

Schelling Games on Graphs*

Edith Elkind¹, Jiarui Gan¹, Ayumi Igarashi²,
Warut Suksompong¹ and Alexandros A. Voudouris¹

¹Department of Computer Science, University of Oxford

²Department of Computer Science and Communication Engineering, Kyushu University

Abstract

We consider strategic games that are inspired by Schelling’s model of residential segregation. In our model, the agents are partitioned into k types and need to select locations on an undirected graph. Agents can be either stubborn, in which case they will always choose their preferred location, or strategic, in which case they aim to maximize the fraction of agents of their own type in their neighborhood. We investigate the existence of equilibria in these games, study the complexity of finding an equilibrium outcome or an outcome with high social welfare, and also provide upper and lower bounds on the price of anarchy and stability. Some of our results extend to the setting where the preferences of the agents over their neighbors are defined by a social network rather than a partition into types.

1 Introduction

In 2015, African Americans constituted 83% of the population of the City of Detroit. At the same time, the neighboring Oakland County was 77% white, and in the city of Dearborn in Detroit metropolitan area about 30% of the residents were Arab Americans. Similar phenomena can be observed in many other major metropolitan areas around the world. In the developed world, the leading cause of such population patterns is not direct discrimination, which is typically illegal; rather, it is the residents themselves who tend to select neighborhoods where their ethnic or social group is well-represented. Schelling [1969, 1971] proposed the following stylized model of this phenomenon: Agents of two different types are placed on a line or on a grid, and are assumed to be happy if at least a fraction τ of the agents within distance w from them are of the same type, for some parameters τ and w ; unhappy agents can either jump to empty positions or swap positions with other agents. Using simple experiments, Schelling showed that, even in cases where the agents are not opposed to integration ($\tau < 1/2$), this behavior leads to almost complete segregation.

In the 50 years since Schelling’s pioneering paper, this segregation model attracted the attention of many researchers, mostly in sociology and economics [Alba and Logan, 1993; Benard and Willer, 2007; Benenson *et al.*, 2009; Clark and Fossett, 2008; Pans and Vriend, 2007; Young, 2001; Zhang, 2004a,b], but recently also in computer science [Barmpalias *et al.*, 2014, 2015; Brandt *et al.*, 2012; Immorlica *et al.*, 2017]. While the early work in this area was mainly empirical, the more recent papers have provided theoretical analysis. In particular, it was proved that the local behavior of unhappy agents is likely to create very large regions consisting of agents of the same type, even when τ is small, i.e., even when the agents themselves are tolerant towards having neighbors of the other type. The vast majority of this work

*This work has been supported by the European Research Council (ERC) under grant number 639945 (ACCORD), and by the KAKENHI Grant-in-Aid for JSPS Fellows number 18J00997.

was based on Schelling’s original model, where agents’ behavior was explained by a simple stochastic model rather than strategic considerations.

An alternative approach is to assume that the behavior of each agent is *strategic*, and exploit tools and techniques from non-cooperative game theory to analyze the induced games. To the best of our knowledge, there are only two papers in the literature that pursue this agenda. Specifically, Zhang [2004b] considered a model with transferable utility where agents prefer to be in a balanced neighborhood. More recently, Chauhan *et al.* [2018] investigated a setting that is closer to Schelling’s motivating scenario, and also incorporates the idea that, in addition to preferences over the composition of their neighborhood, agents may also have preferences over locations. In the model of Chauhan *et al.* [2018], there are two types of agents, and an agent i ’s *happiness ratio* is defined as the fraction of agents of i ’s type among i ’s neighbors. Each agent has two further parameters: a tolerance threshold $\tau \in (0, 1)$ and a preferred location. An agent’s primary goal is to find a location where her happiness ratio exceeds the tolerance threshold; if no such location is available, she aims to maximize her happiness ratio. An agent’s secondary goal is to minimize the distance to her preferred location. To achieve these goals, agents can either swap locations (swap games) or jump to unoccupied locations (jump games). The main contribution of the paper is to identify conditions under which agents are guaranteed to converge to an equilibrium; for instance, the authors establish that in jump games, convergence is guaranteed if agents have no preferred locations and the underlying network is a ring.

1.1 Our contribution

The model of Chauhan *et al.* [2018] makes an important contribution to the literature by enriching Schelling’s model with two additional components: agents who are fully strategic, and location preferences. However, the resulting model of agents’ preferences is quite complex, and, consequently, not easy to analyze: the positive results in the paper are limited to special cases of the utility function and highly regular networks. In this paper, we propose a simpler model that aims to capture the same phenomena and is more amenable to formal analysis.

Specifically, just as in the work of Chauhan *et al.* [2018], in our basic model the agents are partitioned into k types and the set of available locations is represented by an undirected graph, which we will refer to as the *topology*. We also incorporate location preferences in our model; however, instead of assuming that optimizing the distance to the preferred location is the secondary goal of every agent, we assume that agents are either *stubborn*, in which case they stay at their chosen location irrespective of their surroundings, or *strategic*, in which case they aim to maximize their happiness ratio by jumping to an unoccupied location (we do not consider swaps in this paper). Our model captures the fact that, in practice, many residents are unwilling to move to another area even if they are no longer satisfied with the composition of their neighborhood. Importantly, unlike Chauhan *et al.* [2018] or Schelling in his original work, we do not assume that agents have tolerance thresholds; rather, a strategic agent is willing to move as long as there exists another location with a better happiness ratio. Towards the end of the paper (Section 6), we also discuss several variants of this basic model. In particular, we show that some of our positive results extend to the setting where there are no types, but rather the agents are connected by a social network and care about the fraction of their friends (i.e., their neighbors in the social network) among their neighbors in the topology; we refer to the resulting class of games as *social Schelling games*.

The rest of the paper is organized as follows. We define our model in Section 2. Then, in Section 3, we show that for some classes of topologies, such as stars and graphs of maximum degree two, our games always admit a pure Nash equilibrium, i.e., the strategic agents can be assigned to the nodes of the topology so that none of them wants to move to a different location; this result holds even for social Schelling games. In contrast, an equilibrium may fail to exist even if the topology is acyclic and has maximum degree four. In Section 4, we complement this result by presenting a dynamic programming

algorithm that decides whether an equilibrium exists on a tree topology; this algorithm runs in polynomial time if the number of types is bounded by a constant. For more general topologies, we prove that deciding whether an equilibrium exists is an NP-complete problem. Similar hardness and easiness results hold for the problem of maximizing the social welfare (the total utility of all strategic agents). In Section 5, we study the effect of the strategic behavior on the social welfare, by bounding the price of anarchy [Koutsoupias and Papadimitriou, 1999] and the price of stability [Anshelevich *et al.*, 2008]. In particular, we show that even in the absence of stubborn agents it may be impossible to achieve the maximum social welfare in equilibrium. In Section 6 we discuss several variants and extensions of our model and establish some preliminary results for these new models, as well as outline directions for future work.

1.2 Other related work

For an accessible introduction to the Schelling model and a survey of the literature on non-strategic variants, see chapter 4 in the book of Easley and Kleinberg [2010], and the papers by Brandt *et al.* [2012] and Immorlica *et al.* [2017].

Besides the work of Chauhan *et al.* [2018], which was discussed in detail earlier, our model shares a number of properties with hedonic games [Drèze and Greenberg, 1980; Bogomolnaia and Jackson, 2002]; these are games where agents split into coalitions, and each agent’s utility is determined by the composition of her coalition. Specifically, in *fractional hedonic games* [Aziz *et al.*, 2014] the relationships among the agents are described by a weighted directed graph, where the weight of an edge (i, j) is the value that agent i assigns to agent j , and an agent’s utility for a coalition is her average value for the other members in the coalition. If the graph is undirected and all edge weights take values in $\{0, 1\}$, it can be interpreted as a friendship relation; then an agent’s utility in a coalition is computed as the fraction of her friends among the coalition members, which is very similar to how utilities are defined in social Schelling games. On the other hand, the type-based model is closely related to the Bakers and Millers game discussed by Aziz *et al.* [2014]. This connection between Schelling games and hedonic games motivates much of the discussion in Section 6. Of course, a fundamental difference between hedonic games and our setting is that in the former agents derive their utilities from pairwise disjoint coalitions, whereas in our model utilities are derived from (overlapping) neighborhoods.

2 The Model

Let $N = \{1, \dots, n\}$ be a set of $n \geq 2$ agents. The agents are partitioned into $k \geq 2$ different types T_1, \dots, T_k so that $\cup_{j=1, \dots, k} T_j = N$; we write $\mathcal{T} = (T_1, \dots, T_k)$. We say that two agents $i, j \in N$, $i \neq j$, are *friends* if $i, j \in T_\ell$ for some $\ell \in [k]$; otherwise we say that i and j are *enemies*. For each $i \in N$, we denote the set of all friends of agent i by $F(i)$.

A *topology* is an undirected graph $G = (V, E)$ with no self-loops. Each agent in N has to select a node of this graph so that there are no collisions. The agents are classified as either *strategic* or *stubborn*; let R and S denote these sets of agents so that $R \cup S = N$. Stubborn agents care about their location only: each stubborn agent has a preferred node and never moves away from that node. Thus, the preferences of stubborn agents can be described by an injective mapping $\lambda : S \rightarrow V$; for each $i \in S$ the node $\lambda(i)$ is the preferred node of agent i . In contrast, strategic agents do not care about their location, but want to be in a neighborhood that has a large proportion of their friends, and are willing to move to a currently unoccupied node in order to increase their utility.

Formally, given a set of agents $N = R \cup S$ with $|N| = n$, a topology $G = (V, E)$ with $|V| > n$ and a mapping $\lambda : S \rightarrow V$, an *assignment* is a vector $\mathbf{v} = (v_1, \dots, v_n) \in V^n$ such that (1) $v_i = \lambda(i)$ for each $i \in S$ and (2) $v_i \neq v_j$ for all $i, j \in N$ such that $i \neq j$; here, v_i is the node of the topology where agent i is positioned. A node $v \in V$ is *occupied* by agent i if $v = v_i$. For a given assignment

\mathbf{v} and an agent $i \in N$, let $N_i(\mathbf{v}) = \{j \in N : \{v_i, v_j\} \in E\}$ be the set of neighbors of agent i . Let $f_i(\mathbf{v}) = |N_i(\mathbf{v}) \cap F(i)|$ be the number of neighbors of i in \mathbf{v} who are her friends. Similarly, let $e_i(\mathbf{v}) = |N_i(\mathbf{v})| - f_i(\mathbf{v})$ be the number of neighbors of i in \mathbf{v} who are her enemies. Following Chauhan *et al.* [2018], we define the utility $u_i(\mathbf{v})$ of an agent $i \in R$ in \mathbf{v} to be 0 if $f_i(\mathbf{v}) = 0$; otherwise, her utility is defined as the fraction of her friends among the agents in the neighborhood:

$$u_i(\mathbf{v}) = \frac{f_i(\mathbf{v})}{f_i(\mathbf{v}) + e_i(\mathbf{v})}.$$

A tuple $I = (R, S, \mathcal{T}, G, \lambda)$, where R is the set of strategic agents, S is the set of stubborn agents, $\mathcal{T} = (T_1, \dots, T_k)$ is a list of types, $G = (V, E)$ is a topology that satisfies $|V| > |R| + |S|$, and λ is an injective mapping from S to V , is called a *k-typed Schelling game* or *k-typed instance*; let \mathcal{I} be the set of all possible games. We say that an assignment \mathbf{v} is a *pure Nash equilibrium* (or, simply, *equilibrium*) of I if no strategic agent i has an incentive to unilaterally deviate to an empty node z of G in order to increase her utility, i.e., for every $i \in R$ and for every node $z \in V$ such that $z \neq v_j$ for all $j \in R \cup S$ it holds that $u_i(\mathbf{v}) \geq u_i(z, \mathbf{v}_{-i})$, where (z, \mathbf{v}_{-i}) is the assignment obtained by changing the i -th entry of \mathbf{v} to z . Let $\text{EQ}(I)$ denote the set of all equilibria of game I .

The *social welfare* of an assignment \mathbf{v} is defined as the total utility of all strategic agents:

$$\text{SW}(\mathbf{v}) = \sum_{i \in R} u_i(\mathbf{v}).$$

Let $\mathbf{v}^*(I)$ be an assignment that maximizes the social welfare for a given game I ; we refer to it as an *optimal* assignment.

The *price of anarchy* (PoA) of game I with at least one equilibrium is the ratio between the optimal social welfare and the social welfare of the *worst* equilibrium; its *price of stability* (PoS) is defined as the ratio between the optimal social welfare and the social welfare of the *best* equilibrium:

$$\begin{aligned} \text{PoA}(I) &= \sup_{\mathbf{v} \in \text{EQ}(I)} \frac{\text{SW}(\mathbf{v}^*(I))}{\text{SW}(\mathbf{v})}, \\ \text{PoS}(I) &= \inf_{\mathbf{v} \in \text{EQ}(I)} \frac{\text{SW}(\mathbf{v}^*(I))}{\text{SW}(\mathbf{v})}. \end{aligned}$$

The price of anarchy and the price of stability are the suprema of $\text{PoA}(I)$ and $\text{PoS}(I)$ over all $I \in \mathcal{I}$ such that $\text{EQ}(I) \neq \emptyset$, respectively.

3 Existence of Equilibria

In this section, we focus on the existence of equilibria. We warm up by observing that for highly structured topologies such as paths, rings, and stars, there is always at least one equilibrium assignment, and some such assignment can be computed efficiently. This can be shown directly, and also follows from a more general result established in Section 6 (Theorem 6.1).

Theorem 3.1. *Every k-typed Schelling game where the topology is a star or a graph of maximum degree 2 admits at least one equilibrium assignment, which can be computed in polynomial time.*

However, in general, an equilibrium may fail to exist; this holds even if the topology is acyclic and there are no stubborn agents.

Theorem 3.2. *For every $k \geq 2$ there exists a k-typed instance $(R, S, \mathcal{T}, G, \lambda)$ where $S = \emptyset$ and G is a tree that does not admit an equilibrium.*

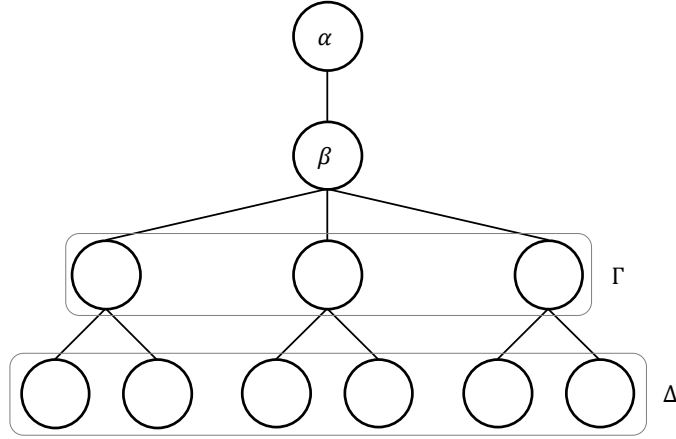


Figure 1: Example of the topology used in the proof of Theorem 3.2 for $k = 2$.

Proof. Given $k \geq 2$, we construct an instance with $2k + 1$ agents per type; the total number of agents is $n = k(2k + 1)$. The topology $G = (V, E)$ is a tree that consists of $|V| = n + 1$ nodes, which are distributed over four layers. Specifically, the tree has a root α , which has one child β . Node β has $2k - 1$ children; we denote the set of its children by Γ . Each node in Γ has k children, which are the leaves of the tree; we denote the set of all leaves by Δ . Figure 1 depicts the topology for $k = 2$. Now, assume that there is an equilibrium assignment; note that exactly one node is left empty. We consider four cases depending on the location of the empty node.

Node α is empty. Assume that the agent occupying node β is of type T . Then, since there are $2k$ other agents of type T and there are only $2k - 1$ nodes in Γ , there must exist some subtree rooted at a node in Γ that contains both agents of type T and agents that belong to other types. Then an agent of type T from this subtree has an incentive to deviate to α .

Node β is empty. Assume that the agent occupying node α is of type T ; note that her utility is 0. If she does not have an incentive to deviate to β , it follows that no agent of type T occupies a node in Γ . But then there is an agent of type T who occupies a node in Δ ; as her parent is not of type T , her utility is 0, and she can increase it by moving to β .

Some node $\gamma \in \Gamma$ is empty. Consider the agents occupying the children of γ ; note that their utility is 0. If at least two of them have the same type, each of them has an incentive to deviate to γ in order to increase her utility to at least $\frac{1}{k}$. If all of them have different types, then there is exactly one agent of each type in this set. In particular, there is an agent i who has the same type as the agent occupying β ; then i can move to γ to increase her utility.

Some node $\delta \in \Delta$ is empty. Let γ denote the parent of this node, and suppose that γ is occupied by an agent i of type T . We say that an agent j of type T is *hungry* if $j \neq i$ and j is adjacent to at least one agent of a different type; note that a hungry agent has an incentive to deviate to δ . We claim that at least one agent is hungry. Indeed, if β is occupied by an agent j of type T , then either j is hungry or every agent in $\Gamma \setminus \{\gamma\}$ is hungry. If the agent in β is not of type T and there is an agent ℓ of type T in $\Gamma \setminus \{\gamma\}$, then ℓ is hungry. Finally, if no agent in $\Gamma \setminus \{\gamma\}$ is of type T , there exists a leaf node not in γ 's subtree that is occupied by an agent r of type T ; r is then hungry.

The proof is complete. □

4 Computational Complexity

We now turn our attention to the computational complexity of k -typed Schelling games. The main result of this section is that finding an equilibrium assignment is computationally intractable.

Theorem 4.1. *For every $k \geq 2$, given a k -typed Schelling game I , it is NP-complete to decide whether I admits an equilibrium assignment. The hardness result holds even if all strategic agents belong to the same type.*

Proof. We give a proof for $k = 2$; it is straightforward to extend it to $k \geq 2$. We will use a reduction from the CLIQUE problem. An instance of this problem is an undirected graph $H = (X, Y)$ and an integer s ; it is a yes-instance if H has a complete subgraph of size s . Given an instance $\langle H, s \rangle$ of CLIQUE with $H = (X, Y)$, we assume without loss of generality that $s \geq 5$ and construct an instance of our problem as follows:

- There are two agent types: red and blue.
- There are s strategic red agents; all remaining agents are stubborn. We will describe the stubborn agents and their locations when defining the topology.
- The topology $G = (V, E)$ consists of three disjoint components G_1, G_2 , and G_3 such that
 - $G_1 = (V_1, E_1)$, where $V_1 = X \cup W$, $|W| = s - 2$, $E_1 = Y \cup \{\{v, w\} : v \in X, w \in W\}$. There is a stubborn blue agent at each node $w \in W$;
 - G_2 is a complete bipartite graph with parts L and R , $|L| = s - 2$, $|R| = 4s$. Of the $4s$ nodes in R , $2s + 1$ nodes are occupied by red agents and $2s - 1$ nodes are occupied by blue agents;
 - G_3 has three empty nodes, denoted x, y , and z , and 121 nodes — 41 red and 80 blue — occupied by stubborn agents. There is an edge between nodes x and y ; also, x is connected to 1 red agent and 2 blue agents; y is connected to 41 red agents and 80 blue agents, and z is connected to 5 red agents and 7 blue agents.

Note that a strategic red agent obtains a utility of $\frac{2s+1}{4s} = \frac{1}{2} + \frac{1}{4s}$ by choosing an available node in G_2 and a utility of $\frac{5}{12}$ by choosing z . If she chooses x , her utility is $\frac{1}{3}$ if y is unoccupied and $\frac{1}{2}$ otherwise. Similarly, if she chooses y , her utility is $\frac{41}{121}$ if x is unoccupied and $\frac{42}{122}$ otherwise; note that $\frac{1}{3} < \frac{41}{121} < \frac{42}{122} < \frac{5}{12}$.

Now, suppose that G contains a clique of size s . If strategic red agents occupy the nodes of that clique, the utility of each such agent is $\frac{s-1}{(s-1)+(s-2)} = \frac{1}{2} + \frac{1}{4s-6}$. Thus, by our choice of parameters, no agent has a profitable deviation.

On the other hand, suppose that G does not contain a clique of size s . Assume for the sake of contradiction that there is an equilibrium assignment \mathbf{v} .

Suppose first that in \mathbf{v} some strategic agents are located in G_1 . It cannot be the case that each of them is adjacent to $s - 1$ friends, as this would mean that their locations form a clique of size s . Hence, at least one of these agents is adjacent to at most $s - 2$ friends. As this agent is also adjacent to the $s - 2$ stubborn blue agents in W , her utility is at most $\frac{1}{2}$. By our choice of parameters, all unoccupied nodes of G_2 offer a higher utility, namely, $\frac{1}{2} + \frac{1}{4s}$. Thus, if there are strategic agents in G_1 , all $s - 2$ nodes of G_2 that are available to strategic agents must be occupied. But then, there are at most two strategic agents in G_1 , which means that their utility is at most $\frac{1}{s-1} < \frac{1}{3}$ (recall that we assume that $s \geq 5$). This leads to a contradiction, as these strategic agents would be better off moving to G_3 where their utility would be at least $\frac{1}{3}$.

Therefore, in equilibrium no strategic agent can be located at a node of G_1 . Further, since all unoccupied nodes of G_2 always offer more utility than any unoccupied nodes of G_3 can offer, in equilibrium

all nodes of G_2 are occupied, and the two remaining strategic agents must be in G_3 , with one of x , y , and z left empty.

Suppose that z is empty. Then the agent located at y can increase her utility from $\frac{42}{122}$ to $\frac{5}{12}$ by moving to z , a contradiction. If y is empty, the agent located at x can increase her utility from $\frac{1}{3}$ to $\frac{41}{121}$ by moving to y , a contradiction. Finally, if x is empty, the agent located at z can increase her utility from $\frac{5}{12}$ to $\frac{1}{2}$ by moving to x , a contradiction. As we have exhausted all possibilities, it follows that if G does not have a clique of size s , then there is no equilibrium assignment. \square

The proof of Theorem 4.1 can be adapted to show that maximizing social welfare in Schelling games is NP-hard as well.

Theorem 4.2. *For every $k \geq 2$, given a k -typed Schelling game I and a rational value s , it is NP-complete to decide whether I admits an assignment with social welfare at least s . The hardness result holds even if $k = 2$, all strategic agents belong to one type, and the other type consists of a single stubborn agent.*

Proof. We modify the reduction in the proof of Theorem 4.1 by removing the gadgets G_2 and G_3 and replacing the set W with a single node w . That is, given an instance $\langle H, s \rangle$ of CLIQUE, we construct an instance of our social welfare maximization problem as follows:

- There are two agent types: red and blue.
- There are s strategic red agents and one stubborn blue agent.
- The topology $G = (V, E)$ is defined so that $V = X \cup \{w\}$ and $E = Y \cup \{\{v, w\} : v \in X\}$.
- The single stubborn blue agent is positioned at node w .

Note that the utility of a red agent p in an assignment \mathbf{v} is $\frac{r}{r+1}$, where r is the number of red agents that p is adjacent to in \mathbf{v} ; the function $\frac{r}{r+1}$ is increasing in r and we have $r \leq s - 1$ for any assignment. Hence, the social welfare of $s - 1$ can be achieved if and only if the red agents can be placed in G so that each agent is adjacent to every other red agent, in which case the utility of each strategic agent is $\frac{s-1}{s}$; this is possible if and only if H contains a clique of size s . \square

On the positive side, for small k we can efficiently decide whether an equilibrium exists if the topology G is a tree. Our algorithm is based on dynamic programming: it selects an arbitrary node of G to be the root, and then for every node v of G , it fills out a multidimensional table whose dimension is linear in the number of types, proceeding from the leaves to the root. It decides whether the given instance admits an equilibrium by scanning the table at the root node. The details of the algorithm are given in the appendix.

Theorem 4.3. *Given a k -typed Schelling game I with n agents, where the topology G is a tree, we can decide whether I admits an equilibrium (and compute one if it exists) in time $\text{poly}(n^k)$, i.e., this problem lies in the complexity class XP with respect to the number of types k .*

By slightly modifying our algorithm, we can compute an assignment that maximizes the social welfare, either among all assignments or among equilibria.

Corollary 4.4. *Given a k -typed Schelling game, where G is a tree, the problems of computing an equilibrium with maximum social welfare or a socially optimal assignment are in XP with respect to k .*

We have not been able to determine whether the problem of computing an equilibrium assignment is fixed-parameter tractable with respect to the number of types; we leave this question for future work.

5 Price of Anarchy and Stability

In this section, we investigate the loss in social welfare caused by strategic behavior, as measured by the price of anarchy and the price of stability; unless otherwise specified, the topology is assumed to be a connected graph.¹ We start by establishing bounds on the price of anarchy for instances with no stubborn agents.

Theorem 5.1. *For k -typed Schelling games with no stubborn agents and n strategic agents, the PoA*

- *can be unbounded for each $k \geq 2$;*
- *is $\Theta(n)$ when there are at least two agents per type;*
- *is $k + o(1)$ if each type has the same number of agents.*

We prove each statement separately.

Lemma 5.2. *For k -typed Schelling games with no stubborn agents and n strategic agents, the price of anarchy can be unbounded for each $k \geq 2$.*

Proof. Consider an instance with two agents of type T_1 and one agent of every other type T_i , $i \geq 2$. The topology is a star with $k + 2$ nodes. Then, any assignment where the center node is occupied by an agent of type T_i with $i \geq 2$, is an equilibrium with zero social welfare. In contrast, any assignment where the center node is occupied by an agent of type T_1 is again an equilibrium, but the social welfare is now strictly positive (since the two agents of type T_1 are connected to each other). \square

Lemma 5.3. *For k -typed Schelling games with no stubborn agents and n strategic agents, in which there are at least two agents per type, the price of anarchy is $\Theta(n)$ for every $k \geq 2$.*

Proof. For the lower bound, consider an instance with $|T_i| = 2$ for $i = 1, \dots, k - 1$ and $|T_k| = n - 2(k - 1)$. The topology is a star with $n + 1$ nodes. Then any assignment where an agent that belongs to one of the first $k - 1$ types is at the center of the star is an equilibrium with social welfare $1 + 1/(n - 1) \leq 2$, while for any assignment where an agent of type T_k is at the center of the star the social welfare is $n - 2(k - 1) - 1 + 1/(n - 1) \geq n - 2k$. Hence, the price of anarchy is at least $n/2 - k$.

For the upper bound, consider a k -typed instance with $n_i \geq 2$ agents of each type T_i , so that $n = \sum_{i \in [k]} n_i$. We will show that the social welfare of any equilibrium assignment is at least 1. This implies our bound on the price of anarchy, since the optimal social welfare is at most n .

Let \mathbf{v} be an arbitrary equilibrium assignment. Recall that we assume that the number of available nodes exceeds the number of agents and the topology is connected, so there must exist some empty node v with at least one non-empty neighbor. Suppose that v is connected to x_i agents of type T_i , for $i \in [k]$, and let $s = \sum_{i \in [k]} x_i$. By deviating to v , an agent of type T_i would get utility $\frac{x_i}{s}$ if she is not connected to v , and utility $\frac{x_i - 1}{s - 1}$ otherwise; again, for readability, we use the convention that $\frac{0}{0} = 0$. Since at equilibrium no agent has an incentive to deviate, her utility is at least the utility she would get by deviating to v . Therefore, the social welfare at equilibrium is at least

$$\begin{aligned} \text{SW}(\mathbf{v}) &\geq \sum_{i \in [k]} \left((n_i - x_i) \frac{x_i}{s} + x_i \frac{x_i - 1}{s - 1} \right) \\ &\geq \frac{1}{s} \sum_{i \in [k]} (n_i - 1)x_i \geq 1, \end{aligned}$$

where the last inequality holds since $n_i \geq 2$ for every $i \in [k]$. This completes the proof. \square

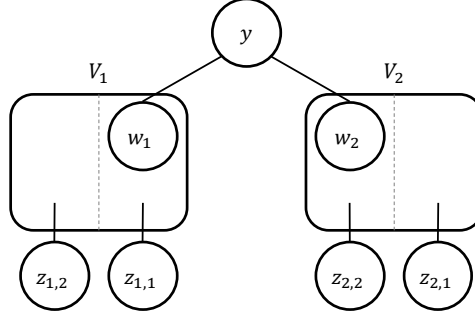


Figure 2: The topology used in the proof of Theorem 5.4 for $k = 2$.

Lemma 5.4. *For k -typed Schelling games with no stubborn agents and n strategic agents, the PoA is $k + o(1)$ if each type has the same number of agents.*

Proof. For the lower bound, fix $k \geq 2$, let $\ell \geq 2$ be a parameter and consider an instance with $k(\ell + 1)$ agents per type; altogether there are $n = k^2(\ell + 1)$ agents. The topology consists of $n + 1$ nodes and is defined as follows. There are k cliques V_1, \dots, V_k of size $k\ell$ each, and a node y . In each clique V_i there is a special node w_i that is connected to y . Also, for each $i \in [k]$ there are k auxiliary nodes $z_{i,1}, \dots, z_{i,k}$; each of these nodes is connected to a distinct set of ℓ nodes in V_i . Let $z_{i,i}$ be the auxiliary node that is connected to w_i . Figure 2 illustrates this topology for $k = 2$.

There is an optimal assignment where all $k(\ell + 1)$ agents of type T_i are placed at the nodes of clique V_i and the corresponding auxiliary nodes, so that all agents are connected only to agents of the same type and have maximum utility 1. Therefore, the optimal social welfare is $k^2(\ell + 1)$.

In contrast, consider the following equilibrium assignment: node y is empty, and for each $i, j \in [k]$ all ℓ nodes in V_i that are connected to the auxiliary node $z_{i,j}$ as well as $z_{i,j}$ itself are occupied by agents of type T_j . Since node y is connected to k nodes that are occupied by agents of different types, any agent would get utility $1/k$ by deviating there. No agent occupying an auxiliary node has an incentive to deviate since she is connected only to agents of her type. For every clique, each agent is connected to exactly ℓ agents of the same type ($\ell - 1$ of whom occupy nodes of the clique and one that occupies the corresponding auxiliary node) and $(k-1)\ell$ agents of different type; thus, her utility is $1/k$. Consequently, no agent has an incentive to deviate, and the social welfare is $k \cdot k\ell \cdot \frac{1}{k} + k^2 = k(\ell + k)$. Hence, the PoA is at least $\frac{k\ell+k}{\ell+k}$; this expression becomes arbitrarily close to k as ℓ grows.

For the upper bound, consider an arbitrary instance with n agents and $k \geq 2$ types so that there are n/k agents per type. We will show that the social welfare of any equilibrium assignment is at least $n/k - 1$. The bound on the PoA then follows, since the optimal social welfare is at most n .

Recall that we assume that the number of available nodes exceeds the number of agents and the topology is connected, so there must exist some empty node v with at least one non-empty neighbor. Suppose that v is connected to x_i agents of type T_i , for $i \in [k]$, and let $s = \sum_{i \in [k]} x_i$. Consider an agent of type T_i . A deviation to v would give her utility $\frac{x_i}{s}$ if she is not connected to v , and utility $\frac{x_i-1}{s-1}$ otherwise (for readability we use the convention that $\frac{0}{0} = 0$). Since at equilibrium no agent has any incentive to deviate, her utility is at least the utility she would get by deviating to v . Therefore, the social welfare at equilibrium is at least

$$\text{SW}(\mathbf{v}) \geq \sum_{i \in [k]} \left(\left(\frac{n}{k} - x_i \right) \frac{x_i}{s} + x_i \frac{x_i - 1}{s - 1} \right)$$

¹We can easily observe that the price of anarchy can be unbounded for not connected topologies. For instance, consider a topology with one isolated node and a connected subgraph $W \subset V$ such that $|W| = n$. Then, any assignment of the n agents at the nodes of W is an equilibrium. Hence, there might exist equilibria with zero social welfare.

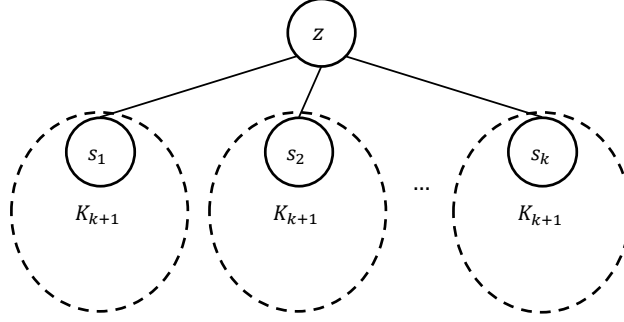


Figure 3: The topology used in the proof of Lemma 5.7.

$$\geq \frac{1}{s} \sum_{i \in [k]} \left(\left(\frac{n}{k} - x_i \right) x_i + x_i(x_i - 1) \right) = \frac{n}{k} - 1.$$

The proof is complete. \square

In the setting considered in Theorem 5.1 the PoA improves significantly if we require each type to have the same number of agents. In the presence of stubborn agents, to ensure that the price of anarchy does not depend on the number of agents n , we additionally require that this constraint holds both for strategic and for stubborn agents.

Theorem 5.5. *For k -typed Schelling games with n agents the PoA*

- *is $\Omega(n)$ for each $k \geq 2$ even if there is an equal number of agents per type;*
- *is $k + o(1)$ if each type has the same number of strategic agents and the same number of stubborn agents.*

Again, we prove each statement separately.

Lemma 5.6. *For k -typed Schelling games with no stubborn agents and n strategic agents, the price of anarchy is $\Omega(n)$ for each $k \geq 2$, even if there is an equal number of agents per type.*

Proof. Pick a positive integer ℓ and consider an instance with $n = k\ell$ agents such that there are ℓ strategic agents of type T_1 , one strategic agent and $\ell - 1$ stubborn agents of type T_2 , and ℓ stubborn agents of type T_i for each $i = 3, \dots, k$. The topology is a star with $n + 1$ nodes, and all stubborn agents occupy leaf nodes. Then, any assignment where the strategic agent of type T_2 occupies the center node is an equilibrium with social welfare $\frac{\ell-1}{k\ell-1} < \frac{1}{k}$, while the social welfare of any assignment where the center node is occupied by an agent of type T_1 is $\ell - 1 + \frac{\ell-1}{k\ell-1} > \ell - 1$. Hence, the price of anarchy is at least $k(\ell - 1) = n - k$. \square

Lemma 5.7. *For k -typed instances, where all types have the same number of strategic agents and the same number of stubborn agents, the price of anarchy is $k + o(1)$.*

Proof. We first establish the lower bound. Suppose that k is odd, and consider an instance with k types of agents such that there are k strategic agents and one stubborn agent s_i per type T_i . The number of strategic agents is $s = k^2$. The topology is depicted in Figure 3 and consists of k cliques of size $k + 1$, which are connected to each other via an auxiliary node z . The stubborn agent s_i of type T_i occupies the node of the i -th clique that is adjacent to z .

In an optimal assignment all strategic agents of type T_i occupy the nodes of the i -th clique: this ensures that the utility of each strategic agent is 1 and the social welfare is equal to s . In contrast,

consider an assignment where the auxiliary node z is left empty, and the i -th clique includes one agent of type T_i and $\frac{k-1}{2}$ pairs of agents of different types.² This is an equilibrium since all strategic agents have utility $1/k$, which is exactly the utility they would get by deviating to z . Therefore, the social welfare achieved by this equilibrium assignment is $\frac{s}{k}$, and the price of anarchy is at least k .

When k is even, we can modify the instance as follows. For each $i \in [k]$, there are $k-1$ strategic agents and one stubborn agent per type T_i . The topology consists of $k-1$ cliques of size k , which are connected to each other via an auxiliary node z , together with k dummy nodes each connected to a single node occupied by a stubborn player. For $i \in [k-1]$, the stubborn agent of type T_i occupies the node of the i -th clique that is adjacent to z , and the stubborn agent of type T_k occupies one of the dummy nodes.

If all strategic agents of type T_i , for $i \in [k-1]$, occupy the nodes of the i -th clique, and the agents of type T_k occupy the dummy nodes, then the social welfare is equal to $(k-1)^2$. On the other hand, there is an equilibrium where agents of type T_k occupy dummy nodes and agents of other types are distributed over the cliques as in the equilibrium for odd k . Then, for $i \in [k-1]$, the utility of each strategic agent of type T_i is $\frac{1}{k-1}$. The social welfare in this case is $k-1$, and so the price of anarchy is at least $k-1$.

For the upper bound, consider an arbitrary k -typed instance with t strategic and ℓ stubborn agents per type, for some integers $t > 0$ and $\ell \geq 0$. We will show that the social welfare of any equilibrium assignment is at least $t-1$. The bound then follows since the utility of every strategic agent is at most 1, meaning that the optimal social welfare is at most kt .

Let \mathbf{v} be an arbitrary equilibrium assignment. Since the number of available nodes exceeds the number of agents and the topology is connected, there must exist some empty node v with at least one non-empty neighbor. Suppose that v is connected to x_i agents of type T_i , for $i \in [k]$, and x_i^R of them are strategic. Also, let $s = \sum_{i \in [k]} x_i$. Now, consider a strategic agent of type T_i . A deviation to v would give her utility $\frac{x_i}{s}$ if she is not connected to v , and utility $\frac{x_i-1}{s-1}$ otherwise; again, for readability, we use the convention $\frac{0}{0} = 0$. Since at equilibrium no strategic agent has any incentive to deviate, her utility is at least the utility she would get by deviating to v . Therefore, the social welfare at equilibrium is at least

$$\begin{aligned} \text{SW}(\mathbf{v}) &\geq \sum_{i \in [k]} \left((t - x_i^R) \frac{x_i}{s} + x_i^R \frac{x_i - 1}{s - 1} \right) \\ &\geq \frac{1}{s} \sum_{i \in [k]} \left(tx_i - x_i^R \right) \geq t - 1, \end{aligned}$$

where the last inequality follows since $x_i^R \leq x_i$. The proof is complete. \square

Finally, we show that in Schelling games even the best equilibrium need not be socially optimal, even if all agents are strategic.³

Theorem 5.8. *For k -typed Schelling games the PoS*

- *can be unbounded for each $k \geq 2$;*
- *is at least 3 for each even $k \geq 2$, if there is the same number of stubborn agents per type;*
- *is at least $34/33$ for each $k \geq 2$, even in the absence of stubborn agents.*

The proof of the above theorem follows by the next three lemmas.

²To see how such an assignment can be computed, split $k-1$ agents of type T_i into pairs and think of each pair as one ball of color T_i and each clique as a bin. Then, there are $\frac{k-1}{2}$ balls of each color, which must be placed in k bins so that each bin contains $\frac{k-1}{2}$ balls of different color. To accomplish this, we can order the balls so that balls of type T_i appear in positions $i, k+i, 2k+i, \dots$; hence, we can simply put the first $\frac{k-1}{2}$ balls in the first bin, the next $\frac{k-1}{2}$ balls in the second bin, etc.

³Note that the assumption of a connected topology is no longer necessary for meaningful bounds on the price of stability, since the PoS deals with the best-case equilibrium assignment rather than the worst-case one.

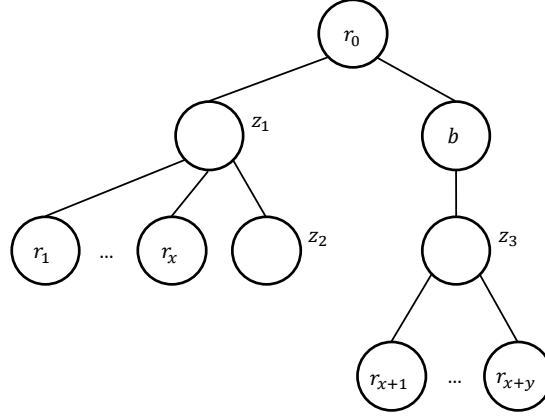


Figure 4: The topology used in the proof of Lemma 5.9 for $k = 2$.

Lemma 5.9. *For k -typed Schelling games, the price of stability can be unbounded for each $k \geq 2$.*

Proof. We prove this lemma only for $k = 2$; our instance can then be generalized to any number of types by adding isolated nodes in the topology which are occupied by stubborn players of different types.

Let $\varepsilon > 0$ be a parameter such that $x = \frac{2}{\varepsilon} - 2$ and $y = \frac{1}{\varepsilon} - 1$ are integer numbers. Consider an instance with $x + y + 1$ stubborn red agents, one stubborn blue agent, and two strategic blue agents. The topology and the placement of the stubborn agents is depicted in Figure 4. There are only three possible assignments depending on which pair of nodes (out of the three available) the two strategic blue agents occupy.

We claim that the only equilibrium assignment is the one where node z_1 is left empty with social welfare $\frac{1}{y+1} = \varepsilon$. First, observe that this assignment is indeed an equilibrium since no strategic agent has any incentive to deviate: node z_1 can give utility $\frac{1}{x+2} = \frac{\varepsilon}{2}$ to the agent occupying node z_3 and utility 0 to the agent occupying node z_2 . Since the two agents get utility $\frac{1}{y+1} = \varepsilon$ and 0, respectively, none of them has any incentive to deviate. To verify the uniqueness of the equilibrium, observe that in the other two possible assignments there exists a strategic agent that can deviate to the empty node in order to increase her utility from $\frac{\varepsilon}{2}$ to ε in case the strategic blue agents are connected, or from ε to 1 in case the strategic blue agents are not connected.

In contrast, the assignment according to which the strategic blue agents are connected to each other (by occupying nodes z_1 and z_2) is the optimal one with social welfare $1 + \frac{\varepsilon}{2}$. Therefore, the price of stability is at least $\frac{1}{\varepsilon} + \frac{1}{2}$, which tends to infinity as ε tends to zero. \square

Lemma 5.10. *For k -typed Schelling games, the price of stability is at least 3, for each even $k \geq 2$, if there is the same number of stubborn agents per type.*

Proof. For simplicity, we will prove the lemma for $k = 2$. Let $x \geq 1$ be a parameter and consider an instance with two types of agents (red and blue) such that there are $x + 1$ stubborn red agents, $x + 1$ stubborn blue agents, and two strategic blue agents. The topology and the placement of the stubborn agents are depicted in Figure 5. There are only three possible assignments depending on which pair of nodes (out of the three available) the two strategic blue agents occupy.

We claim that the only equilibrium assignment is the one where node y is left empty with social welfare $\frac{x+1}{2x+1}$. First, observe that this assignment is indeed an equilibrium since no strategic agent has any incentive to deviate: node y can give utility $1/2$ to the agent occupying node z and utility 0 to the agent occupying node w . Since the two agents get utility $\frac{x+1}{2x+1} > 1/2$ and 0, respectively, none of them has any incentive to deviate. To verify the uniqueness of the equilibrium, observe that in the other two possible assignments there exists a strategic agent that can deviate to the empty node in order to increase

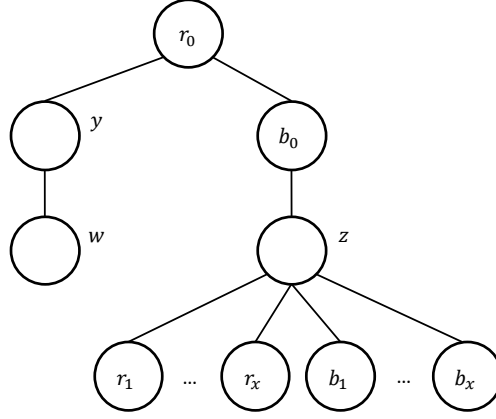


Figure 5: The topology used in the proof of Lemma 5.10 for $k = 2$.

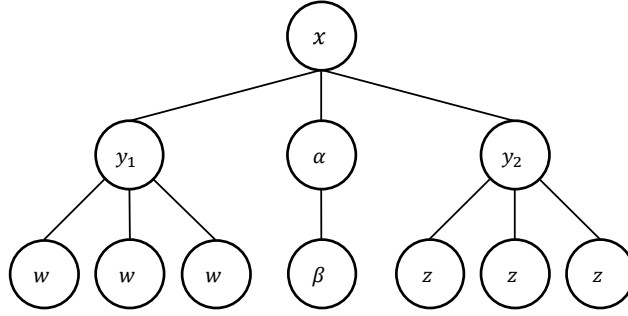


Figure 6: The topology used in the proof of Lemma 5.11 for $k = 2$.

her utility from $1/2$ to $\frac{x+1}{2x+1} > 1/2$ in case the strategic blue agents are connected, or from $\frac{x+1}{2x+1}$ to 1 in case the strategic blue agents are not connected.

In contrast, the assignment according to which the strategic blue agents are connected to each other (by occupying nodes y and w) is the optimal one with social welfare $3/2$. Therefore, the price of stability is at least $\frac{3(2x+1)}{2(x+1)}$, which tends to 3 as x becomes arbitrarily large.

The bound can easily be extended to the case of k types (for even $k \geq 2$) by replicating $k/2$ times the whole instance and connecting the topologies via an empty node. \square

Lemma 5.11. *For k -typed Schelling games, the price of stability is at least $34/33$, for each $k \geq 2$, even in the absence of stubborn agents.*

Proof. For the sake of simplicity, we prove the lemma for $k = 2$, for which the desired lower bound is $34/33$; we will discuss how to generalize our construction to $k > 2$ at the end of the proof.

Consider an instance with two types of agents (red and blue) such that there are five red and five blue agents; the topology is depicted in Figure 6.

Let \mathbf{v} be the following assignment: node x , node y_1 and all three w -type nodes are occupied by red agents, while node y_2 , all z -type nodes and node β are occupied by blue agents. One can easily verify that \mathbf{v} is an equilibrium since no agent has any incentive to deviate to the empty node α ; the social welfare is $\text{SW}(\mathbf{v}) = 33/4$.

Let \mathbf{v}' be the following assignment: node x , node y_1 and all three w -type nodes are occupied by red agents, while node y_2 , two of the z -type nodes, node α and node β are occupied by blue agents. This is not an equilibrium assignment since the blue agent occupying α has utility $1/2$ and hence has an

incentive to deviate to the empty z -type node in order to increase her utility to 1. However, it achieves an improved social welfare of $SW(\mathbf{v}') = 34/4$.

In order to complete the proof, we need to argue that \mathbf{v} is an equilibrium with the maximum social welfare. To this end, we establish some properties of equilibrium assignments.

- Node x must be occupied. Assume otherwise that x is left empty. If nodes y_1 , α and y_2 are occupied by agents of the same type, then at least one of them will be connected to some agent of the other type, and therefore will have an incentive to deviate to x in order to connect only to agents of the same type. Hence, without loss of generality (due to symmetry), at y_1 , α and y_2 there are two red agents and one blue agent. Trivially, the blue agent cannot be connected to any red agents, since otherwise any such red agent would get zero utility and have an incentive to deviate to x in order to increase her utility to $2/3$. Since there are four remaining blue agents, at least one of them must be connected to one of the two red agents occupying nodes at the second layer. Hence, this blue agent gets zero utility and has an incentive to deviate to x in order to increase her utility to $1/3$.
- Nodes y_1 and y_2 must be occupied. Assume otherwise that one of these nodes, say y_1 , is left empty, while node x is occupied by a red agent (without loss of generality). If all w -type nodes are occupied by agents of the same type, then all these agents get zero utility and have an incentive to deviate to y_1 in order to increase their utility to at least $2/3$. So, agents of both types must appear at the w -type nodes. But then, the red such agent has an incentive to deviate to y_1 in order to increase her utility from zero to at least $1/3$.
- Agents of both types must appear at the nodes of the second layer. Assume otherwise that only agents of the same type appear at these nodes, while node x is occupied by a red agent (without loss of generality). Let us further assume that y_1 , α and y_2 are all occupied by blue agents. Then, since the empty node is one of those at the third layer, two of the blue agents occupying nodes y_1 , α and y_2 have an incentive to deviate in order to increase their utility from strictly less than 1 (since they are connected to the red agent occupying node x) to 1. In case y_1 , α and y_2 are all occupied by red agents, then all blue agents occupy nodes at the third layer, meaning that at least two of the red agents occupying y_1 , α and y_2 get utility strictly less than 1, and only one of them can be connected to the empty node. Hence, the other such red agent has an incentive to deviate to the empty node and increase her utility to 1. Therefore, the empty node must be one of those at the second layer.

Since y_1 and y_2 are occupied, α has to be the empty node. If β is occupied by a red agent, then this agent gets zero utility and has an incentive to deviate to α in order to connect to the red agent occupying x . Hence, β must be occupied by a blue agent. If y_1 and y_2 are occupied by blue agents, then either one of them is connected only to red agents or both are connected to three red agents (including the one at x) and one blue agent. In any case, at least one of them has an incentive to deviate to α and increase her utility from 0 or $1/4$ to $1/2$. So, nodes y_1 and y_2 must be occupied by red agents. But then, all blue agents occupy nodes at the third layer, get zero utility, and the four of them that are connected to the red agents occupying y_1 and y_2 have an incentive to deviate to α in order to increase their utility to $1/2$.

- The type of agents that appears at node x can appear only once more at nodes y_1 , α or y_2 . Assume otherwise that x is occupied by a red agent (without loss of generality) and two nodes at the second layer, say y_1 and α , are occupied by red agents as well; the case where y_1 and y_2 are occupied by red players is similar. By the discussion above, y_2 must then be occupied by a blue agent. Observe that since there are four remaining blue agents, one of them (agent p) has to be connected to one of the red agents occupying y_1 and α . Trivially, none of the z -type nodes can be empty, since this

would give an incentive to p to deviate there in order to increase her utility from zero to 1. But then, this means that one of the w -type nodes or node β is empty, thus giving an incentive to the red agent occupying x to deviate in order to increase her utility from $2/3$ to 1.

Given the above structural properties, there can only be two equilibria (and two more symmetric ones, produced by exchanging agents of different types):

- Nodes x and y_1 are occupied by red agents, while nodes α and y_2 are occupied by blue agents; the assignment for the nodes of the third layer is then trivially defined. Such an equilibrium has social welfare $97/12$.
- Nodes x and y_1 are occupied by red agents, node y_2 is occupied a blue agent, and node α is empty; the assignment for the nodes of the third layer is trivially defined so that node β is occupied by the last blue agent that gets zero utility. Such an equilibrium has social welfare $33/4$.

Hence, the second type of equilibrium assignment is the one with maximum social welfare, and the lower bound on the price of stability follows.

We can generalize the above instance to $k > 2$ agent types as follows. Let G_0 be the topology used in the above instance for two agent types. Now, consider an instance where the topology consists of G_0 and $k - 2$ isolated nodes. There are five agents of type T_1 , five agents of type T_2 , and one agent per type T_i for $i \in \{3, \dots, k\}$; the agents of type T_1 and T_2 correspond to the red and blue agents in the instance for $k = 2$.

Observe that the agents of type T_i for $i \in \{3, \dots, k\}$ get zero utility in any possible assignment, since they are unique of their type. Consequently, even though there are many equilibrium assignments where these agents occupy nodes of G_0 , none of these equilibria achieve higher utility than the ones where these agents occupy isolated nodes, and agents of types T_1 and T_2 occupy the nodes of G_0 . Consequently, following the same reasoning as in the above instance for $k = 2$, we can conclude that the price of stability is at least $34/33$. \square

6 Variants and Extensions

Throughout this paper, we focused on a setting where agents are classified into k types and their utilities are defined by the proportion of their friends among their neighbors. In this section, we introduce three variants of this model and briefly discuss some preliminary results; a more thorough investigation of these alternative models is left for future work.

6.1 Schelling games with social networks

In k -typed Schelling games, the friendship relation is defined by types: an agent's set of friends consists of all agents of the same type. One can also consider a more general friendship relation, defined by an arbitrary undirected graph \mathcal{G} with vertex set N , which we will refer to as the *social network*: the set of friends of agent i consists of all neighbors of i in \mathcal{G} . We refer to the resulting class of games as *social Schelling games*.

By definition, k -typed Schelling games form a subclass of social Schelling games: a k -typed game corresponds to a social network consisting of k cliques. Hence, our next theorem implies Theorem 3.1 in Section 3.

Theorem 6.1. *Every social Schelling game where the topology is a star or a graph of maximum degree 2 admits at least one equilibrium assignment, which can be computed in polynomial time.*

Proof. Consider a social Schelling game with a set of agents $N = R \cup S$, where R is the set of strategic agents and S is the set of stubborn agents, a topology $G = (V, E)$, a social network \mathcal{G} , and a function $\lambda : S \rightarrow V$ that describes the locations of the stubborn agents; for each $i \in R$, let $F(i)$ denote the set of nodes of \mathcal{G} that are adjacent to i .

Suppose that G is a star with center v . Consider an assignment \mathbf{v} such that $v = v_i$ for some $i \in R \cup S$. All strategic agents are indifferent among the leaves, so no agent in $R \setminus \{i\}$ has a beneficial deviation. Now, consider agent i . If i is stubborn, she cannot deviate; if i is strategic, she does not want to deviate, as any leaf node would give her zero utility. Hence, \mathbf{v} is an equilibrium.

Now, suppose that $G = (V, E)$ is a graph of maximum degree 2. Our analysis for this case is inspired by Theorem 6 in the work of Chauhan *et al.* [2018]. For each $v \in V$, let $\deg(v)$ denote the degree of a vertex v in G . Given an assignment \mathbf{v} , for each edge $e = \{v, w\}$, we define

$$\phi(\mathbf{v}, e) = \begin{cases} 1 & \text{if } w = v_i, v = v_j \text{ and } i \in F(j) \\ 0 & \text{if } w = v_i, v = v_j \text{ and } i \notin F(j) \\ \frac{1}{3} & \text{if } v \text{ or } w \text{ is unoccupied in } \mathbf{v}. \end{cases}$$

Let $\Phi(\mathbf{v}) = \sum_{e \in E} \phi(\mathbf{v}, e)$. We claim that $\Phi(\mathbf{v})$ is an ordinal potential function for our setting, i.e., if an agent deviates to increase her utility, the potential function increases.

To see this, consider an assignment \mathbf{v} and an agent i with $v_i = v$ that deviates to an empty node w ; denote the resulting assignment by \mathbf{v}' . Given an edge $e \in E$, let

$$\Delta(e) = \phi(\mathbf{v}', e) - \phi(\mathbf{v}, e).$$

Also, for $z \in \{v, w\}$, let $\Delta(z) = \sum_{e: z \in e} \Delta(e)$. Note that i 's move only changes the potential of edges incident to v and w . Hence, if v and w are not adjacent, we have $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = \Delta(v) + \Delta(w)$. We will now prove that $\Delta(v) + \Delta(w) > 0$; if v and w are not adjacent, this establishes our claim; towards the end of the proof we will explain how to handle the case $\{v, w\} \in E$. We make the following observations.

- As no agent benefits from moving to an isolated node, it must be $\deg(w) > 0$.
- If $\deg(w) = 1$, let $e_w \in E$ be the edge that is adjacent to w . Since w is empty in \mathbf{v} , we have that $\phi(\mathbf{v}, e_w) = \frac{1}{3}$. Since agent i benefits from moving to w , we have that $\phi(\mathbf{v}', e_w) = 1$. Hence, $\Delta(e_w) = \frac{2}{3}$ and, consequently, $\Delta(w) = \frac{2}{3}$.
- If $\deg(w) = 2$, let $e_{w,1}$ and $e_{w,2}$ be the two edges incident to w . Since w is empty in \mathbf{v} , we have that $\phi(\mathbf{v}, e_{w,1}) = \phi(\mathbf{v}, e_{w,2}) = \frac{1}{3}$. Since agent i benefits from moving to w , we have that $\phi(\mathbf{v}', e_{w,1}) + \phi(\mathbf{v}', e_{w,2}) \geq 1$. Hence, $\Delta(w) \geq \frac{1}{3}$.
- If $\deg(v) = 0$ then by definition $\Delta(v) = 0$.
- If $\deg(v) = 1$, let $e_v \in E$ be the edge that is incident to v . Since i benefits from moving away from v , we have that $\phi(\mathbf{v}, e_v) \leq \frac{1}{3}$. Since v is left empty in \mathbf{v}' , we have that $\phi(\mathbf{v}', e_v) = \frac{1}{3}$ and, consequently, $\Delta(v) \geq 0$.
- If $\deg(v) = 2$, let $e_{v,1}$ and $e_{v,2}$ be the two edges incident to v . Since v is left empty in \mathbf{v}' , we have that $\phi(\mathbf{v}', e_{v,1}) = \phi(\mathbf{v}', e_{v,2}) = \frac{1}{3}$. Since agent i benefits from moving away from v , we have that $\phi(\mathbf{v}, e_{v,1}) + \phi(\mathbf{v}, e_{v,2}) \leq 1$. Thus, $\Delta(v) \geq -\frac{1}{3}$.

By the above observations, it follows that $\Delta(v) + \Delta(w) > 0$ unless $\Delta(v) = -\frac{1}{3}$ and $\Delta(w) = \frac{1}{3}$. However, this is impossible: $\Delta(v) = -\frac{1}{3}$ only if in \mathbf{v} agent i is adjacent to one friend and one enemy, and $\Delta(w) = \frac{1}{3}$ only if in \mathbf{v}' agent i is adjacent to one friend and one enemy; but in such a case, agent i would have no incentive to move, a contradiction. This completes the analysis for when $\{v, w\} \notin E$.

Now, suppose that v and w are adjacent. In this case we have that $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = \Delta(v) + \Delta(w) - \Delta(\{v, w\})$. However, since w is empty in \mathbf{v} and v is empty in \mathbf{v}' , it must be $\Delta(\{v, w\}) = \frac{1}{3} - \frac{1}{3} = 0$, and hence $\Delta(v) + \Delta(w) > 0$ implies $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) > 0$ in this case as well.

Finally, note that the potential function takes values in the set $\{\frac{\ell}{3} \mid \ell = 0, \dots, 3|V|\}$, where $|V|$ is the number of nodes of the topology graph. Therefore, any best response dynamics starting from an arbitrary initial configuration converges to an equilibrium in $O(|V|)$ steps. \square

Conversely, all our non-existence results (Theorem 3.2), hardness results (Theorems 4.1 and 4.2) and lower bounds on the PoA and PoS (Section 5) apply to social Schelling games as well. In fact, maximizing the social welfare in social Schelling games is NP-hard even if all agents are strategic (whereas our hardness reduction for k -typed games uses stubborn agents). Moreover, this hardness result holds even if G is a graph of maximum degree 2, i.e., social welfare maximization may be hard even when finding equilibria is easy.

Theorem 6.2. *Given a social Schelling game I and a rational value s , it is NP-complete to decide whether I admits an assignment with social welfare at least s . The hardness result holds even if all agents are strategic and even if G is a graph of maximum degree 2.*

Proof. It is immediate that our problem is in NP. To show NP-hardness, we will use a reduction from the HAMILTONIAN CYCLE (HC) problem. An instance of HC is an undirected graph $H = (X, Y)$; it is a yes-instance if and only if the vertices of this graph can be ordered as $x_1, \dots, x_{|X|}$ so that $\{x_{|X|}, x_1\} \in E$ and for each $i \in [|X| - 1]$ it holds that $\{x_i, x_{i+1}\} \in E$.

Given an instance $H = (X, Y)$ of HC, where X is the set of nodes and Y is the set of edges, we construct an instance of our social welfare maximization problem as follows:

- For every node $v \in X$, we have a strategic agent p_v with set of friends $F(p_v) = \{p_z : \{z, v\} \in Y\}$.
- The topology $G = (V, E)$ is a cycle consisting of $|X|$ nodes together with an isolated node w .

By construction, a social welfare of $|X|$ can be achieved if and only if the agents can be assigned to the nodes of the cycle so that each of them is adjacent to two friends; this is possible if and only if H admits a Hamiltonian cycle. \square

Identifying special classes of social Schelling games that allow for good upper bounds on the price of anarchy and the price of stability is an interesting research direction. We note that the upper bounds in Section 5 only apply to k -typed instances with further restrictions on the structure of each type, so they cannot be extended to the social setting.

6.2 Schelling games with enemy aversion

In our model, if an agent is not adjacent to any friends, it does not matter how many enemies she is adjacent to. This is also the case in fractional hedonic games: agents are indifferent between being alone and being in coalitions consisting of their enemies. This assumption makes sense when the “enemies” of an agent are simply agents that do not contribute to her welfare. However, an agent may prefer being alone to being in a group full of enemies. In the context of hedonic games, such preferences are modeled by *modified fractional hedonic games* [Olsen, 2012; Elkind *et al.*, 2016; Bredereck *et al.*, 2019], where the utility of an agent in a coalition with f friends and e enemies is $\frac{f+1}{f+e+1}$, i.e., the agent herself is included in the set of her friends.

Many of our results extend to this definition of utility. For example, we can construct instances without equilibria even for 2-typed games, using ideas similar to those in the reduction of Theorem 4.1. Further, for k -typed games with a tree topology and a constant number of types, equilibrium existence

can be decided in polynomial time, by adapting the proof of Theorem 4.3. However, it remains an open question if instances with no stubborn agents always admit an equilibrium in this model.

6.3 Schelling games with linear utilities

Throughout the paper we assume that an agent's utility is determined by the fraction of her friends among her neighbors. Alternatively, an agent may simply care about the number of friends in her neighborhood or the difference between the number of friends f_i and the number of enemies e_i ; more broadly, her utility may be an arbitrary linear function of f_i and e_i (in the context of hedonic games, this model corresponds to a subclass of *additively separable hedonic games*; e.g., see [Aziz and Savani, 2016]). It turns out that games of this form are potential games and therefore have at least one equilibrium; furthermore, in the absence of stubborn agents there is always an equilibrium that is socially optimal.

Theorem 6.3. *Consider a variant of the (social) Schelling model where the utility of each agent i , who is adjacent to f_i friends and e_i enemies, is $\alpha f_i - \beta e_i$ for some $\alpha, \beta \geq 0$. Then, every instance has an equilibrium assignment which can be computed in polynomial time. Moreover, if no agent is stubborn, the price of stability is 1.*

Proof. Consider a game with a set of strategic agents R , a set of stubborn agents S , a topology $G = (V, E)$ and a friendship relation that is defined by a social network \mathcal{G} . Fix non-negative constants α and β such that the utility of an agent who is adjacent to f friends and e enemies in the topology is given by $\alpha f - \beta e$. Our analysis is inspired by Proposition 2 in the work of Bogomolnaia and Jackson [2002], showing that a Nash stable partition always exists in symmetric additively separable hedonic games.

Let $N = R \cup S$ and \mathbf{v} be an assignment. For each $i \in N$ let $\phi_i(\mathbf{v}) = \alpha f_i(\mathbf{v}) - \beta e_i(\mathbf{v})$, and $\Phi(\mathbf{v}) = \sum_{i \in N} \phi_i(\mathbf{v})$. We will argue that Φ is an ordinal potential function for our game. Note that if all agents are strategic, $\Phi(\mathbf{v})$ is equal to the social welfare of \mathbf{v} . However, in general this is not the case: intuitively, Φ ascribes “strategic” utilities to the stubborn agents.

Consider an assignment \mathbf{v} and an agent i with $v_i = v$. Suppose that i has a beneficial deviation from v to another node $w \in V$, which is empty in \mathbf{v} ; denote the resulting assignment by \mathbf{v}' . Suppose that agent i has f friends and e enemies at \mathbf{v} , and f' friends and e' enemies at \mathbf{v}' . Then, since the deviation is profitable, it holds that $\phi_i(\mathbf{v}') - \phi_i(\mathbf{v}) = \alpha(f' - f) - \beta(e' - e) > 0$. We claim that $\Phi(\mathbf{v}') > \Phi(\mathbf{v})$.

Indeed, consider an agent $j \in N \setminus \{i\}$. If j is a neighbor of i in both \mathbf{v} and \mathbf{v}' , or if j is not a neighbor of i in both \mathbf{v} and \mathbf{v}' , then $\phi_j(\mathbf{v}) = \phi_j(\mathbf{v}')$.

Now, suppose that j is adjacent to i in \mathbf{v} , but not in \mathbf{v}' . If j is a friend of i , then $\phi_j(\mathbf{v}') = \phi_j(\mathbf{v}) - \alpha$, and if j is an enemy of i , then $\phi_j(\mathbf{v}') = \phi_j(\mathbf{v}) + \beta$. Similarly, if j is adjacent to i in \mathbf{v}' , but not in \mathbf{v} , then if j is a friend of i , then $\phi_j(\mathbf{v}') = \phi_j(\mathbf{v}) + \alpha$, and if j is an enemy of i , then $\phi_j(\mathbf{v}') = \phi_j(\mathbf{v}) - \beta$. Thus, the overall change in potential can be computed as

$$\begin{aligned} \Phi(\mathbf{v}') - \Phi(\mathbf{v}) &= \phi_i(\mathbf{v}') - \phi_i(\mathbf{v}) - \alpha f + \beta e + \alpha f' - \beta e' \\ &= 2 \left(\alpha(f' - f) - \beta(e' - e) \right) > 0. \end{aligned}$$

It follows that, if the strategic agents follow the best response dynamics starting from any initial configuration, they will converge to an equilibrium. Moreover, the assignment that maximizes Φ is an equilibrium, so if all agents are strategic, this equilibrium maximizes the social welfare. Note also that the function Φ takes values in the set $\{\alpha i - \beta j \mid 0 \leq i, j \leq n^2\}$, where n is the number of agents. Thus, any best response dynamics converges in $O(n^4)$ iterations. \square

7 Conclusions

In this paper, we investigated Schelling games on graphs, both from the perspective of equilibrium analysis and from the perspective of social welfare. Concerning equilibrium existence, our positive results are rather limited in scope: while an equilibrium always exists for very simple topologies, such as stars and paths, it may fail to exist even if the topology does not contain cycles. It would be interesting to obtain a complete characterization of topologies that guarantee existence of equilibria.

For welfare maximization, a natural question is whether one can efficiently compute assignments with nearly optimal social welfare. We note that our NP-hardness reductions are not approximation preserving, so they do not rule out this possibility. Another interesting algorithmic question is whether the problem of computing equilibria in k -typed games remains hard in the absence of stubborn agents; we conjecture that this is indeed the case, but were unable to prove it.

References

- R. Alba and J. Logan. Minority proximity to whites in suburbs: An individual-level analysis of segregation. *American Journal of Sociology*, 98(6):1388–1427, 1993.
- E. Anshelevich, A. Dasgupta, J. M. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.
- H. Aziz and R. Savani. Hedonic games. In *Handbook of Computational Social Choice*, pages 356–376. 2016.
- H. Aziz, F. Brandt, and P. Harrenstein. Fractional hedonic games. In *Proceedings of the 2014 International conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 5–12, 2014.
- G. Barmpalias, R. Elwes, and A. Lewis-Pye. Digital morphogenesis via Schelling segregation. In *Proceedings of the 55th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 156–165, 2014.
- G. Barmpalias, R. Elwes, and A. Lewis-Pye. From randomness to order: unperturbed Schelling segregation in two or three dimensions. *CoRR*, abs/1504.03809, 2015.
- S. Benard and R. Willer. A wealth and status-based model of residential segregation. *Journal of Mathematical Sociology*, 31(2):149–174, 2007.
- I. Benenson, E. Hatna, and E. Or. From Schelling to spatially explicit modeling of urban ethnic and economic residential dynamics. *Sociological Methods and Research*, 37(4):463–497, 2009.
- A. Bogomolnaia and M. O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.
- C. Brandt, N. Immorlica, G. Kamath, and R. Kleinberg. An analysis of one-dimensional Schelling segregation. In *Proceedings of the 44th Symposium on Theory of Computing Conference (STOC)*, pages 789–804, 2012.
- R. Bredereck, E. Elkind, and A. Igarashi. Hedonic diversity games. In *Proceedings of the 2019 International conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2019.
- A. Chauhan, P. Lenzner, and L. Molitor. Schelling segregation with strategic agents. In *Proceedings of the 11th International Symposium on Algorithmic Game Theory (SAGT)*, pages 137–149, 2018.

- W. Clark and M. Fossett. Understanding the social context of the Schelling segregation model. *Proceedings of the National Academy of Sciences*, 105(11):4109–4114, 2008.
- J. H. Drèze and J. Greenberg. Hedonic coalitions: optimality and stability. *Econometrica*, 48(4):987–1003, 1980.
- D. A. Easley and J. M. Kleinberg. *Networks, Crowds, and Markets – Reasoning about a Highly Connected World*. Cambridge University Press, 2010.
- E. Elkind, A. Fanelli, and M. Flammini. Price of Pareto optimality in hedonic games. In *Proceedings of the 20th AAAI Conference on Artificial Intelligence (AAAI)*, pages 475–481, 2016.
- N. Immorlica, R. Kleinberg, B. Lucier, and M. Zadomighaddam. Exponential segregation in a two-dimensional Schelling model with tolerant individuals. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 984–993, 2017.
- E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 404–413, 1999.
- M. Olsen. On defining and computing communities. In *Proceedings of the 18th Computing: Australasian Theory Symposium (CATS)*, pages 97–102, 2012.
- R. Pans and N. Vriend. Schelling’s spatial proximity model of segregation revisited. *Journal of Public Economics*, 91(1–2):1–24, 2007.
- T. C. Schelling. Models of segregation. *American Economic Review*, 59(2):488–493, 1969.
- T. C. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2):143–186, 1971.
- H. P. Young. *Individual Strategy and Social Structure: an Evolutionary Theory of Institutions*. Princeton University Press, 2001.
- J. Zhang. A dynamic model of residential segregation. *Journal of Mathematical Sociology*, 28(3):147–170, 2004.
- J. Zhang. Residential segregation in an all-integrationist world. *Journal of Economic Behavior and Organization*, 54(4):533–550, 2004.

A Proof of Theorem 4.3

For readability, we will present a polynomial-time algorithm that can decide whether an equilibrium exists for instances with two agent types (red and blue) and no stubborn agents; towards the end of the proof, we will explain how to extend it to instances with a constant number of agent types that may contain stubborn agents. Let T_R denote the set of all red agents and let T_B denote the set of all blue agents. Throughout the proof, we use the convention that a fraction of the form $\frac{a}{b}$ evaluates to 0 whenever $a = 0$.

Consider an instance I with n agents and tree topology $G = (V, E)$. Pick an arbitrary node r to be the root of G . Let $tree(v)$ denote the set of descendants of v (including v), and let $child(v)$ be the set of children of v . Observe that the utility of a strategic agent takes values in the set $\mathcal{U} = \{i/j : i \in [n], j \in [n], i \leq j\} \cup \{0\}$; note that $|\mathcal{U}| \leq n^2$.

We use the following dynamic programming approach. For each node $v \in V$, we fill out a table τ_v , which contains an entry $\tau_v(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$ for each tuple $(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$, where

- $C \in \{blue, red, empty\}$,
- $\mathbf{n} = (n_B, n_R) \in [n]^2$,
- $\mathbf{k} = (k_B, k_R) \in [n]^2$,
- $\check{\mathbf{u}} = (\check{u}_B, \check{u}_R, \check{u}_{B^\dagger}, \check{u}_{R^\dagger}) \in \mathcal{U}^4$, and
- $\hat{\mathbf{u}} = (\hat{u}_B, \hat{u}_R, \hat{u}_{top}) \in \mathcal{U}^3$.

Thus, the number of entries in each table is $3 \cdot n^4 \cdot |\mathcal{U}|^7$, which is polynomial in the input size.

The value of each entry is either *true* or *false*. Specifically, $\tau_v(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \text{true}$ if and only if there exists an assignment of a subset of agents to the nodes in $tree(v)$ that satisfies the following conditions:

1. If $C = \text{empty}$, then node v is empty and otherwise it is assigned to an agent of color C .
2. Exactly n_B nodes of $tree(v)$ are assigned to blue agents, and exactly n_R nodes of $tree(v)$ are assigned to red agents.
3. Exactly k_B nodes of $child(v)$ are assigned to blue agents, and exactly k_R nodes of $child(v)$ are assigned to red agents.
4. Every blue agent in a node of $child(v)$ gets utility at least \check{u}_{B^\dagger} and every red agent in a node of $child(v)$ gets utility at least \check{u}_{R^\dagger} .
5. Every blue agent in a node of $tree(v) \setminus (child(v) \cup \{v\})$ gets utility at least \check{u}_B and every red agent in a node of $tree(v) \setminus (child(v) \cup \{v\})$ gets utility at least \check{u}_R .
6. If a blue agent that is *not* already in $tree(v)$ moves to an empty node of $tree(v) \setminus \{v\}$, her utility would be at most \hat{u}_B , and if a red agent that is *not* already in $tree(v)$ moves to an empty node of $tree(v) \setminus \{v\}$, her utility would be at most \hat{u}_R .
7. If node v is not empty, then the agent occupying v can get utility at most \hat{u}_{top} by moving to an empty node of $tree(v) \setminus \{v\}$.
8. All agents in nodes of $tree(v) \setminus \{v\}$ do not have an incentive to deviate to empty nodes of $tree(v) \setminus \{v\}$.

Condition 8 directly relates to stability of $tree(v) \setminus \{v\}$, whereas conditions 1–7 are auxiliary, providing the necessary information that we need in order to determine the stability of node v , and fill out the dynamic programming table for the parent of v .

Consider the table τ_r at the root node r . The game admits an equilibrium if and only if there exists $(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$ such that $n_B = |T_B|$, $n_R = |T_R|$, $\tau_r(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \text{true}$ for the root node r of G , and, moreover,

- if $C = \text{blue}$, then

$$\frac{k_B}{k_B + k_R} \geq \hat{u}_{top};$$

- if $C = \text{red}$, then

$$\frac{k_R}{k_B + k_R} \geq \hat{u}_{top};$$

- if $C = \text{empty}$, then for each $X \in \{R, B\}$ with $k_X > 0$ it holds that

$$\frac{k_X}{k_B + k_R} \leq \check{u}_X, \quad \frac{k_X - 1}{k_B + k_R - 1} \leq \check{u}_{X^\dagger}.$$

The first two conditions ensure that if the root node is not empty, the agent in that node does not have an incentive to move to another node of the tree, and the last condition ensures that if the root node is empty, no agent has an incentive to deviate there (the exact form of this condition depends on whether the potential deviator is located in a child of r). Together with condition 8, these conditions ensure that no agent wants to deviate.

The existence of a tuple $(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$ with these properties can be decided in polynomial time by going through all entries of τ_r . It remains to show that τ_r can be filled in in polynomial time.

Given $C \in \{\text{red}, \text{blue}, \text{empty}\}$, we write $\mathbb{1}_B(C) = 1$ if $C = \text{blue}$ and 0 otherwise; similarly, $\mathbb{1}_R(C) = 1$ if $C = \text{red}$ and 0 otherwise, and $\mathbb{1}_E(E) = 1$ if $C = \text{empty}$ and 0 otherwise.

We fill the tables in all nodes starting from the leaf nodes of G . For every leaf node v , we have

$$T_v(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \begin{cases} \text{true}, & \text{if } \mathbf{n} = (\mathbb{1}_B(C), \mathbb{1}_R(C)), \mathbf{k} = (0, 0) \text{ and } \hat{\mathbf{u}} = (0, 0, 0) \\ \text{false}, & \text{otherwise.} \end{cases} \quad (1)$$

Suppose now that for a node w we have constructed the table τ_v for each $v \in \text{child}(w)$. We will construct τ_w using these tables as follows. Let $\text{child}(w) = \{v_1, \dots, v_L\}$. We create an intermediate table θ_w^ℓ for each $\ell \in \{0, 1, \dots, L\}$. This table has an entry $\theta_w^\ell(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$ for every tuple $(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$. The entry $\theta_w^\ell(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$ is set to *true* if and only if conditions 1–8 hold for the subtree $tree_\ell(w)$ obtained from $tree(w)$ by deleting the subtrees rooted at $v_{\ell+1}, \dots, v_L$. Note that, by construction, we have $\tau_w(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \theta_w^L(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}})$.

We construct θ_w^ℓ sequentially for $\ell = 0, \dots, L$. We can fill out θ_w^0 using Equation (1). Next, suppose that we have filled out the first ℓ tables, i.e., $\theta_w^0, \dots, \theta_w^{\ell-1}$. We combine $\theta_w^{\ell-1}$ and τ_{v_ℓ} in order to build θ_w^ℓ as follows: $\theta_w^\ell(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \text{true}$ if and only if there exist a pair of tuples $(C', \mathbf{n}', \mathbf{k}', \check{\mathbf{u}}', \hat{\mathbf{u}}')$ and $(C'', \mathbf{n}'', \mathbf{k}'', \check{\mathbf{u}}'', \hat{\mathbf{u}}'')$ such that $\theta_w^{\ell-1}(C', \mathbf{n}', \mathbf{k}', \check{\mathbf{u}}', \hat{\mathbf{u}}') = \tau_{v_\ell}(C'', \mathbf{n}'', \mathbf{k}'', \check{\mathbf{u}}'', \hat{\mathbf{u}}'') = \text{true}$ and the following conditions hold:

1. $C' = C$.
2. $\mathbf{n}'' + \mathbf{n}' = \mathbf{n}$.
3. $\mathbb{1}_B(C'') + k'_B = k_B$ and $\mathbb{1}_R(C'') + k'_R = k_R$.

4. For each $X \in \{B, R\}$,

$$\check{u}'_{X^\dagger} \geq \check{u}_{X^\dagger}$$

so that the agents occupying nodes $v_1, \dots, v_{\ell-1}$ have utility at least \check{u}_{X^\dagger} . Additionally, if $C'' = \text{blue}$, then

$$\frac{k''_B + \mathbb{1}_B(C'')}{k''_B + k''_R + (1 - \mathbb{1}_E(C''))} \geq \check{u}_{B^\dagger}$$

and, if $C'' = \text{red}$, then

$$\frac{k''_R + \mathbb{1}_R(C'')}{k''_B + k''_R + (1 - \mathbb{1}_E(C''))} \geq \check{u}_{R^\dagger}$$

so that the agent occupying node v_ℓ has utility at least \check{u}_{B^\dagger} if she is blue or at least \check{u}_{R^\dagger} if she is red. Therefore, if these conditions hold, all agents occupying the first ℓ children of w have utility at least \check{u}_{B^\dagger} or \check{u}_{R^\dagger} , according to their type.

5. For each $X \in \{B, R\}$,

$$\check{u}'_X, \check{u}''_X, \check{u}''_{X^\dagger} \geq \check{u}_X,$$

so that all agents of type X occupying the nodes of $\text{tree}_\ell(w) \setminus (\text{child}(w) \cup \{w\})$ have utility at least \check{u}_X .

6. For each $X \in \{B, R\}$,

$$\hat{u}'_X \leq \hat{u}_X, \hat{u}''_X \leq \hat{u}_X$$

and, if $C'' = \text{empty}$, then

$$\frac{k''_X + \mathbb{1}_X(C'')}{k''_B + k''_R + (1 - \mathbb{1}_E(C''))} \leq \hat{u}_X,$$

so that the agents that do not occupy nodes of $\text{tree}_\ell(w)$ have no incentive to deviate to any node in the first $\ell - 1$ branches, any node other than v_ℓ in the ℓ -th branch, or node v_ℓ .

7. If $C' = \text{blue}$, then

$$\hat{u}'_{\text{top}} \leq \hat{u}_{\text{top}}, \quad \hat{u}''_B \leq \hat{u}_{\text{top}},$$

and, if $C'' = \text{empty}$, then

$$\frac{k''_B}{k''_B + k''_R} \leq \hat{u}_{\text{top}},$$

so that the blue agent i^* occupying node w has utility at most \hat{u}_{top} if she deviates to a node in the first $\ell - 1$ branches, a node in the ℓ -th branch (excluding node v_ℓ), or node v_ℓ . Similarly, if $C' = \text{red}$, then

$$\hat{u}'_{\text{top}} \leq \hat{u}_{\text{top}}, \quad \hat{u}''_R \leq \hat{u}_{\text{top}}$$

and, if $C'' = \text{empty}$,

$$\frac{k''_R}{k''_B + k''_R} \leq \hat{u}_{\text{top}}.$$

8. $\check{u}''_B, \check{u}''_{B^\dagger} \geq \hat{u}'_B$ so that blue agents occupying nodes in the ℓ -th branch (excluding v_ℓ) have no incentive to deviate to any node in the first $\ell - 1$ branches. Also, if $C'' = \text{empty}$, then

$$\check{u}''_B \geq \frac{k''_B + \mathbb{1}_B(C'')}{k''_B + k''_R + \mathbb{1}_B(C'')},$$

and if $k''_B > 0$ then

$$\check{u}''_{B^\dagger} \geq \frac{k''_B + \mathbb{1}_B(C'') - 1}{k''_B + k''_R + \mathbb{1}_B(C'') - 1},$$

so that blue agents occupying nodes other than v_ℓ in the ℓ -th branch have no incentive to deviate to v_ℓ . Since $\tau_{v_\ell}(C'', \mathbf{n}'', \mathbf{k}'', \check{\mathbf{u}}'', \hat{\mathbf{u}}'') = \text{true}$ means that these agents already have no incentive to deviate to other empty nodes in the ℓ -th branch, now these agents have no incentive to deviate to any empty node in $\text{tree}_\ell(w) \setminus \{w\}$. Further, if $C'' = \text{blue}$, then

$$\frac{k_B'' + \mathbb{1}_B(C'')}{k_B'' + k_R'' + \mathbb{1}_B(C'')} \geq \hat{u}_B', \hat{u}_{top}''$$

so that if there is a blue agent at node v_ℓ , she has no incentive to deviate as well. Similar constraints must hold for red agents.

9. $\check{u}_B', \check{u}_{B^\dagger}' \geq \hat{u}_B''$ so that blue agents occupying nodes in the first $\ell - 1$ branches have no incentive to deviate to nodes in the ℓ -th branch (excluding node v_ℓ). Additionally, if $C'' = \text{empty}$, then

$$\check{u}_B', \check{u}_{B^\dagger}' \geq \frac{k_B'' + \mathbb{1}_B(C'')}{k_B'' + k_R''},$$

so that blue agents in the first $\ell - 1$ branches have no incentive to deviate to v_ℓ if it is empty. Similar constraints must hold for red agents as well.

These constraints can be verified in polynomial time by checking each pair of entries of the tables $\theta_w^{\ell-1}$ and τ_{v_ℓ} . This completes the proof for instances with two agent types and no stubborn agents.

To extend the algorithm to instances with stubborn agents, we can set the entry values of the table τ_v to *false* if v is occupied by a stubborn agent of a type other than C , and only consider possible deviations by strategic agents. The algorithm can trivially be extended to instances with constant number of different agent types; the size of the tables would scale exponentially with the number of types. \square