scale invariance, richness, and consistency) that many classic network clustering methods do not satisfy simultaneously. Other than providing a novel perspective on network clusterings, FHGs also capture natural social, economic, and political settings in which agents 'seek to maximize the average agreement with the members of their coalition' [Brandl *et al.*, 2015].

Fractional hedonic games have recently received interest in various papers with the focus primarily being on stability concepts [Brandl *et al.*, 2015; Bilò *et al.*, 2014; 2015]. In this paper, we focus on maximizing various welfare objectives including *utilitarian* (sum of utilities), *egalitarian* (utility of worst off agent) and *Nash welfare* (product of utilities).

**Contributions** We present several results on FHGs, and in particular for simple symmetric FHGs. An FHG is *symmetric* if every two agents have the same valuation for each other, and it is *simple* if all valuations are 0 or 1. Our results are summarized as follows. We present (i) simple examples that show that utilitarian, egalitarian, and Nash welfare maximizing outcomes need not coincide, even in simple symmetric FHGs; (ii) a reduction that shows that maximizing utilitarian welfare, egalitarian welfare, or Nash welfare is NP-hard, even for simple symmetric FHGs; (iii) a polynomial-time 2-approximation algorithm for maximizing the utilitarian welfare of simple symmetric FHGs; (iv) a polynomial-time 4-approximation algorithm for maximizing the utilitarian welfare of symmetric FHGs; and (v) a polynomial-time 3-approximation algorithm for maximizing the egalitarian welfare of simple symmetric FHGs. The computational results are summarized in Table 1.

| Objective | Restriction | Complexity | Reference |
|---|---|---|---|
| utilitarian | simple symmetric | NP-hard | Th. 3 |
| utilitarian | simple symmetric | 2-approx | Th. 6 |
| utilitarian | symmetric | 4-approx | Th. 7 |
| egalitarian | simple symmetric | NP-hard | Th. 3 |
| egalitarian | simple symmetric | 3-approx | Th. 8 |
| Nash | simple symmetric | NP-hard | Th. 4 |

Table 1: Our results for welfare maximization for FHGs.

**Related Work** Network clustering and community detection are vast fields with many clustering measures and algorithms introduced for the problems. We recommend [Newman, 2004] and [Schaeffer, 2007] for surveys on the topic.

Other classes of hedonic games based on graphs have been examined from a social welfare perspective. In particular, additively separable hedonic games are a related class of games that can also be represented by graphs [Aziz *et al.*, 2013]. Aziz *et al.* [2014] pointed out axiomatic justification to consider welfare maximizing or stable partitions of FHGs and why stable or efficient outcomes of FHGs provide better clusterings than their counterparts for additively separable hedonic games. Olsen [2012] examined a variant of FHGs and considered computation of Nash stable outcomes. In the games he considered, however, every perfect matching is a best possible outcome even if there are large cliques present in the graph. By contrast, in FHGs, agents have an incentive to form large cliques.

In prior work on FHGs, most of the focus has been on stable partitions. Although stability concepts like core stability capture incentive aspects, a disadvantage of these stability concepts is that a stable outcome may not be guaranteed to exist [Aziz *et al.*, 2014; Bilò *et al.*, 2014; Brandl *et al.*, 2015] or may suggest the partition consisting of the grand coalition [Bilò *et al.*, 2014]. In these situations the stability concept may not always suggest a desirable clustering or partition.

## 2 Preliminaries

Let $N$ be a set $\{1, \ldots, n\}$ of agents or players. With a slight abuse of terminology we refer to both the set $N$ of all players and the partition $\{N\}$ as the *grand coalition*. A *hedonic game* is a pair $(N, \succsim)$, where $\succsim = (\succsim_1, \ldots, \succsim_n)$ is a profile of complete and transitive relations $\succsim_i$, modeling the preferences of the players. An outcome of a hedonic game is a partition $\mu$ of the agents in $N$. Each set in $\mu$ is called a *coalition*. The coalition in which agent $i$ participates in partition $\mu$ is denoted by $\mu(i)$.

A *value function* for a player $i$ is defined as a function $v_i \colon N \to \mathbb{R}$ assigning a real value to each of the players. Unless stated otherwise we assume $v_i(i) = 0$. A value function $v_i$ induces a value function over coalitions where for each coalition $S$ containing agent $i$, the *utility* agent $i$ derives from participating in coalition $S$ is $v_i(S) = \frac{\sum_{j \in S} v_i(j)}{|S|}$. A hedonic game $(N, \succsim)$ is now said to be a *fractional hedonic game (FHG)* if for each player $i$ in $N$ there is a value function $v_i$ such that for all coalitions $S, T \subseteq N$, $S \succsim_i T$ if and only if
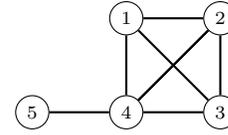


Figure 1: The partition $\{\{1,2,3,4\}, \{5\}\}$ maximizes utilitarian welfare. However, its egalitarian and Nash welfares are 0. The maximum egalitarian/Nash welfare partition $\{\{1,2,3\}, \{4,5\}\}$.

$v_i(S) \geq v_i(T)$. Hence a fractional hedonic games can simply be defined by the value function $v$.

An FHG is said to be *symmetric* if $v_i(j) = v_j(i)$ and *simple* if $v_i(j) \in \{0, 1\}$. A simple FHG $(N, \succsim_i)$ can be represented by a *directed* graph $(N, E)$ in which $(i, j) \in E$ if and only if $v_i(j) = 1$. In a much similar fashion, if $(N, \succsim_i)$ is both symmetric and simple, it can be represented by an *undirected* graph $(N, E)$ such that $\{i, j\} \in E$ if and only if $v_i(j) = v_j(i) = 1$. The complete undirected graph on $n$ vertices is denoted by $K_n$ whereas an undirected cycle on $n$ vertices is denoted by $C_n$.

We consider the following notions of welfare of a partition $\mu$ of $N$: (i) utilitarian welfare: $\sum_{i \in N} v_i(\mu(i))$; (ii) egalitarian welfare: $\min_{i \in N} v_i(\mu(i))$; (iii) Nash welfare: $\prod_{i \in N} v_i(\mu(i))$. Before we formalize computational problems to maximize the three welfare notions, we point out that even in the restricted domain of simple symmetric FHGs, the three welfare notions are different. We first present examples that show that the utilitarian, egalitarian, and Nash welfare objectives are not equivalent even for simple symmetric FHGs.

**Example 1** *Consider a clique $K_k$ of size $k$ with an additional vertex $v$ attached using a single edge (see Fig. 1). If $k \geq 4$, the partition $\{K_k, \{v\}\}$ maximizes utilitarian welfare, which is $\frac{k(k-1)}{k} + \frac{0}{1} = k - 1$. But it gives zero egalitarian and Nash welfare, while other partitions have greater egalitarian and Nash welfare.*

**Example 2** *For the FHG in Fig. 2, we know that a perfect matching gives maximum egalitarian welfare $1/2$ and Nash welfare $(1/2)^{10} = 0.0009765625$. However, $\{\{1, 2, 3, b_4, b_5\}, \{4, 5, b_1, b_2, b_3\}\}$ has greater Nash welfare, namely $\left(\frac{4}{5}\right)^3 \left(\frac{3}{5}\right)^2 \left(\frac{4}{5}\right)^2 \left(\frac{2}{5}\right)^3 \approx 0.0075497472$.*

We consider the following computational problems for welfare objective $\mathcal{W} \in \{\text{UTILITARIAN, EGALITARIAN, NASH}\}$.

**Name**: $\mathcal{W}$ WELFARE
**Instance**: An FHG represented by a directed weighted graph $G = (N, E)$.
**Output**: A partition that maximizes the $\mathcal{W}$ welfare.

When considering the restriction of these problems to symmetric (simple) FHGs, $G$ will be undirected (and unweighted).
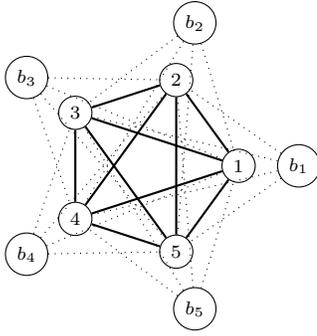
Figure 2: The maximum egalitarian welfare partition $\{\{1, b_2\}, \{2, b_3\}, \{3, b_4\}, \{4, b_5\}, \{5, b_1\}\}$ does not maximize Nash welfare. Some edges are dotted for easier visualization.

## 3 Hardness

In this section, we examine partitions with the maximum utilitarian welfare (which maximize the sum of the utilities of the players), the maximum egalitarian welfare (which maximize the utility of the player who is worst off) and the Nash welfare (which maximizes the product of utilities of the players). First we point out an alternative way to define maximum utilitarian partitions of simple FHGs: $\max_\mu \sum_{i \in N} v_i(\mu(i)) = \max_\mu \sum_{C \in \mu} \frac{2|E(C)|}{|C|}$. Here, $E(C)$ denotes the edges that have both endpoints in $C$. We define the *utility* of a cluster $C \in \mu$ to be the contribution of the vertices in $C$ to the objective; in other words, the utility of $C$ is defined as $u(C) = \frac{2|E(C)|}{|C|}$. When the setting is weighted, we define the utility of $C$ as $u(C) = \frac{2w(E(C))}{|C|}$.

**Theorem 3** *For simple symmetric FHGs,* UTILITARIAN WELFARE *and* EGALITARIAN WELFARE *are both NP-hard.*

*Proof:* Our proof is by a polynomial reduction from PARTITIONINTOTRIANGLES [Garey and Johnson, 1979].

**Name**: PARTITIONINTOTRIANGLES
**Instance**: An undirected weighted graph $G = (V, E)$ with $|V| = 3q = n$ for some integer $q$.
**Output**: 'Yes' if the vertices of $G$ can be partitioned into $q$ disjoint sets $V_1, V_2, ..., V_q$, each containing exactly 3 vertices, such that each of these $V_i$ is the node set of a triangle in $G$. 'No' otherwise.

It is well-known that PARTITIONINTOTRIANGLES is NP-complete even for 3-partite graphs in which each of the independent three sets is specified. We reduce PARTITIONINTOTRIANGLES for 3-partite graphs to computing a partition with the maximum utility. We denote the three vertex sets in $G$ as $X, Y$, and $Z$.

The claim is that a partition into triangles is a partition with the maximum utility. It is clear that each player gets utility $2/3$. Therefore, the total utility of the partition is $2n/3$. Now assume that there exists a partition $\mu$ other than a partition into triangles which achieves utility greater than $2n/3$. Then,

$\mu$ must contain at least one coalition $S$ which has average utility greater than $2/3$. However, we show that this is not possible. For a coalition of size one or two, this certainly does not hold. Let us assume that we have a coalition $S$ with $x$ vertices from $X$, $y$ from $Y$, and $z$ vertices from $Z$ such that $x \geq y \geq z \geq 1$ and $x \geq 2$. In the worst case (for partition into triangles to not maximize welfare), the graph induced by vertices in $S$ is a complete 3-partite graph. Then the utility of vertices in $S$ of type $X$ is $\frac{y+z}{x+y+z}$, the utility of vertices in $S$ of type $Y$ is $\frac{x+z}{x+y+z}$, and the utility of vertices in $S$ of type $Z$ is $\frac{x+y}{x+y+z}$. Therefore, the average utility of vertices in $S$ is

$$\frac{\frac{z(x+y)}{x+y+z} + \frac{x(y+z)}{x+y+z} + \frac{y(x+z)}{x+y+z}}{x+y+z} = \frac{2(xy+yz+xz)}{(x+y+z)^2} .$$

We need to prove that this value is less than or equal $2/3$. Therefore, it is sufficient to prove that

$$(x+y+z)^2 \geq 3(xy+yz+xz)$$
iff $\quad\quad x^2 + y^2 + z^2 \geq xy + yz + xz$
iff $\quad (x-z)(x-y) + (y-z)(y-z) \geq 0.$

Since $x \geq y \geq z$, the last inequality holds so we know that for any coalition, the average utility can be at most $2/3$. The inequality is tight, i.e., average utility is exactly $2/3$ only if $x = y = z$ and the graph induced by the vertices in $S$ is a complete 3-partite graph. But in that case, $S$ can be partitioned into $x$ triangles. Therefore, we have proved that if we want to check whether $G$ can be partitioned into triangles, it is enough to compute a partition with the maximum utilitarian welfare and check whether each vertex has utility exactly $2/3$.

The same reduction and argument can be used to prove that computing a partition with the maximum egalitarian utility is also NP-hard. $\quad\quad\square$

We use a similar argument to prove the following.

**Theorem 4** *For simple symmetric FHGs,* NASH WELFARE *is NP-hard.*

*Proof:* Consider the reduction in the proof of Theorem 3. The claim is that a partition into triangles is the partition with the maximum Nash welfare. It is clear that each player gets utility $2/3$. Therefore, the Nash welfare of the partition is $(2/3)^n$. Now assume that there exists a partition $\mu$ other than a partition into triangles which achieves Nash welfare utility greater than $(2/3)^n$ which means that the geometric mean of the Nash welfare utilities is greater than $2/3$. Then, $\mu$ must contain at least one coalition in which the geometric mean of the Nash welfare utilities of the coalition is greater than $2/3$. However, we show that this is not possible. Let us assume again that we have a coalition $S$ of size at least three with $x$ vertices from $X$, $y$ from $Y$, and $z$ vertices from $Z$ such that $x \geq y \geq z \geq 1$ and $x \geq 2$. In the worst case (for partition into triangles to not maximize welfare), the graph induced by vertices in $S$ is a complete 3-partite graph. Therefore, the geometric mean of the Nash welfare of vertices in $S$ is $\sqrt[x+y+z]{\left(\frac{x+y}{x+y+z}\right)^z \left(\frac{y+z}{x+y+z}\right)^x \left(\frac{x+z}{x+y+z}\right)^y}$. We need to prove

that this value is less than or equal $2/3$. We use the inequality of arithmetic and geometric means:

$$\sqrt[n]{x_1 x_2 \cdots x_n} \le \frac{x_1 + x_2 + \cdots + x_n}{n} \quad \forall x_1, \ldots, x_n \ge 0.$$

Therefore, for any coalition, the geometric mean of the Nash welfare utilities of the vertices in the coalition is at most the arithmetic mean of the utilitarian welfare valuations which we showed can be at most $2/3$. An average utilitarian welfare of exactly $2/3$ can be achieved only if $x = y = z$ and the graph induced by the vertices in $S$ is complete 3-partite. But in that case, $S$ can be partitioned into $x$ triangles. $\qquad\square$

Bilò *et al.* [2014] presented a more elaborate reduction to show that computing a maximum *utilitarian* welfare Nash stable partition as well as maximizing *utilitarian* welfare is NP-hard. Our reduction is relatively simpler and simultaneously shows that three different measures of welfare are NP-hard to achieve. In subsequent sections, we examine the approximation of maximum utilitarian welfare and egalitarian welfare.

## 4 Approximating utilitarian welfare

In this section, we present approximation algorithms for UTILITARIAN WELFARE for symmetric FHGs. We start with simple symmetric FHGs, where we obtain better approximation ratios.

### 4.1 Utilitarian welfare for simple symmetric FHGs

Our first algorithm takes as input an undirected unweighted graph $G = (N, E)$. It computes a *maximal* matching $M$ and outputs the clustering induced by the matching; namely, for each edge $\{u, v\} \in M$, we output the cluster $\{u, v\}$, and for each unmatched vertex $u \in N$, we output the cluster $\{u\}$.

**Theorem 5** *For simple symmetric FHGs,* UTILITARIAN WELFARE *has a linear-time 4-approximation algorithm.*

*Proof:* We can compute a maximal matching and its induced clustering in $O(|N| + |E|)$ time, so the time complexity claim follows.

Let $M$ be the maximal matching computed by the algorithm and $\mu$ be the clustering induced by $M$. Let $\mu^* = \{C_1, \ldots, C_p\}$ be an optimal clustering of $G$. Let $A$ be the vertex set of $M$ and $B$ be $N \setminus A$. Note that by the maximality of $M$, there are no edges connecting vertices in $B$.

Consider a cluster $C_i \in \mu^*$. It can contain vertices of $A$, denoted by $A_i$, and vertices of $B$, denoted by $B_i$. The utility of $C_i$ is maximized when the number of edges between the vertices of $C_i$ is maximized. The number of edges within $C_i$ is upper bounded by the number of edges between vertices in $A_i$, which is at most $\binom{|A_i|}{2}$, plus the number over edges connecting vertex in $A_i$ with a vertex in $B_i$, which is at most $|A_i| \cdot |B_i|$. The utility of $C_i$ is then

$$
\begin{aligned}
u(C_i) &\le \frac{2\binom{|A_i|}{2} + 2|A_i||B_i|}{|A_i| + |B_i|} \\
&= \frac{|A_i|(|A_i| - 1 + 2|B_i|)}{|A_i| + |B_i|} \le 2|A_i|
\end{aligned}
$$

Hence, the total utility of $\mu^*$ is at most $2|A|$. On the other hand, the total utility of $\mu$ is $|A|/2$. Therefore, the solution output is 4-approximate. $\qquad\square$

Our second algorithm is very similar but instead of computing a maximal matching, it uses a maximum cardinality matching.

**Theorem 6** *For simple symmetric FHGs,* UTILITARIAN WELFARE *has an $O(\sqrt{|N|}|E|)$-time 2-approximation algorithm.*

*Proof:* Notice that the maximum cardinality matching can be computed in $O(\sqrt{|N|}|E|)$ time [Micali and Vazirani, 1980], so the time complexity claim follows. Let $M$ be the maximum matching computed by the algorithm and $\mu$ be the clustering induced by $M$. Let $\mu^* = \{C_1, \ldots, C_p\}$ be an optimal clustering of $G$. Let $M' = M'_1 \cup \ldots \cup M'_p$ where $M'_i$, $1 \le i \le p$, is a maximum matching of the subgraph of $G$ induced by the vertices in $C_i$. Finally, let $\mu'_i$ be the clustering induced by $M'_i$ and let $\mu'$ be the clustering induced by $M'$.

Consider an arbitrary cluster $C_i$ of $\mu^*$ with vertex set $N_i$. Let $A_i$ be the vertex set of $M'_i$ and let $B_i = N_i \setminus A_i$.

**Claim 1** *For any $i = 1, \ldots, p$ and any edge $\{u, v\} \in M'_i$ we have $\deg_{B_i}(u) + \deg_{B_i}(v) \le |B_i| + 1$, where $\deg_{B_i}(u)$ is the number of neighbours of $u$ in $G$ that belong to $B_i$.*

*Proof:* If $u$ and $v$ do not have a common neighbor in $B_i$, the bound follows since $\deg_{B_i}(u) + \deg_{B_i}(v) \le |B_i|$. Suppose then that $u$ and $v$ are both connected to a vertex $x \in B_i$. If this is the case, then neither $u$ or $v$ can be connected to any other vertex in $B_i$, since this would create an augmenting path of length 3 (e.g., if $u$ was adjacent to some $y \in B_i - x$ then we have the augmenting path $\langle y, u, v, x \rangle$), contradicting the maximum cardinality of $M'_i$. Notice that in this case the bound also holds since $\deg_{B_i}(u) + \deg_{B_i}(v) = 2 \le |B_i| + 1$. $\qquad\square$

We upper bound the utility of $C_i$ as follows

$$
\begin{aligned}
u(C_i) &= \frac{1}{|C_i|} \sum_{u \in C_i} \deg_{C_i}(u) \\
&= \frac{1}{|C_i|} \sum_{u \in A_i} \left( \deg_{A_i}(u) + 2\deg_{B_i}(u) \right) \\
&= \frac{1}{|C_i|} \sum_{(u,v) \in M'_i} \big( \deg_{A_i}(u) + \deg_{A_i}(v) \\
&\qquad\qquad + 2\deg_{B_i}(u) + 2\deg_{B_i}(v) \big) \\
&\le \frac{1}{|A_i| + |B_i|} \sum_{(u,v) \in M'_i} (2(|A_i| - 1) + 2(|B_i| + 1)) \\
&= |A_i| \ .
\end{aligned}
$$

Therefore the total utility of the optimal solution is

$$u(\mu^*) \le \sum_{i=1}^p u(C_i) \le \sum_{i=1}^p |A_i| = 2 \cdot |M'|,$$

while the utility of $\mu'$ is

$$u(\mu') = \sum_{i=1}^{p} u(\mu_i) = \sum_{i=1}^{p} \frac{|A_i|}{2} = |M'|,$$

Since $|M| \geq |M'|$ we get

$$u(\mu^*) \leq 2 \cdot u(\mu') \leq 2 \cdot u(\mu).$$

Therefore the solution output is 2-approximate. $\qquad\square$

### 4.2 Utilitarian welfare for symmetric FHGs

We show how the 4-approximation for unweighted graphs can be extended to weighted graphs. Again, the algorithm is very similar but in this case we compute a maximum weight matching.

**Theorem 7** *For symmetric FHGs,* UTILITARIAN WELFARE *has a polynomial-time 4-approximation algorithm.*

*Proof:* Notice that a maximum weight matching can be computed in $O(|N|(|E| + |N| \log |N|))$ time [Gabow, 2009], so the time complexity claimed follows.

Let $M$ be the maximum weight matching computed by the algorithm and $\mu$ be the clustering induced by $M$. Let $\mu^* = \{C_1, \ldots, C_p\}$ be an optimal clustering. Let $M' = M_1 \cup \ldots \cup M_p$ where $M_i$, $1 \leq i \leq p$, is a greedy matching of the subgraph induced by the vertices in $C_i$, where a greedy matching is obtained by iteratively picking the heaviest edge. Finally, let $\mu'_i$ be the clustering induced by $M'_i$ and let $\mu'$ be the clustering induced by $M'$.

Let $A_i$ be the set of vertices matched in $M_i$ and $B_i = C_i \setminus A_i$. Furthermore, let $\{u_1, v_1\}, \ldots, \{u_k, v_k\}$ be the edges of $M_i$ in the order they were chosen, and let $w_1, \ldots, w_k$ denote their weight. Since $M_i$ is maximal, it follows that there are no edges within $B_i$. Therefore, for every edge $(x, y) \in E[C_i]$ it must be the case that $x = u_j$ or $y = v_j$ for some $i$; furthermore, by the greedy nature in which $M_i$ is computed, we have $w(x, y) \leq w_j$. In other words, every edge $(x, y)$ in $C_i$ has an endpoint $u_j$ or $v_j$ such that $w(e) \leq w_j$.

Therefore, the utility of cluster $C_i$ is at most

$$u(C_i) \leq \frac{1}{|C_i|} \sum_{j=1}^{k} 4(|C_i| - 1) w_j < 4 \cdot w(M_i).$$

On the other hand, the total utility of $\mu'_i$ is $w(M'_i)$. Furthermore, since $w(M') \leq w(M)$ we get

$$u(\mu^*) \leq 4 \cdot u(\mu') \leq 4 \cdot u(\mu).$$

Therefore the solution output is 4-approximate. $\qquad\square$

## 5 Approximating egalitarian welfare

In this section we consider the problem of maximizing egalitarian welfare for simple symmetric FHGs; that is, we want to maximize the value that the worst-off agent gets.

The algorithm takes as input an undirected unweighted graph $G = (N, E)$. If $G$ contains an isolated vertex, then output the grand coalition. The next step of the algorithm is to compute a maximum cardinality matching $M$ in $G$. If

the matching is perfect, the algorithm stops and returns the clustering induced by $M$. Otherwise, we construct a bipartite graph from $G$ by contracting every edge $\{u, v\} \in M$ into a super-node $\{u, v\}$ and deleting edges between super-nodes. Let $H = (S, U, F)$ be the resulting graph, where $S$ is the set of super-nodes, $U$ is the set of unmatched vertices in $G$, and $F$ is the set of edges in the contracted graph.

Next, we compute in $H$ an assignment of $U$ into $S$ whose first objective is to minimize the maximum number $\ell$ of vertices of $U$ any super-node in $S$ receives, and whose secondary objective is to minimize the number of super-nodes in $S$ that receive $\ell$ vertices from $U$. This can be done using a minimum cost network flow algorithm. For a given value of $\ell$, start from $H$ with all edges having capacity 1 and cost 0, add a source node pushing one unit of flow to each vertex from $U$ at no cost, and add a sink node that is connected to each vertex from $M$ with one edge with capacity $\ell - 1$ and cost 0 and another edge with capacity 1 and cost 1.

Finally, for each super-node $\{x, y\} \in S$ we look at the vertices in $U$ that were assigned to it, say $u_1, \ldots, u_k$, and put all the vertices into a single cluster, namely, $\{x, y, u_1, \ldots, u_k\}$.

**Theorem 8** *For simple symmetric FHGs,* EGALITARIAN WELFARE *has a polynomial-time 3-approximation algorithm.*

*Proof:* The running time of the described algorithm is clearly polynomial. We will show that it computes a partition whose egalitarian welfare is at least $1/3$ the maximum egalitarian welfare. If $G$ contains an isolated vertex $v$, then every partition maximizes egalitarian welfare, which is 0, since $v$ has utility 0 in every partition.

Let us now consider the case where $G$ contains no isolated vertex. First observe that for any coalition $T$ and any agent $i \in T$ we have $v_i(T) < 1$. Secondly, note that if the matching $M$ found by the algorithm is perfect, then for the clustering $\mu$ induced by $M$ we get $v_i(\mu(i)) = \frac{1}{2}$. Therefore, if $M$ is perfect the solution output is at least 2-approximate.

Let us then focus on the case where $M$ is not perfect and there is a set of vertices $U$ that are left unmatched. The set $U$ is an independent set in $G$, since $M$ is a maximum cardinality matching. The vertices in $U$ are then assigned to super-nodes in $S$ using the graph $H$. (Recall that each super-node in $S$ corresponds to an edge in $M$.) We note that such an assignment always exists since every vertex $u \in U$ must be adjacent in $H$ to at least one super-node because $U$ is independent and $G$ contains no isolated vertices. Let $\ell$ be the maximum number of vertices that any super-node receives.

**Claim 2** *The egalitarian welfare of $\mu$ is at least $\frac{1}{\ell+2}$.*

*Proof:* Every cluster $C$ we output is connected, so every vertex $i \in C$ has at least one neighbor in $C$. Hence,

$$v_i(C) = \frac{\deg_C(i)}{|C|} \geq \frac{1}{|C|} \geq \frac{1}{\ell + 2} \ .$$

$\qquad\square$

Therefore, if $\ell = 1$ the solution is 3-approximate. Let us assume from now on that $\ell \geq 2$. We will use $N_H(v)$ to denote the set of vertices adjacent to vertex $v$ in $H$.

Let $S' \subseteq S$ denote the super-nodes that receive exactly $\ell$ vertices from $U$ in the assignment. Iteratively add to $S'$ any super-node $s \in S \setminus S'$ such that $s$ is assigned exactly $\ell - 1$ vertices from $U$ and there exists $s' \in S'$ that is assigned a vertex $u \in N_H(s)$, until no such $s$ exists anymore. Let $U' \subseteq U$ denote the vertices from $U$ that are assigned to $S'$.

We will upper bound the maximum egalitarian welfare of any partition in terms of the parameter $\ell$. To that end, we analyze the coalitions containing vertices from $U'$. First, observe that $|U'| > (\ell - 1) \cdot |S'|$. We will now prove two claims about the neighbors of the vertices in $U'$.

**Claim 3** *Let $u \in U'$ and $s = \{x, y\} \in S'$. Then $u$ cannot be adjacent in $G$ to both $x$ and $y$; namely, $|N_G(u) \cap \{x, y\}| \leq 1$.*

*Proof:* For the sake of contradiction, assume that $u$ is adjacent in $G$ to both $x$ and $y$.

First, we will show that either $x$ or $y$ has a neighbor $u'$ in $U' \setminus \{u\}$. If a vertex from $U' \setminus \{u\}$ has been assigned to $s$, then we can take $u'$ to be such a vertex. Otherwise, since $\ell \geq 2$ and every vertex in $S'$ has been assigned at least $\ell - 1$ vertices, we have that $u$ is the unique vertex that has been assigned to $s$ and $\ell = 2$. By the construction of $S'$, we know that $s$ was added to $S'$ in an iteration where some super-node $s'$ was already in $S'$ and $s'$ was assigned a vertex $u' \in N_H(s)$. This is the vertex we were looking for.

Without loss of generality, assume that $u' \in N_G(x)$. Recall that the super-node $s = \{x, y\}$ corresponds to an edge $\{x, y\} \in M$. Now, we obtain an augmenting path $(u, y, x, u')$, which implies that $M$ is not maximum since $(M \setminus \{\{x, y\}\}) \cup \{\{u, y\}, \{x, u'\}\}$ is a larger matching in $G$—a contradiction. $\square$

**Claim 4** *Let $s = \{x, y\}$ be a super-node in $S'$. Then either $x$ or $y$ has neighbors in $U'$ in the graph $G$, but not both.*

*Proof:* For the sake of contradiction, assume that $x$ has a neighbor $u \in U'$ and $y$ has a neighbor $v \in U'$. By the previous claim, we have that $u \neq v$. But then, $(u, x, y, v)$ is an augmenting path for $M$ in $G$, which implies that $M$ is not a maximum matching—a contradiction. $\square$

For a super-node $s = \{x, y\} \in S'$, we call the vertex that has neighbors in $U'$ the *leader* of $s$. Denote by $L$ all the leaders of super-nodes in $S'$. Let us now prove that $L$ contains all the $G$-neighbors of the vertices in $U'$.

**Claim 5** $N_G(U') \subseteq L$.

*Proof:* Suppose $u \in U'$ has a neighbor $v \in V \setminus L$. Since $U$ is an independent set, $v \notin U$. By Claim 4 and since $v \notin L$, we have that $v$ cannot be a vertex inside a super-node in $S'$. The only remaining possibility is that $v$ is a vertex in a super-node $s \in S \setminus S'$. Since $u$ is assigned to some super-node $s' \in S'$ and $s$ has not been added to $S'$, we have that $s$ is assigned at most $\ell - 2$ vertices. But then, by the construction of $S'$, a series of changes can be performed in the assignment that
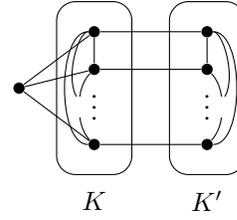


Figure 3: Instance showing that the analysis of the approximation ratio in Theorem 8 is tight.

decreases the number of super-nodes that are assigned $\ell$ vertices from $U$: assign $u$ to $s$ instead of $s'$; if $\ell$ vertices had been assigned to $s'$ we are done, otherwise assign the vertex from $U'$ to $s'$ that caused $s'$ to be added to $S'$ in the iterative construction, and continue in this way until eventually decreasing by one the number of super-nodes that are assigned $\ell$ vertices from $U$. This contradicts the fact that the assignment minimizes the number of edges that are assigned $\ell$ vertices. $\square$

In a partition $\mu$ maximizing the egalitarian welfare, each vertex $u \in U'$ shares a coalition with some vertex from $L$, since $N_G(u) \subseteq L$ and an egalitarian welfare maximizing partition contains no singleton coalitions. Consider a coalition $C \in \mu$ that contains at least one vertex from $L$ and maximizes the ratio $\frac{|C \cap U'|}{|C \cap L|}$. Since $|U'| > (\ell - 1) \cdot |S'| = (\ell - 1) \cdot |L|$, this ratio is greater than $\ell - 1$. Therefore, the value function for a vertex $i \in C \cap U'$ is

$$\frac{|C \cap N_G(i)|}{|C|} \leq \frac{|C \cap L|}{|C|} \leq \frac{|C \cap L|}{|C \cap U'| + |C \cap L|} \leq \frac{1}{\ell} .$$

In other words, the egalitarian welfare of $\mu$ is therefore at most $\frac{1}{\ell}$. On the other hand, the egalitarian welfare of our solution is at least $\frac{1}{\ell + 2}$. Therefore, the approximation ratio of the algorithm is at most $\frac{\ell + 2}{\ell} \leq 2$ for $\ell \geq 2$. $\square$

Fig. 3 shows an instance where our algorithm achieves an approximation ratio arbitrarily close to 3. To construct the instance, start with a disjoint union of two complete graphs $K$ and $K'$ on $r$ vertices each; then add a perfect matching $M$ between $K$ and $K'$; finally, add a new vertex $i$ that is adjacent to all vertices in $K$. If the algorithm starts with the matching $M$, it will add $i$ to one coalition corresponding to an edge of $M$, and the utility for $i$ is $1/3$. In the optimal partition $\{\{i\} \cup K, K'\}$, every vertex has utility at least $(r - 1)/r$. This gives an approximation ratio of $3 - \frac{3}{r}$.

## 6  Conclusions

In this paper, we initiated an analysis of approximation algorithms for welfare maximization in FHGs. Note that for trees, straightforward dynamic programming techniques can be used to compute a maximum welfare partition. It remains to be seen whether one can obtain fixed parametrized tractable results for parameter treewidth. For utilitarian and egalitarian welfare, we presented approximation algorithms. It remains open whether there are similar approximation bounds for maximum Nash welfare and better bounds for utilitarian and egalitarian welfare.

## References

[Aziz and Savani, 2015] H. Aziz and R. 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. Forthcoming.

[Aziz *et al.*, 2013] H. Aziz, F. Brandt, and H. G. Seedig. Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, 195:316–334, 2013.

[Aziz *et al.*, 2014] H. Aziz, F. Brandt, and P. Harrenstein. Fractional hedonic games. In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 5–12. IFAAMAS, 2014.

[Bilò *et al.*, 2014] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, and L. Moscardelli. Nash stability in fractional hedonic games. In *Proceedings of the 10th International Workshop on Internet and Network Economics (WINE)*, Lecture Notes in Computer Science (LNCS), 2014.

[Bilò *et al.*, 2015] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, and L. Moscardelli:. On the price of stability of fractional hedonic games. In *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1239–1247. IFAAMAS, 2015.

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

[Brandes and Erlebach, 2005] U. Brandes and T. Erlebach, editors. *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science (LNCS)*. Springer-Verlag, 2005.

[Brandl *et al.*, 2015] F. Brandl, F. Brandt, and M. Strobel. Fractional hedonic games: Individual and group stability. In *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1219–1227. IFAAMAS, 2015.

[Gabow, 2009] H. N. Gabow. Data structures for weighted matching and nearest common ancestors with linking. In *Proceedings of the 1st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 434–443. ACM Press, 2009.

[Garey and Johnson, 1979] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[Micali and Vazirani, 1980] S. Micali and V. V. Vazirani. An $O(\sqrt{|V|}|e|)$ algorithm for finding maximum matching in general graphs. In *Proceedings of the 21st Symposium on Foundations of Computer Science (FOCS)*, pages 17–27. IEEE Computer Society Press, 1980.

[Newman, 2004] M. E. J. Newman. Detecting community structure in networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38(2):321–330, 2004.

[Olsen, 2012] M. Olsen. On defining and computing communities. In *Proceedings of the 18th Computing: Australasian Theory Symposium (CATS)*, volume 128 of *Conferences in Research and Practice in Information Technology (CRPIT)*, pages 97–102, 2012.

[Papadimitriou, 2001] C. H. Papadimitriou. Algorithms, games, and the internet. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 749–753. ACM Press, 2001.

[Schaeffer, 2007] S. E. Schaeffer. Graph clustering. *Computer Science Review*, I:27–64, 2007.