Schelling Games on Graphs*

Aishwarya Agarwal[1], Edith Elkind[1], Jiarui Gan[2], Ayumi Igarashi[3],
Warut Suksompong[4] and Alexandros A. Voudouris[5]

[1]Department of Computer Science, University of Oxford, UK
[2]Max Planck Institute for Software Systems, Germany
[3]Principles of Informatics Research Division, National Institute of Informatics, Japan
[4]School of Computing, National University of Singapore, Singapore
[5]School of Computer Science and Electronic Engineering, University of Essex, UK

Abstract

We study strategic games inspired by Schelling's seminal model of residential segregation. These games are played on undirected graphs, with the set of agents partitioned into multiple types; each agent either aims to maximize the fraction of her neighbors who are of her own type, or occupies a node of the graph and never moves away. We consider two natural variants of this model: in *jump games* agents can *jump* to empty nodes of the graph to increase their utility, while in *swap games* they can *swap* positions with other agents. We investigate the existence, computational complexity, and quality of equilibrium assignments in these games, both from a social welfare perspective and from a diversity perspective. Some of our results extend to a more general setting where the preferences of the agents over their neighbors are defined by a social network rather than a partition into types.

Keywords: Schelling games; Equilibrium analysis; Price of anarchy; Computational complexity

1 Introduction

In 2019, African Americans constituted 78% of the population of the City of Detroit. At the same time, the neighboring Oakland County was 75% white, and in the city of Hamtramck in the Detroit metropolitan area about 26% of the residents were Asians (compared to 1.7% in the City of Detroit and 8.2% in Oakland County).[1] Similar phenomena of residential segregation have been extensively documented in numerous 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.

To formalize and study how the motives of individuals lead to residential segregation, Schelling [1969, 1971] proposed the following elegant model. There are two types of agents who are to be placed on a line or a grid. An agent is happy with her location if at least a fraction $\tau \in (0, 1]$ of the agents within a

*This paper unifies and expands earlier versions that appeared in Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019) [Elkind et al., 2019] and Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI 2020) [Agarwal et al., 2020]. In particular, it includes new results (5.1, 5.10, 5.11, 6.1), improved bounds (5.3, 5.5), as well as extensions of previous results (7.1, 7.3).

[1]See http://en.wikipedia.org/wiki/Ethnic_groups_in_Metro_Detroit#/media/File:FischerDetroit2010Census.png for a map of ethnic groups in the Detroit metropolitan area.

certain radius are of the same type as her. Happy agents do not want to move, but unhappy agents are willing to do so in the hope of improving their current situation. Schelling described a dynamic process in which at each step, unhappy agents jump to random unoccupied locations or swap positions with other randomly selected agents. He showed via simple experiments that, surprisingly, this process can lead to a strongly segregated placement even when the agents themselves are tolerant of mixed neighborhoods (i.e., $\tau < 1/2$).

In the half-century since Schelling's pioneering papers, this segregation model has attracted the attention of many researchers in sociology and economics [Alba and Logan, 1993, Benard and Willer, 2007, Pancs and Vriend, 2007, Clark and Fossett, 2008, Benenson et al., 2009], who proposed and studied several variants of Schelling's model, mainly via agent-based simulations.[2] Variants of the model have also been theoretically analyzed in a series of papers, including in computer science [Young, 2001, Zhang, 2004a,b, Brandt et al., 2012, Barmpalias et al., 2014, Bhakta et al., 2014, Barmpalias et al., 2015, Immorlica et al., 2017]; one of the main findings is that, with high probability, the random behavior of the agents leads to the formation of large monochromatic regions, thereby implying that strong segregation is likely to occur.

While most of the earlier work has focused on settings where the behavior of the agents is random, in practice it is more realistic to assume that the agents are *strategic* and move only when they have an opportunity to improve their situation. Prior to our work, such a game-theoretic approach has only been followed in a handful of papers. Specifically, Zhang [2004b] considered a game where the agents optimize a single-peaked utility function. Chauhan et al. [2018] studied a strategic setting with two types of agents who have preferred locations and can either swap with other agents or jump to empty positions. For a given tolerance threshold $\tau \in (0, 1]$, each agent's primary goal is to maximize the fraction of her neighbors that are of her own type as long as this fraction is below τ (with all fractions above τ being equally good); her secondary goal is to be as close as possible to her preferred location. For both types of games (swap and jump), Chauhan et al. identified values of τ for which the best response dynamics of the agents leads to an equilibrium when the topology is a ring or a regular graph. More recently, after the publication of an initial version of our work [Elkind et al., 2019], there has been a stream of follow-up papers on Schelling games; we discuss them in Section 1.2.

1.1 Our Contribution

The model considered by Chauhan et al. [2018] makes an important contribution to the literature by enriching Schelling's model with two additional components: (1) agents who are fully strategic and (2) location preferences. However, the resulting model of the agents' preferences is quite complex and, consequently, not easy to analyze: the positive results of Chauhan et al. are mostly limited to special cases of the utility function and highly regular networks. In our work, we focus on a similar model, aiming to capture the same phenomena, but in a way that is more amenable to formal analysis. In particular, our paper is the first to consider more than two types of agents—an important feature of real-life segregation scenarios such as that of the Detroit metropolitan area described earlier.

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 into our model, but 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 either *jumping* to an unoccupied location or *swapping* locations with other agents;

[2] See http://ncase.me/polygons for an easily accessible implementation.

we refer to the two classes of games as k*-jump* and k*-swap* games, respectively. Our model aims to capture the fact that, in practice, many residents (such as older people or those with underwater mortgages) are unwilling to move even if they are no longer satisfied with the composition of their neighborhood. Unlike most of the work on Schelling's model, we do not assume that agents have tolerance thresholds. Instead, a strategic agent is willing to move as long as there exists another location with a better happiness ratio (i.e., $\tau = 1$). We answer a wide range of fundamental questions concerning our model, which we formally define in Section 2. We remark here that some of our negative results (i.e., non-existence, computational intractability, and price of anarchy/stability lower bounds) can be extended for some other values of $\tau < 1$; we discuss this in Section 8.

In Section 3, we show that for some classes of topologies, such as stars and graphs of maximum degree 2, our games always admit an *equilibrium assignment*, that is, the strategic agents can be assigned to the nodes of the topology so that none of them want to move to a different location; this result holds even for the more general "social Schelling games" (see the overview of Section 7 below), and such an assignment can be computed efficiently. In contrast, equilibria may fail to exist even for games in which the topology is acyclic and has a small maximum degree (4 for jump games and 3 for swap games).[3] We complement this negative 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 (so in particular for the case $k = 2$, which is the focus of the prior work). For more general topologies, we prove that deciding whether an equilibrium exists is an NP-complete problem.

In the next three sections, we study the efficiency of (equilibrium) assignments. Following a well-established research agenda in the algorithmic game theory literature, we primarily focus on the *social welfare* objective, which is defined as the sum of agents' utilities. In addition, given that the goal of Schelling's work was to study integration and segregation, it is also natural to ask what level of integration can be achieved at equilibrium. While a number of integration indices have been proposed (see the survey of Massey and Denton [1988]), several of them are specifically defined for settings where the topology is highly regular and there are only two agent types, so it is not clear how to adapt them to our general model. We therefore focus on a simple index that we call the *degree of integration*; it is inspired by the work of Lieberson and Carter [1982] and admits a natural interpretation in our context. This index counts the number of agents who are *exposed* to agents of other types, i.e., have at least one neighbor of a different type.

In Section 4, we show that finding assignments (not necessarily equilibria) with a high social welfare or a high degree of integration is computationally hard. Then, in Section 5, we thoroughly study the effects of 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]. That is, for any given game, we compare the minimum/maximum social welfare over all equilibrium assignments to the maximum social welfare over all (not necessarily equilibrium) assignments. Even though the price of anarchy is unbounded in general, we identify particularly interesting restricted subclasses of jump and swap games in which the price of anarchy (and thus the price of stability) is bounded by the number of agents or the number of types; see Table 1 for an overview of our price of anarchy bounds. In Section 6, we turn our attention to bounding the price of anarchy and the price of stability with respect to the degree of integration—to the best of our knowledge, this is the first such analysis for an integration index in the context of Schelling games. We illustrate that even the best equilibria can be much less diverse than the maximally diverse assignments.

Finally, in Section 7, we discuss several variants of our basic typed model. We show that some of our

[3] This corrects a result of Chauhan et al. [2018, Theorem 8], which claims that equilibria are guaranteed to exist for both 2-jump and 2-swap games regardless of the topology when agents have no preferred location (as in our model).

		fully-strategic		with stubborn agents
		general	balanced	strongly-balanced
Jump games		$\Theta(n)$	$\in [2k-1, 2k]$	$\left[\frac{3k-1}{3k+1} \cdot 2k, 2k\right]$
Swap games	$k = 2$	$\Theta(n)$	$\in (2.058, 4]$	$\in (2.666, 4]$
	$k \geq 3$	unbounded		unbounded

Table 1: An overview of our price of anarchy bounds from Section 5 for games with at least two strategic agents per type and a connected topology. In cases where there is a type with a single strategic agent or the topology is disconnected, the price of anarchy is unbounded. We have two main axes: the strategy space of the agents (*jump* or *swap*), and whether there are stubborn agents (we refer to games without stubborn agents as *fully-strategic*). When all agents are strategic, we further distinguish between *general* games (with possibly unequal number of agents per type) and *balanced* games (where each type consists of the same number of agents). When there are stubborn agents, we present results for *strongly-balanced* games (where each type consists of the same number of stubborn agents, and the same number of strategic agents), as in any other case the price of anarchy is unbounded.

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 (that is, their neighbors in the social network) among their neighbors in the topology—we refer to the resulting class of games as *social Schelling games*. We also present two alternative utility functions that differ conceptually from the canonical one. The first function aims to capture enemy aversion by including each agent in the set of her own friends, while the second aims to capture scenarios in which the agents only care about the *difference* between the number of her friends and the number of her enemies in her neighborhood (as opposed to the ratio). For the difference-based utility function, we show that there is always an equilibrium and, when all agents are strategic, the price of stability is 1.

1.2 Further Related Work

For an accessible introduction to the Schelling model and a survey of the literature on non-strategic variants, we refer the interested reader to 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 paper of Chauhan et al. [2018], which we discussed in detail earlier, a number of other authors have studied similar models after the publication of the conference version of our paper [Elkind et al., 2019]. In particular, Echzell et al. [2019] strengthened the results of Chauhan et al. [2018] and extended them to more than two agent types, and also examined the complexity of computing assignments that maximize the number of happy agents. Kanellopoulos et al. [2021a] investigated a variant of the utility function with enemy aversion that we discuss in Section 7.2. Massand and Simon [2019] considered a class of linear swap games, like the ones we study in Section 7.3, where the utility of the agents is a linear function of the weights they have for their neighbors. A similar setting with agents that derive linear utility both from their location as well as their nearby friends, was recently studied by Elkind et al. [2020b]. Bilò et al. [2020] examined the influence of the topology and locality on the existence of equilibria and the price of anarchy in swap games. Chan et al. [2020] introduced an alternative model wherein multiple agents can occupy the same location and, similarly to our social Schelling games (Section 7.1), there is a friendship network. Bullinger et al. [2021] presented results on the complexity of computing assignments fulfilling welfare guarantees or other efficiency notions. Very recently, Kreisel et al. [2021] built on our work to establish that finding equilibria is hard even if all agents are strategic; their result holds both for jump and for swap games.

Our model shares a number of similarities with hedonic games [Drèze and Greenberg, 1980, Bogomolnaia and Jackson, 2002]; these are games in which 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., 2019], the relationships among the agents are described by a weighted directed graph: the weight of an edge (i, j) represents the value that agent i assigns to agent j, and an agent's utility for a coalition is her average value for the 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. In this case, an agent's utility in a coalition corresponds to the fraction of her friends among the coalition members, which is similar to how utilities are defined in social Schelling games. More precisely, the utility definition in our games is analogous to that in unweighted *modified fractional hedonic games* [Olsen, 2012, Bredereck et al., 2019, Monaco et al., 2019, Elkind et al., 2020a], where the denominator of the fraction that describes the utility of agent i in a coalition X is $|X \setminus \{i\}|$ rather than $|X|$, i.e., the average is computed over coalition members other than i. Moreover, our type-based model is closely related to the Bakers and Millers game discussed by Aziz et al. [2019], and to the recently introduced class of hedonic diversity games [Bredereck et al., 2019, Boehmer and Elkind, 2020]. These connections between Schelling games and hedonic games motivate much of the discussion in Section 7. Nevertheless, a fundamental difference between hedonic games and Schelling games is that in hedonic games, agents derive their utilities from pairwise disjoint coalitions, whereas in Schelling games the utilities are derived from (overlapping) neighborhoods.

2 Preliminaries

A *Schelling game* is given by a set $N = \{1, \dots, n\}$ of $n \geq 2$ *agents* partitioned into $k \geq 2$ pairwise disjoint *types* $T_1, \dots, T_k$, and an undirected simple graph $G = (V, E)$, called the *topology*. We often identify agent types with colors; for example, in Schelling games with two types, the agents are either red (T_1) or blue (T_2). Given an agent $i \in T_\ell$, we refer to all other agents in T_ℓ as *friends of* i and denote the set of i's friends by $F_i = T_\ell \setminus \{i\}$.

An *assignment* is a vector $\mathbf{v} = (v_1, \dots, v_n)$, where $v_i \in V$ for each $i \in N$ and $v_i \neq v_j$ for every pair of agents $i \neq j$; intuitively, v_i is the location of agent i and no two agents can occupy the same location. Let $\pi_v(\mathbf{v})$ denote the agent that occupies the node $v \in V$ according to an assignment $\mathbf{v}$, that is, $\pi_{v_i}(\mathbf{v}) = i$. Given an assignment $\mathbf{v}$, let $N_i(\mathbf{v}) = \{j \in N \setminus \{i\} : \{v_i, v_j\} \in E\}$ be the set of neighbors of agent i.

In addition to their types, the agents are also classified as either *strategic* or *stubborn*. We denote by R the set of strategic agents and by S the set of stubborn agents, so that $R \cup S = N$. In case $S = \varnothing$, the setting is *fully-strategic*. The utility of a strategic agent $i \in R$ for an assignment $\mathbf{v}$ is the fraction of her neighbors who are her friends, i.e.,

$$u_i(\mathbf{v}) = \frac{|N_i(\mathbf{v}) \cap F_i|}{|N_i(\mathbf{v})|}.$$

and 0, in case $N_i(\mathbf{v}) = \varnothing$. Strategic agents aim to maximize their utility, and may move in order to do so. In contrast, each stubborn agent is associated with a node of the topology and never moves away from that node (i.e., the node that each stubborn agent occupies is explicitly given and never changes); we assume that their utility is independent of the assignment and set $u_i(\mathbf{v}) = 0$ for each $i \in S$.

To maximize her utility, a strategic agent may either *jump* to an empty node of the topology or *swap* positions with another agent. Games with k types in which the strategic agents are only allowed to jump to empty nodes are called k*-jump games*, while games in which the strategic agents are only allowed to perform pairwise swaps are called k*-swap games*; in this paper, we do not consider games where both types of moves are allowed. In jump games, we assume that $|V| > n$, so that there always exists at least one empty node to which an agent can jump; in swap games, we assume that $|V| = n$, since empty nodes will remain unoccupied throughout the game and have no impact on the utility calculations.

Observe that in swap games, since $|V| = n$, every node is occupied by some agent, so if G is connected then $N_i(\mathbf{v}) \neq \varnothing$ for every $i \in N$. However, this is not generally true in jump games and it may be that $N_i(\mathbf{v}) = \varnothing$ for some agent i and assignment $\mathbf{v}$, even if G is connected. In such cases, we assume that the utility of agent i is 0.

In a jump game, let $(z, \mathbf{v}_{-i})$ be the assignment that is obtained from $\mathbf{v}$ by changing the location of a strategic agent i from v_i to z, where z is unoccupied in $\mathbf{v}$. We say that an assignment $\mathbf{v}$ is a *pure Nash jump equilibrium* (or, simply, *a jump equilibrium*) if no strategic agent i has an incentive to unilaterally deviate to an empty node in order to increase her utility. That is, $\mathbf{v}$ is a jump equilibrium if 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

$$u_i(\mathbf{v}) \geq u_i(z, \mathbf{v}_{-i}).$$

In a swap game, let $\mathbf{v}^{i\leftrightarrow j}$ be the assignment that is obtained from $\mathbf{v}$ by swapping the positions of the strategic agents i and j: $v_\ell^{i\leftrightarrow j} = v_\ell$ for every $\ell \in N \setminus \{i, j\}$, $v_i^{i\leftrightarrow j} = v_j$ and $v_j^{i\leftrightarrow j} = v_i$. Agents i and j swap positions only if they both strictly increase their utility: $u_i(\mathbf{v}^{i\leftrightarrow j}) > u_i(\mathbf{v})$ and $u_j(\mathbf{v}^{i\leftrightarrow j}) > u_j(\mathbf{v})$. Clearly, agents of the same type cannot both increase their utilities by swapping, so swaps always involve agents of different types. An assignment $\mathbf{v}$ is a *swap equilibrium* if no pair of agents i, j want to swap positions. That is, $\mathbf{v}$ is a swap equilibrium if and only if for every $i, j \in R$ we have

$$u_i(\mathbf{v}) \geq u_i(\mathbf{v}^{i\leftrightarrow j}) \text{ or } u_j(\mathbf{v}) \geq u_j(\mathbf{v}^{i\leftrightarrow j}).$$

In what follows, we omit the words 'jump'/'swap' when the context is clear, and denote the set of all (jump or swap) equilibrium assignments of the (jump or swap) game $\mathcal{G}$ by $\text{EQ}(\mathcal{G})$.

For every assignment $\mathbf{v}$, we define two benchmarks that aim to capture, respectively, the happiness of the agents and the societal diversity. The first one is the well-known *social welfare*, defined as the total utility of all strategic agents:

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

Our second benchmark is the *degree of integration*. We say that a strategic agent is *exposed* if she has at least one neighbor of a different type, and count the number of exposed agents:

$$\text{DI}(\mathbf{v}) = |\{i \in R : N_i(\mathbf{v}) \setminus F_i \neq \varnothing\}|.$$

Note that we ignore the stubborn agents in the definitions of our benchmarks, as their utility is always the same and they never want to move somewhere else.

For $f \in \{\text{SW}, \text{DI}\}$, let $\mathbf{v}_f^*(\mathcal{G})$ be an *optimal* assignment in terms of the benchmark f for a given game $\mathcal{G}$. The *price of anarchy* (PoA) in terms of the benchmark f is the worst-case ratio (over all games $\mathcal{G}$ such that $\text{EQ}(\mathcal{G}) \neq \varnothing$) between the optimal performance (among all assignments) and the performance of an equilibrium assignment that *minimizes* this benchmark. Similarly, the *price of stability* (PoS) in terms of f is the worst-case ratio between the optimal performance and the performance of an equilibrium assignment that *maximizes* this benchmark:

$$\text{PoA}_f = \sup_{\mathcal{G}:\text{EQ}(\mathcal{G}) \neq \varnothing} \max_{\mathbf{v} \in \text{EQ}(\mathcal{G})} \frac{f(\mathbf{v}_f^*(\mathcal{G}))}{f(\mathbf{v})},$$

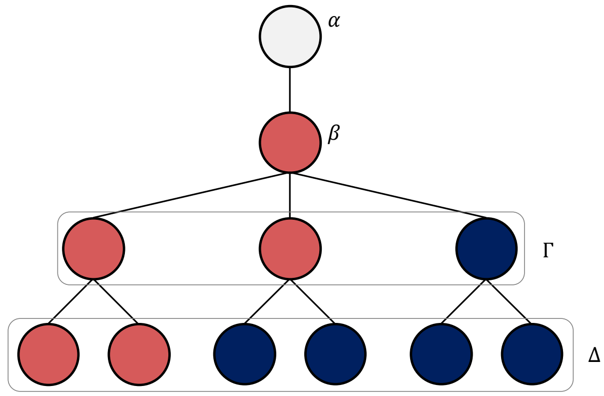

Figure 1: Example of the topology used in the proof of Theorem 3.2 for $k = 2$. The assignment depicted here in which node α is empty is not an equilibrium, since the middle red agent occupying a node of Γ wants to deviate to α and thereby increase her utility from $1/3$ to 1.

$$\mathrm{PoS}_f = \sup_{\mathcal{G}:\mathrm{EQ}(\mathcal{G})\neq\varnothing} \min_{\mathbf{v}\in\mathrm{EQ}(\mathcal{G})} \frac{f(\mathbf{v}_f^*(\mathcal{G}))}{f(\mathbf{v})}.$$

For readability, we refer to the quantity $\mathrm{PoA}_{\mathrm{SW}}$ as the *social price of anarchy* and to $\mathrm{PoA}_{\mathrm{DI}}$ as the *integration price of anarchy*, and use similar terminology for the price of stability.

3 Existence and Complexity of Computing Equilibria

We begin the presentation of our technical results by discussing the existence of equilibria and studying the complexity of computing equilibrium assignments. We warm up by observing that for topologies such as paths, rings, and stars, there is always at least one equilibrium assignment, and such an assignment can be computed efficiently. This can be shown directly, and also follows from a more general result established later (Theorem 7.1).

Theorem 3.1. *For every $k \geq 2$, every k-jump or k-swap game with a topology that is a star or a graph of maximum degree 2 admits at least one equilibrium assignment, which can be computed in polynomial time.*

In what follows, we consider jump and swap games separately in Sections 3.1 and 3.2, respectively. While the results we present are similar for both classes of games, in many cases the proofs exploit different techniques that are specifically designed for each type of games.

3.1 Jump Games

Let us first focus on jump games. We show that for this class of games, an equilibrium may fail to exist in general. This negative result 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-jump game that does not admit an equilibrium assignment, even when all agents are strategic and the topology is a tree.*

Proof. Given $k \geq 2$, we construct an instance with $2k+1$ agents per type; thus, 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 leaves of the tree; we denote the set of all leaves (excluding the root) by Δ. Figure 1 depicts the topology for $k = 2$.

Assume for the sake of contradiction that there is an equilibrium assignment. Since exactly one node is left empty, we consider four cases depending on its location. We show that each case cannot result in an equilibrium, which yields the desired contradiction.

- 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 only $2k-1$ nodes in Γ, there must exist some subtree rooted at a node in Γ that contains an agent of type T as well as 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 of type T is hungry. Indeed, if β is occupied by an agent j of type T, then j 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. □

Our next result shows that deciding whether a jump game admits an equilibrium assignment is an intractable problem in general.

Theorem 3.3. *For every $k \geq 2$, it is* NP*-complete to decide whether a given k-jump game admits an equilibrium assignment, even if all strategic agents belong to the same type.*

Proof. We give a proof for $k = 2$; one can extend it to $k \geq 2$ by adding isolated stubborn agents of different types. One can verify whether a given assignment is an equilibrium simply by checking all possible deviations, so our problem is in NP. To prove that it is NP-hard, we provide a reduction from the Clique problem. An instance of this problem is an undirected graph $H = (X, Y)$ and an integer ξ; it is a yes-instance if and only if H has a complete subgraph of size ξ. This problem remains NP-hard if we require that $\xi \geq 5$ [Garey and Johnson, 1979]. Given an instance $\langle H, \xi \rangle$ of Clique with $H = (X, Y)$ and $\xi \geq 5$, we construct an instance of our problem as follows:

- There are two agent types: red and blue.

- There are ξ 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 $V_1 = X \cup W$, $|W| = \xi - 2$, $E_1 = Y \cup \{\{v, w\} : v \in X, w \in W\}$. There is a stubborn blue agent at each node $w \in W$;
  - G_2 is a complete bipartite graph with parts L and R, $|L| = \xi - 2$, $|R| = 4\xi$. Of the 4ξ nodes in R, $2\xi + 1$ nodes are occupied by stubborn red agents and $2\xi - 1$ nodes are occupied by stubborn blue agents;
  - G_3 has three empty nodes, denoted x, y, and z, and 136 nodes occupied by stubborn agents—47 red and 89 blue. 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.

  Figure 2 depicts the topology G defined above. We can connect G_1, G_2 and G_3 by adding edges between nodes occupied by stubborn agents, so that the resulting topology is connected; this has no impact on the strategic agents' behavior. This shows that our hardness result holds even if we require the topology to be connected.

Note that a strategic red agent obtains a utility of $\frac{2\xi+1}{4\xi} = \frac{1}{2} + \frac{1}{4\xi}$ 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 H contains a clique of size ξ. If strategic red agents occupy the nodes of the corresponding clique in G_1, the utility of each such agent is

$$\frac{\xi - 1}{(\xi - 1) + (\xi - 2)} = \frac{1}{2} + \frac{1}{4\xi - 6} > \frac{1}{2} + \frac{1}{4\xi}.$$

Thus, by our choice of parameters, no agent has a profitable deviation.

On the other hand, suppose that H does not contain a clique of size ξ. 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 $\xi - 1$ friends, as this would mean that their locations form a clique of size ξ. Hence, at least one of these agents is adjacent to at most $\xi - 2$ friends. As this agent is also adjacent to the $\xi - 2$ 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, $\frac{1}{2} + \frac{1}{4\xi}$. Thus, if there are strategic agents in G_1, all $\xi - 2$ 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 $\frac{1}{\xi-1} < \frac{1}{3}$ (recall that we assume that $\xi \geq 5$). 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 node 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 H does not have a clique of size ξ, then there is no equilibrium assignment. $\square$

We note that the hardness result of Theorem 3.3 is established for jump games with stubborn agents, and our hardness reduction makes heavy use of stubborn agents. Very recently, Kreisel et al. [2021] showed

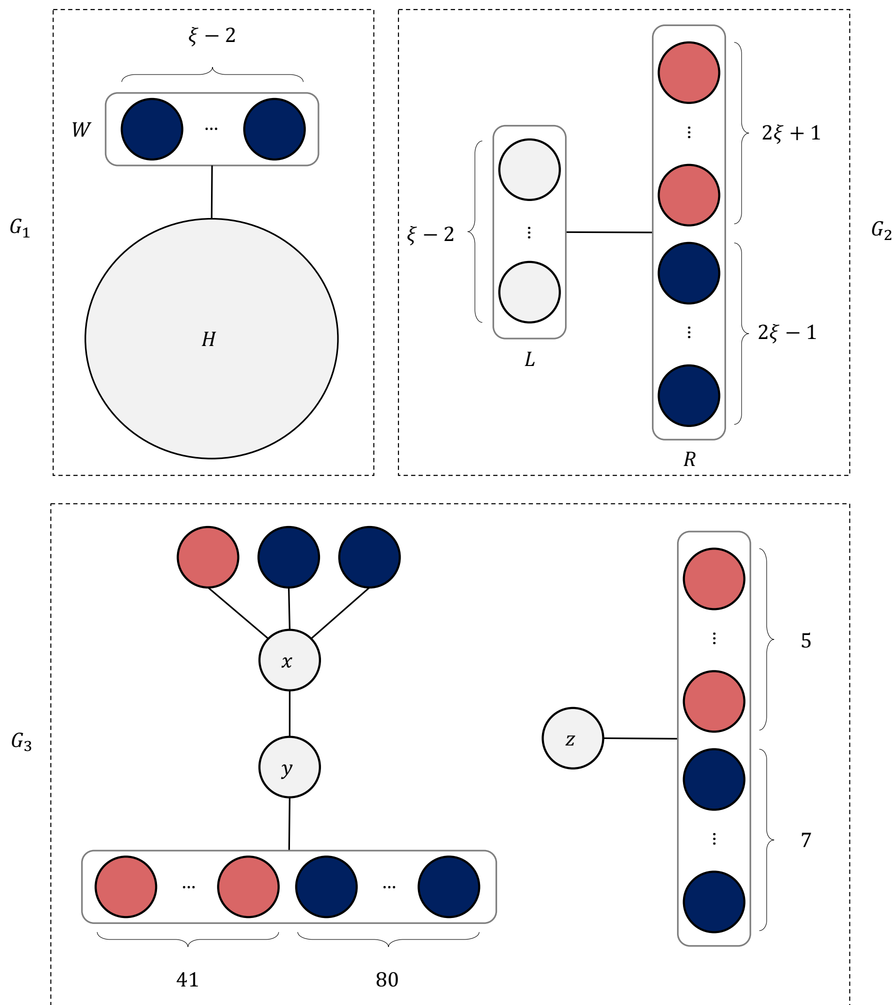

Figure 2: The topology G used in the reduction of Theorem 3.3, which consists of three subgraphs G_1, G_2 and G_3. An edge between two components indicates that each node in one component is connected to every node of the other component. Colored nodes are occupied by stubborn agents of the corresponding type; all other nodes are empty.

that deciding the existence of equilibria is hard even if all agents are strategic, thereby answering an open question from the conference version of our paper. Their construction relies on ours, but is significantly more involved to avoid using stubborn agents.

On the positive side, for small k we can efficiently decide whether an equilibrium exists if the topology G is a tree.[4] 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 then decides whether the given instance admits an equilibrium by scanning the table at the root node. The details of the algorithm are rather involved, so we relegate the algorithm and its analysis to the appendix.

[4]The algorithm used in the proof Theorem 3.4 is presented for k-jump games, but can also be adapted for k-swap games by not enumerating empty nodes; we omit the details, and only present the algorithm for the harder case of jump games.

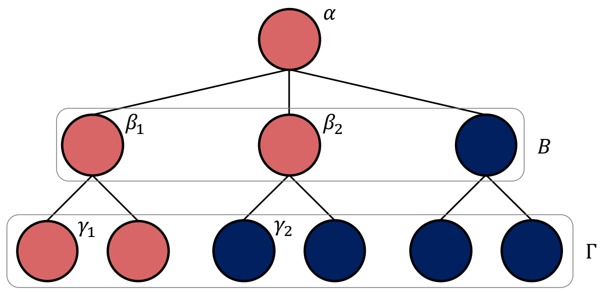

Figure 3: The topology of the 2-swap game considered in the proof of Theorem 3.5, and an assignment that corresponds to the last case in the analysis; it is not an equilibrium since the red agent at node α and the blue agent at node γ_2 would like to swap.

Theorem 3.4. *Given a k-jump game with a tree topology, we can decide whether there exists an equilibrium (and compute one if it exists) in time $poly(n^k)$, i.e., this problem lies in the complexity class* XP *with respect to the number of types k.*

3.2 Swap Games

We now consider swap games, and again start with a proof of non-existence of equilibria for every $k \geq 2$.

Theorem 3.5. *For every $k \geq 2$, there exists a k-swap game that does not admit an equilibrium assignment, even when all agents are strategic and the topology is a tree.*

Proof. We start with the case of $k = 2$. Consider a 2-swap game with 10 strategic agents: 5 red agents and 5 blue agents. The topology is a tree with a root node α, which has three children nodes (set B), each of which has two children of its own (set Γ); see Figure 3. Suppose for the sake of contradiction that this game admits an equilibrium assignment $\mathbf{v}$.

Since $|B| = 3$ and there are only two types of agents, at least two nodes in B, say β_1 and β_2, must be occupied by agents of the same type, say red. Now assume that nodes γ_1 (a child of β_1) and γ_2 (a child of β_2) are occupied by blue agents. Then the red agent $\pi_{\beta_1}(\mathbf{v})$ and the blue agent $\pi_{\gamma_2}(\mathbf{v})$ can swap positions to increase their utility from strictly smaller than 1 and 0 to 1 and positive, respectively. Therefore, for at least one of these nodes (say, β_1) it must be the case that both of its children are occupied by red agents; as there are only five red agents, it follows that at least one of the children of β_2, say γ_2, is occupied by a blue agent.

If node α is occupied by a blue agent, then the red agent $\pi_{\beta_1}(\mathbf{v})$ and the blue agent $\pi_{\gamma_2}(\mathbf{v})$ can both increase their utility by swapping. Hence, node α must be occupied by a red agent (see Figure 3). However, this assignment is not an equilibrium either, since the red agent $\pi_{\alpha}(\mathbf{v})$ and the blue agent $\pi_{\gamma_2}(\mathbf{v})$ have an incentive to swap.

For $k \geq 3$, consider a k-swap game with $n = k(k^2 - 2)$ agents such that there are $k^2 - 2$ agents of each type. The topology is a tree whose nodes are distributed over three layers, just like in the case $k = 2$. Specifically, there is a root node α, which has a set B of $k(k-1) - 1$ children. Each node $\beta \in B$ has a set Γ_β of k children leaf nodes; let $\Gamma = \bigcup_{\beta \in B} \Gamma_\beta$. Next, we will identify some configurations that cannot arise in an equilibrium.

Lemma 3.6. *Consider an assignment $\mathbf{v}$ such that there are two nodes $\beta_1, \beta_2 \in B$ that are occupied by agents of the same type T_x, and there exist nodes $\gamma_1 \in \Gamma_{\beta_1}$ and $\gamma_2 \in \Gamma_{\beta_2}$ that are occupied by agents of some type T_y, $y \neq x$. Then $\mathbf{v}$ is not an equilibrium.*

Proof. Let $\mathbf{v}$ be an assignment that satisfies the conditions in the statement of the lemma. Then the utility of agent $\pi_{\beta_1}(\mathbf{v})$ is strictly less than 1, while the utility of agent $\pi_{\gamma_2}(\mathbf{v})$ is 0. Therefore, they would like to swap positions in order to increase their utility to 1 and positive, respectively. $\square$

Lemma 3.7. *Consider an assignment $\mathbf{v}$ such that for every $\ell \in [k]$ there exists an agent of type T_ℓ that occupies some node of B. Then $\mathbf{v}$ is not an equilibrium.*

Proof. Let $\mathbf{v}$ be an assignment that satisfies the condition in the statement of the lemma. Without loss of generality, assume that the agent $\pi_\alpha(\mathbf{v})$ is of type T_x. We now distinguish between the following two cases.

- There exist nodes $\beta \in B$ and $\gamma \in \Gamma_\beta$ such that $\pi_\beta(\mathbf{v})$ is of type T_x and $\pi_\gamma(\mathbf{v})$ is of type T_y, $y \neq x$. By the assumption of the lemma, there exists at least one agent of type T_y located at some node $\beta' \in B \setminus \{\beta\}$. Therefore, agents $\pi_\alpha(\mathbf{v})$ and $\pi_\gamma(\mathbf{v})$ would like to swap positions in order to increase their utility from strictly less than 1 and 0 to 1 and positive, respectively.

- For every node $\beta \in B$ occupied by an agent of type T_x, all agents occupying the nodes of Γ_β are of type T_x. Since α is occupied by an agent of type T_x, there are $k^2 - 3 = (k-1)(k+1) - 2$ other agents of type T_x. Since each subtree rooted at a node of B has $k+1$ nodes, there are at most $k-2$ such subtrees that can be completely filled up by agents of type T_x. Consequently, there are at least $k^2 - 3 - (k-2)(k+1) = k-1$ agents of type T_x located at leaf nodes whose (unique) neighbor is not of type T_x.

  Now, assume that one of these agents of type T_x occupies a node $\gamma \in \Gamma_\beta$ such that $\beta \in B$ is occupied by an agent of type T_y, with $y \neq x$. We will now argue that there must exist another node $\beta' \in B \setminus \{\beta\}$ occupied by an agent of type T_y. Indeed, assume that there is no such node. Then all agents of type T_y are located in leaf nodes of the tree. There are at least $k^2 - 2 - k$ agents of type T_y that are not located in the subtree rooted in β. As $k^2 - k - 2 > k(k-2)$ for $k > 2$, they appear in at least $k-1$ different subtrees; let $B' \subset B$ be the set of roots of these subtrees. By Lemma 3.6, if two agents of type T_y appear in different subtrees, the roots of these subtrees have to be occupied by agents of different types. Thus, the nodes of B' are occupied by agents of $k-1$ different types; moreover, by our assumption, none of them is occupied by an agent of type T_x. Thus, at least one node in B' is occupied by an agent of type T_y, a contradiction. Consequently, there exists a node $\beta' \in B \setminus \{\beta\}$ occupied by an agent of type T_y. But then agents $\pi_\gamma(\mathbf{v})$ and $\pi_{\beta'}(\mathbf{v})$ can swap positions in order to increase their utility from 0 and strictly less than 1 to positive and 1, respectively.

This completes the proof of the lemma. $\square$

Lemma 3.8. *Consider an assignment $\mathbf{v}$ such that for some $\ell \in [k]$ there is no node in B that is occupied by an agent of type T_ℓ. Then $\mathbf{v}$ is not an equilibrium.*

Proof. Let $\mathbf{v}$ be an assignment that satisfies the condition in the statement of the lemma, and assume for the sake of contradiction that $\mathbf{v}$ is an equilibrium.

Suppose first that $k \geq 4$. As there is at most one agent of type T_ℓ occupying α, there are at least $k^2 - 3$ agents of type T_ℓ that must occupy nodes of Γ. Further, $k^2 - 3 > k(k-1)$ for $k > 3$, so these agents appear in at least k subtrees rooted in a node of B. By Lemma 3.6, if there exist two distinct nodes $\beta, \beta' \in B$ such that a child of β and a child of β' are occupied by agents of type T_ℓ, then β and β' cannot be occupied by agents of the same type. Hence, B must contain nodes occupied by agents of k different types, a contradiction with the assumption that no node in B is occupied by an agent of type T_ℓ.

Now, consider the case $k = 3$. Assume without loss of generality that $\ell = 3$. If α is not occupied by an agent of type T_3, then all $k^2 - 2 = 7$ agents of this type must occupy nodes of Γ, and the same argument as above leads to a contradiction. Hence, assume that $\pi_\alpha(\mathbf{v})$ is of type T_3, so there are 6 agents of type T_3 occupying nodes of Γ. By Lemma 3.6, there exist two nodes $\beta, \beta' \in B$ such that β is occupied by an agent of type T_1, β' is occupied by an agent of type T_2, and all nodes in $\Gamma_\beta \cup \Gamma_{\beta'}$ are occupied by agents of type T_3. Now, consider a node $\beta'' \in B \setminus \{\beta, \beta'\}$, and assume without loss of generality that it is occupied by an agent of type T_1. Then an agent in $\gamma \in \Gamma_\beta$ and the agent in β'' can swap positions to increase their utilities from 0 and strictly less than one to positive and 1, respectively. □

By Lemmas 3.7 and 3.8, we conclude that no assignment can be an equilibrium. □

The topology used in the proof of Theorem 3.5 for $k = 2$ is utilized as a subgraph in the proof of the following theorem, which shows that the problem of deciding whether an equilibrium exists is computationally hard.

Theorem 3.9. *For every $k \geq 2$, it is* NP*-complete to decide whether a given k-swap game admits an equilibrium assignment.*

Proof. Membership in NP is immediate: we can verify whether a given assignment is an equilibrium by simply checking if there exists a pair of agents that would like to swap positions. To prove NP-hardness, as in Theorem 3.3, we give a reduction from the Clique problem. Specifically, we will show how to map an instance $\langle H, \xi \rangle$ of Clique with $H = (X, Y)$ and $\xi > 5$ to a k-swap game so that H has a clique of size ξ if and only if our game admits a swap equilibrium.

Given an instance $\langle H, \xi \rangle$ of Clique with $H = (X, Y)$ and $\xi > 5$, we will construct a 2-swap game as follows (the reduction can be extended to any $k > 2$ by adding isolated agents of different types). Let d_v denote the degree of node v in H, and set $d_H = \max_{v \in X} d_v$.

- There are ξ strategic red agents and $t = |X| + 5$ strategic blue agents; all other agents are stubborn, and will be defined in conjunction with the topology.
- The topology $G = (V, E)$ consists of three components G_1, G_2 and G_3. These are connected to each other via stubborn agents, and their internal structure is defined as follows:
  - To define $G_1 = (V_1, E_1)$, let W_v be a set of $2d_H - d_v + 2\xi - 3$ nodes for each $v \in X$. Then, $V_1 = X \cup \left(\bigcup_{v \in X} W_v\right)$ and $E_1 = Y \cup \{\{v, w\} : v \in X, w \in W_v\}$. For every $v \in X$, d_H nodes of W_v are occupied by stubborn red agents, while the remaining $d_H - d_v + 2\xi - 3$ nodes are occupied by stubborn blue agents. Observe that every empty node of G_1 has degree $\delta = 2d_H + 2\xi - 3$.
  - The subgraph $G_2 = (A \cup B, E_2)$ is a complete bipartite graph with $|A| = \xi - 5$ and $|B| = 4\delta$. Out of the 4δ nodes of B, $2\delta + 1$ nodes are occupied by stubborn red agents, while the remaining $2\delta - 1$ nodes are occupied by stubborn blue agents.

    Hence, a strategic red agent occupying a node of A has utility $\chi_r = \frac{2\delta+1}{4\delta} = \frac{1}{2} + \frac{1}{4\delta}$. Similarly, a strategic blue agent has utility $\chi_b = \frac{2\delta-1}{4\delta} = \frac{1}{2} - \frac{1}{4\delta}$.
  - To define $G_3 = (V_3, E_3)$, let $G_3' = (V_3', E_3')$ be the graph used in the proof of Theorem 3.5, for which there is no equilibrium assignment; see Figure 3. For every *non-leaf* node $v \in V_3'$, let Z_v

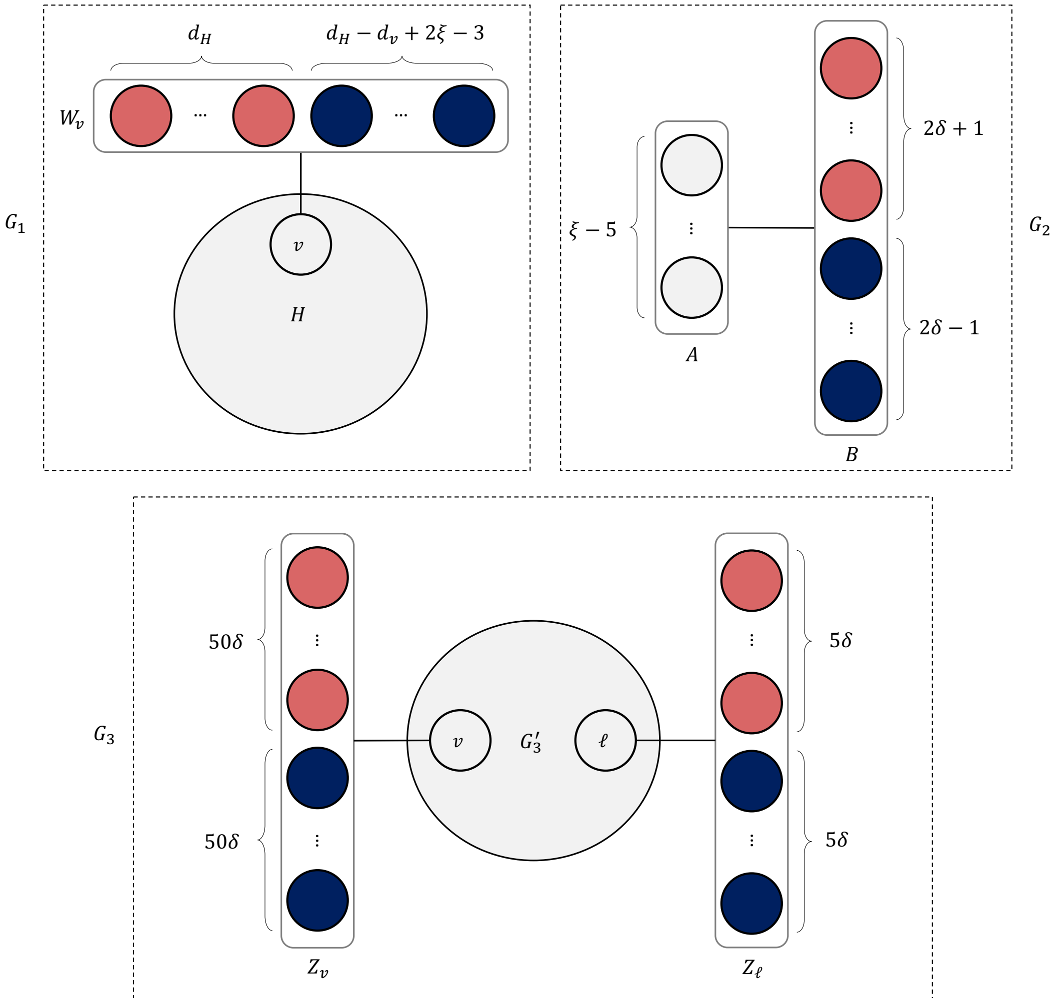

Figure 4: The topology G used in the reduction of Theorem 3.9. An edge between two components indicates that each node in one component is connected to every node of the other component. Colored nodes are occupied by stubborn agents of the corresponding type. In G_3, v is an internal node, whereas ℓ is a leaf node.

be a set of 100δ nodes such that 50δ of these nodes are occupied by stubborn red agents, while the remaining 50δ nodes are occupied by stubborn blue agents. For every *leaf* node $v \in V_3'$, let Z_v be a set of 10δ nodes such that 5δ of these nodes are occupied by stubborn red agents, while the remaining 5δ nodes are occupied by stubborn blue agents. Then, $V_3 = V_3' \cup \left(\bigcup_{v \in V_3'} Z_v\right)$ and $E_3 = E_3' \cup \{\{v, w\} : v \in V_3', w \in Z_v\}$.

One can verify that the utility of a strategic agent (red or blue) occupying a node of G_3 is at least $\psi_0 = \frac{5\delta}{10\delta+1} > \frac{1}{2} - \frac{1}{4\delta}$ and at most $\psi_1 = \frac{5\delta+1}{10\delta+1} < \frac{1}{2} + \frac{1}{4\delta}$.

Figure 4 depicts the topology G defined above.

Now, assume that H has a clique of size ξ, and let $\mathbf{v}$ be an assignment in which the strategic red agents occupy the nodes of the clique, and the strategic blue agents occupy the remaining available nodes. Each strategic red agent is connected to $\xi - 1 + d_H$ other red agents (strategic and stubborn) in G_1, and thus has utility

$$u = \frac{\xi - 1 + d_H}{\delta} = \frac{d_H + \xi - 1.5 + 0.5}{2d_H + 2\xi - 3} = \frac{1}{2} + \frac{1}{2\delta}.$$

Clearly, since $u > \chi_r$ and $u > \psi_1$, no strategic red agent would be willing to swap positions with another strategic agent in G_2 or G_3. Moreover, if a red agent were to swap positions with a blue agent within G_1, she would still be adjacent to at most $\xi - 1 + d_H$ red agents, and since every node in G_1 has the same degree, her utility cannot be improved. Hence, no strategic red agent can benefit from a swap, so $\mathbf{v}$ is an equilibrium.

Conversely, assume that H does not contain a clique of size ξ, and for the sake of contradiction also assume that there is an equilibrium assignment $\mathbf{v}$.

Suppose that some strategic red agents are located in G_1. It cannot be the case that each of them is adjacent to $\xi - 1$ other strategic red agents, as this would mean that the nodes they occupy form a clique of size ξ. Hence, at least one of these agents, say agent i, is adjacent to at most $\xi - 2$ strategic red agents. Since every node of G_1 has degree δ and every node is adjacent to d_H stubborn red agents, the utility of i is

$$u_i \leq \frac{d_H + \xi - 2}{\delta} = \frac{d_H + \xi - 1.5 - 0.5}{2d_H + 2\xi - 3} = \frac{1}{2} - \frac{1}{2\delta}.$$

We have $u_i < \chi_r$ and $u_i < \psi_0$, and hence agent i has an incentive to move to G_2 or G_3. On the other hand, the utility that a strategic blue agent j that is currently located in G_2 or G_3 (there is always such an agent) can obtain by swapping positions with i is

$$u_j = 1 - u_i \geq \frac{1}{2} + \frac{1}{2\delta}.$$

Since $u_j > \chi_b$ and $u_j > \psi_1$, agent j also has an incentive to swap positions with agent i, and hence $\mathbf{v}$ cannot be an equilibrium assignment. Therefore, no strategic red agent is located in G_1.

Similarly, observe that $\chi_r > \psi_1$ and $\chi_b < \psi_0$, meaning that strategic red agents would prefer to be in G_2, while strategic blue agents would prefer to be in G_3. Thus, for $\mathbf{v}$ to be an equilibrium assignment, it must be the case that if a node of G_2 is not occupied by a stubborn agent, it is occupied by a strategic red agent. As a result, there are 5 strategic red agents and 5 strategic blue agents in G_3. However, similarly to the proof of Theorem 3.5, we can argue that there is no equilibrium assignment for these agents in G_3. Since we have exhausted all possibilities, it follows that if H does not have a clique of size ξ, then there is no equilibrium assignment. □

Just as for jump games, Theorem 3.9 has recently been strengthened by Kreisel et al. [2021], who showed that the hardness result holds even for fully-strategic swap games. In addition, when the topology is a tree, we can decide the existence of an equilibrium using a dynamic programming algorithm similar to the one used for jump games (see also Footnote 4).

4 Maximizing Social Welfare and Degree of Integration

In this section, we focus on the problems of computing assignments with high social welfare or high degree of integration. Observe that the complexity of these problems does not depend on the set of strategic actions available to the agents, i.e., it does not matter whether we consider jump games or swap games.

We start by showing that, for $k \geq 3$, maximizing the social welfare is hard. Our hardness result holds when the number of locations $|V|$ is equal to the number of agents n, which corresponds to the framework of swap games. We can extend it to the case $|V| > n$ (so as to capture the framework of jump games) by introducing isolated nodes.

Theorem 4.1. *For every $k \geq 3$, given a Schelling game with k types and a rational value ξ, it is* NP*-complete to decide whether the game admits an assignment with social welfare at least ξ. The hardness result holds even if all but two types consist of stubborn agents only.*

Proof. Membership in NP is immediate: given an assignment, we can sum up the utilities of the strategic agents and check whether the social welfare is at least ξ. To prove NP-hardness, we give a reduction from an NP-complete variant of the min-cut problem with additional cardinality constraints on the size of the subsets; this problem is known as the Equal-Min-Cut problem [Garey et al., 1974]. An instance of Equal-Min-Cut consists of a graph $H = (X, Y)$, two distinguished nodes $s, t \in X$, and an integer β. It is a yes-instance if and only if there exist disjoint subsets of nodes X_1 and X_2 such that $X_1 \cup X_2 = X$, $|X_1| = |X_2|$, $s \in X_1$, $t \in X_2$ and $|\{\{v, z\} \in Y : v \in X_1, z \in X_2\}| \leq \beta$. To simplify notation, we write vz to denote an edge $\{v, z\}$. Without loss of generality, we assume that $|X|$ is an even number. Assume also that $k = 3$ (to extend our hardness proof to $k > 3$, we can create stubborn agents of additional types placed in isolated nodes), and the three types are red, blue, and green.

Given an instance $\langle H, s, t, \beta\rangle$ of Equal-Min-Cut with $H = (X, Y)$, we construct an instance of our social welfare maximization problem as follows:

- There are $|X|/2 - 1$ strategic red and $|X|/2 - 1$ strategic blue agents.
- The topology $G = (V, E)$ consists of H with additional nodes and edges:
  - Let s^* and t^* be two auxiliary nodes, and $X^* = X \cup \{s^*, t^*\}$, and $Y^* = Y \cup \{s^*v : sv \in Y\} \cup \{t^*v : tv \in Y\}$.
  - For every $v \in X \setminus \{s, t\}$, let $d_v = |\{e \in Y^* : v \in e\}|$ be the degree of v on Y^*, and $d^* = \max_{v \in X\setminus\{s,t\}} d_v$. Let Z_v be a set of $d^* - d_v$ nodes.
  - Let $G = (V, E)$, where $V = X^* \cup \left(\bigcup_{v\in X\setminus\{s,t\}} Z_v\right)$ and $E = Y^* \cup \{vz : v \in X\setminus\{s, t\}, z \in Z_v\}$. Hence, the sets Z_v ensure that every node in $X \setminus \{s, t\}$ has the same degree d^* on G.
- Let nodes s and s^* be occupied by stubborn red agents, nodes t and t^* be occupied by stubborn blue agents, and all nodes in $\bigcup_{v\in X\setminus\{s,t\}} Z_v$ be occupied by stubborn green agents.
- Finally, let $\xi = \frac{2}{d^*}(|Y| - \beta)$.

We will argue that this instance admits an assignment $\mathbf{v}$ with $\mathrm{SW}(\mathbf{v}) \geq \xi$ if and only if $\langle H, s, t, \beta\rangle$ is a yes-instance of Equal-Min-Cut.

We first show that for any assignment $\mathbf{v}$, the social welfare is

$$\mathrm{SW}(\mathbf{v}) = \frac{2}{d^*}|\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}|. \tag{$\star$}$$

For ease of description, we let $\chi_v(\mathbf{v})$ be the type of the agent occupying node v. Since strategic agents occupy nodes in $X \setminus \{s, t\}$, we have

$$\mathrm{SW}(\mathbf{v}) = \sum_{v\in X\setminus\{s,t\}} \frac{|\{vz \in E : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}|}{d^*}$$

$$
\begin{aligned}
&= \sum_{v \in X \setminus \{s,t\}} \frac{|\{vz \in Y^* : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}|}{d^*} \\
&= \frac{1}{d^*} \sum_{v \in X \setminus \{s,t\}} |\{vz \in Y^* : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), z \in X \setminus \{s,t\}\}| \\
&\quad + \frac{1}{d^*} \sum_{v \in X \setminus \{s,t\}} |\{vz \in Y^* : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), z \in \{s,t,s^*,t^*\}\}|, \qquad (1)
\end{aligned}
$$

where the second transition holds since every edge not in Y^* is occupied by a green agent on one end.

Consider an arbitrary node $v \in X \setminus \{s,t\}$. Edge vs is in Y if and only if both vs and vs^* are in Y^*. The same can be said for the edges vt and vt^*. The way the stubborn agents are positioned gives $\chi_s(\mathbf{v}) = \chi_{s^*}(\mathbf{v})$ and $\chi_t(\mathbf{v}) = \chi_{t^*}(\mathbf{v})$. It follows that

$$
|\{vz \in Y^* : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), z \in \{s,t,s^*,t^*\}\}| = 2|\{vz \in Y : \chi_v(\mathbf{v}) \ = \chi_z(\mathbf{v}), z \in \{s,t\}\}|. \qquad (2)
$$

Moreover, if $z \in X \setminus \{s,t\}$, then $vz \in Y$ if and only if $vz \in Y^*$. Hence,

$$
|\{vz \in Y^* : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), z \in X \setminus \{s,t\}\}| = |\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), z \in X \setminus \{s,t\}\}|. \qquad (3)
$$

Substituting (3) and (2) into (1) gives

$$
\begin{aligned}
\mathrm{SW}(\mathbf{v}) &= \frac{2}{d^*}|\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), v, z \in X \setminus \{s,t\}\}| \\
&\quad + \frac{2}{d^*}|\{vz \in Y : \chi_v(\mathbf{v}) \ = \chi_z(\mathbf{v}), v \in X \setminus \{s,t\}, z \in \{s,t\}\}| \\
&= \frac{2}{d^*}|\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), v \in X \setminus \{s,t\}\}| \\
&= \frac{2}{d^*}|\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}|,
\end{aligned}
$$

where the last transition holds because we have $\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v}), v, z \in \{s,t\}\} = \varnothing$, given that $\chi_s(\mathbf{v}) \neq \chi_t(\mathbf{v})$. Note that in the first transition there is a factor 2 for the first term on the right side. This is because each vz is counted twice in the summation in (1). Therefore, ($\star$) holds.

Now, assume that the input instance $\langle H, s, t, \beta \rangle$ is a yes-instance of Equal-Min-Cut, and let $X = X_1 \cup X_2$ be a partition that witnesses this; hence, $|\{vz \in Y : v \in X_1, z \in X_2\}| \leq \beta$. We construct an assignment $\mathbf{v}$ by placing the strategic red agents into nodes of $X_1 \setminus \{s\}$ and the strategic blue agents into nodes of $X_2 \setminus \{t\}$. We have

$$
|\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}| = |Y| - |\{vz \in Y : v \in X_1, z \in X_2\}| \geq |Y| - \beta.
$$

It follows by ($\star$) that $\mathrm{SW}(\mathbf{v}) \geq \frac{2}{d^*}(|Y| - \beta) = \xi$.

Conversely, assume that there exists an assignment $\mathbf{v}$ with $\mathrm{SW}(\mathbf{v}) \geq \xi = \frac{2}{d^*}(|Y| - \beta)$. Let X_1 consist of s and the nodes occupied by strategic red agents, and let X_2 consist of t and the nodes occupied by strategic blue agents. Then, $X_1 \cap X_2 = \varnothing$, and since there is an equal number of strategic red and blue agents, we also have $|X_1| = |X_2|$. By ($\star$), we have $|Y| - |\{vz \in Y : \chi_v(\mathbf{v}) = \chi_z(\mathbf{v})\}| \leq \beta$, or equivalently,

$$
|\{vz \in Y : v \in X_1, z \in X_2\}| \leq \beta,
$$

which means that $\langle H, s, t, \beta \rangle$ is a yes-instance. □

Next, we establish the hardness of maximizing social welfare for the case $k = 2$ and $|V| > n$; very recently, Bullinger et al. [2021] complemented our result by showing that the problem is hard, even when $k = 2$ and $|V| = n$.

Theorem 4.2. *Given a Schelling game with two types and a rational value ξ, it is* NP*-complete to decide whether the game admits an assignment with social welfare at least ξ. The hardness result holds even if 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 3.3 by removing the gadgets G_2 and G_3 and replacing the set W with a single node w. That is, given an instance $\langle H, \xi' \rangle$ of Clique, we construct an instance of our social welfare maximization problem as follows:

- There are two agent types: red and blue.
- There are ξ' 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.
- $\xi = \xi' - 1$.

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 $r \leq \xi' - 1 = \xi$ for any assignment. Hence, the social welfare of ξ 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 $\frac{\xi'-1}{\xi'}$; this is possible if and only if H contains a clique of size ξ'. □

We now focus on the complexity of computing assignments with a high degree of integration, and show that this task is also computationally intractable for $k = 2$. The hardness holds even if all agents are strategic and the topology $G = (V, E)$ is such that $|V| = n$.

Theorem 4.3. *For every $k \geq 2$, given a Schelling game with k types, it is* NP*-complete to decide whether the game admits an assignment in which every agent is exposed. The hardness result holds even if all agents are strategic, and even if the number of locations is equal to the number of agents.*

Proof. The problem is clearly in NP: for a given assignment we can verify whether each of the n agents has at least one neighbor of a different type in time $O(n^2)$. For the NP-hardness proof, we give a reduction from the Vertex Cover problem, which is known to be NP-complete [Garey and Johnson, 1979]. An instance of Vertex Cover consists of an undirected graph $H = (X, Y)$ and an integer $\lambda \leq |X|$. It is a yes-instance if and only if there exists a set $X' \subseteq X$ such that $|X'| \leq \lambda$ and $\{v, w\} \cap X' \neq \varnothing$ for every edge $\{v, w\} \in Y$. Such a set X' is called a vertex cover of H. Without loss of generality, we assume that H has no isolated vertices.

Given an instance $\langle H, \lambda \rangle$ of the Vertex Cover problem with $H = (X, Y)$, we construct a Schelling game as follows:

- We have $|X| + |Y| - \lambda$ red agents and λ blue agents, for a total of $n = |X| + |Y|$ agents.
- To construct the topology $G = (V, E)$, we start with the graph H. Then, for each edge $e = \{v, w\} \in Y$, we add a node s_e, and two edges connecting s_e to v and w. Let $Q = \{s_e : e \in Y\}$. Then, $V = X \cup Q$, $X \cap Q = \varnothing$, and $|V| = |X| + |Q| = |X| + |Y| = n$.

We show that H has a vertex cover of size at most λ if and only if there exists an assignment in which every agent is exposed.

First, suppose that there exists a vertex cover $X' \subseteq X$ of H of size at most λ; by adding nodes, we can assume that $|X'| = \lambda$. Consider the assignment $\mathbf{v}$ in which the nodes of X' are occupied by blue agents, while all other nodes of $V \setminus X'$ are occupied by red agents. In this assignment, every agent is exposed:

- For every blue agent i occupying a node $v \in X'$, since H has no isolated nodes, there must exist an edge $e \in Y$ such that $v \in e$, and hence v is connected to s_e, which is occupied by a red agent.
- For every red agent i occupying a node $v \in X \setminus X'$, since X' is a vertex cover, v must be connected to a node $z \in X'$, which is occupied by a blue agent.
- For every red agent i occupying a node $v = s_e \in Q$, since X' is a vertex cover, at least one of e's endpoints must be in X' and is therefore occupied by a blue agent; this endpoint is connected to s_e.

Conversely, suppose that $\mathbf{v}$ is an assignment of the agents to the nodes of the topology such that every agent is exposed.

For each edge $e = \{v, w\} \in E$, let $\ell(e)$ be an arbitrary element of $\{v, w\}$. Let $X' = \{v \in X : \pi_v(\mathbf{v}) \text{ is blue}\}$ and $X_Q = \{z \in X : z = \ell(e) \text{ for some } e \text{ such that } \pi_{s_e}(\mathbf{v}) \text{ is blue}\}$. Since there are $\lambda - |X'|$ nodes in Q that are occupied by blue agents, we have $|X_Q| \leq \lambda - |X'|$ and hence $|X' \cup X_Q| \leq \lambda$. We claim that $X' \cup X_Q$ is a vertex cover for H. Indeed, consider an arbitrary edge $e = \{v, w\} \in Y$; we will argue that $e \cap (X' \cup X_Q) \neq \varnothing$. If $v \in X'$ or $w \in X'$, we are done. Otherwise, both v and w are occupied by red agents; since $\pi_{s_e}(\mathbf{v})$ is adjacent to an agent of a different type, it follows that s_e is occupied by a blue agent and $\ell(e) \in X_Q$. Hence at least one of v and w is in X_Q. This completes the proof. □

5 Social Welfare at Equilibrium

In this section, we consider the efficiency of equilibrium assignments in terms of social welfare, and provide bounds on the social price of anarchy and stability. In order to obtain meaningful bounds on the price of anarchy, we will need to make several assumptions about the agents and the structure of the topology. The next proposition provides a justification for this: it shows that the price of anarchy can be unbounded even for some very simple classes of jump and swap games.

Proposition 5.1. *For both jump and swap games, the social price of anarchy may be unbounded if there exists an agent type that consists of a single agent, or if the topology is a disconnected graph.*

Proof. We will first construct a Schelling game where there is a type T that consists of a single agent and the social price of anarchy is unbounded. In our construction, the topology is a star and there exists another type T' with $|T'| > 1$. Then any assignment that places the agent of type T in the center node of the star is an equilibrium with social welfare 0. On the other hand, it is possible to achieve positive social welfare by assigning one of the agents of type T' to the center of the star.

Next, we construct a Schelling game where the topology is disconnected and the social price of anarchy is unbounded. Our game has two agents per type. The topology consists of (at least) $n - 2$ isolated nodes, and a path with two nodes. Consider an assignment that places two agents of different types in the nodes of the path. This assignment is an equilibrium with social welfare 0: indeed, no agent has an incentive to deviate to an empty isolated node, and neither of the two agents on the path wants to swap with any other agent. However, it is possible to achieve positive social welfare by assigning two agents of the same type to the nodes of the path. □

In light of Proposition 5.1, in the rest of this section we will focus on games where each type consists of at least two agents and the topology is a connected graph. We will distinguish between jump and swap games, as they require different techniques.

5.1 Jump Games

We start by establishing bounds on the price of anarchy for fully-strategic jump games.

Theorem 5.2. *For any fixed $k \geq 2$, the social price of anarchy of fully-strategic k-jump games with connected topology and at least two agents per type is $\Theta(n)$.*

Proof. For the lower bound, consider the following fully-strategic k-jump game. There are $n > 2k$ agents: $|T_i| = 2$ agents of type T_i for each $i \in [k-1]$, and $|T_k| = n - 2(k-1)$ agents of type T_k. The topology is a star with $n+1$ nodes. It is easy to see that any assignment that does not place an agent of type T_k in the center of the star is an equilibrium with social welfare $1 + \frac{1}{n-1} = \frac{n}{n-1}$. On the other hand, any assignment that places an agent of type T_k in the center of the star has social welfare

$$n - 2(k-1) - 1 + \frac{n-2(k-1)-1}{n-1} = n - 2(k-1)\left(1 + \frac{1}{n-1}\right) = \frac{n}{n-1} \cdot (n - 2k + 1).$$

Hence, the price of anarchy is at least $n - 2k + 1$, i.e., $\Omega(n)$ for any fixed k.

For the upper bound, consider an arbitrary fully-strategic k-jump game with $n_i \geq 2$ agents of type T_i for each $i \in [k]$, 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 equilibrium assignment. Recall that we assume that the number of available nodes exceeds the number of agents, and the topology is connected. Hence, there must exist some empty node v with at least one non-empty neighbor. Suppose that v is connected to x_i nodes containing agents of type T_i, for $i \in [k]$, and let $s = \sum_{i \in [k]} x_i \geq 1$. By deviating to v, an agent of type T_i would get utility $\frac{x_i}{s}$ if her current location 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 an incentive to deviate, the utility of each agent is at least the utility she would get by deviating to v. Therefore, the social welfare at equilibrium is at least

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

where the second inequality holds since

$$(n_i - x_i)\,\frac{x_i}{s} + x_i \cdot \frac{x_i - 1}{s-1} \geq (n_i - 1)\,\frac{x_i}{s} \quad \text{for all } i \in [k],$$

and the third inequality holds since $n_i \geq 2$ for every $i \in [k]$. $\square$

To establish the lower bound of $\Omega(n)$ in Theorem 5.2, we considered a family of games with $\Theta(n)$ agents of type T_k and just two agents of any other type. Hence, one may expect that the price of anarchy is lower in games with the same number of agents per type; we refer to such games as *balanced*. Our next theorem shows that this is indeed the case.

Theorem 5.3. *The social price of anarchy of fully-strategic balanced k-jump games with connected topology is between $2k - 1$ and $2k$.*

Proof. The upper bound follows directly from Theorem 5.5, which establishes this bound for a more general setting.

For the lower bound, consider the following fully-strategic balanced k-jump game. For each $i \in [k]$, there are two agents of type T_i; hence there are $n = 2k$ agents in total. The topology (V, E) is a spider graph with $3k$ branches, among which k branches have length 2 while $2k$ branches have length 1; formally, $V = \{\alpha_0\} \cup \{\alpha_i : i \in [3k]\} \cup \{\beta_i : i \in [k]\}$, and $E = \{\{\alpha_0, \alpha_i\} : i \in [3k]\} \cup \{\{\alpha_i, \beta_i\} : i \in [k]\}$.

Consider the assignment where, for each $i \in [k]$, the agents of type T_i occupy the nodes of the i-th long branch. This assignment is optimal, as the utility of each agent is 1. Hence, the optimal social welfare is exactly $n = 2k$. On the other hand, the assignment where the agents occupy α_0 (the center of the star) as well as the nodes $\alpha_{k+1}, \dots, \alpha_{3k-1}$ is an equilibrium. Let us assume that in this assignment the center of the star is occupied by an agent of type T_1. Then, this agent has utility $\frac{1}{n-1}$, the other agent of type T_1 has utility 1, and all other agents have utility 0. Thus, the social welfare in this equilibrium is $\frac{1}{n-1} + 1 = \frac{n}{n-1}$. The price of anarchy is therefore at least $n - 1 = 2k - 1$. $\square$

As we saw, the social price of anarchy of fully-strategic jump games improves significantly if we focus on balanced games, i.e., games where there is the same number of agents of each type. We can extend the concept of balancedness to games with stubborn agents. Recall that R and S denote the set of strategic and stubborn agents, respectively. We say that a game is *strongly-balanced* if $|T_i \cap R| = |T_j \cap R|$ and $|T_i \cap S| = |T_j \cap S|$ for all $i, j \in [k]$. The next result shows that, in order to obtain a meaningful bound on the price of anarchy in strongly-balanced jump games, we have to also require that there are at least two strategic agents per type.

Proposition 5.4. *The social price of anarchy of strongly-balanced k-jump games is unbounded for every $k \geq 2$.*

Proof. To construct the topology, we create $k - 1$ paths with two nodes and one path with three nodes; let α_i and β_i be the first two nodes of the i-th path. For $i \in [k-1]$, we connect α_i to α_{i+1}. There are two agents per type (one stubborn agent and one strategic agent), and, for $i \in [k]$, node α_i is occupied by the stubborn agent of type T_i. Thus, in the optimal assignment, the strategic agent of type T_i occupies node β_i and has utility 1, so the social welfare is k. However, there is also an equilibrium in which for each $i \in [k]$ the node β_i is occupied by the strategic agent of type T_{i+1} (where the subscript is modulo k). In this equilibrium all strategic agents have utility 0. $\square$

On the other hand, if there are at least two strategic agents per type, we obtain an almost-tight bound on the price of anarchy in strongly-balanced games.

Theorem 5.5. *The social price of anarchy of strongly-balanced k-jump games with at least two strategic agents per type and connected topology is between $\frac{3k-1}{3k+1} \cdot 2k$ and $2k$.*

Proof. For the lower bound, consider a slight variation of the construction in the proof of Theorem 5.3 (where we consider balanced fully-strategic k-jump games). There are two strategic agents and one stubborn agent per type, so $n = 3k$. To construct the topology, we start with a star with center α_0 and leaves $\alpha_1, \dots, \alpha_n$, create nodes β_i and γ_i for $i \in [k]$, and add edges $\{\alpha_n, \beta_i\}$ and $\{\beta_i, \gamma_i\}$ for each $i \in [k]$. There is a stubborn agent of type T_i in node α_i for each $i \in [k]$. In the optimal assignment, the strategic agents of type T_i are placed in β_i and γ_i, so that the social welfare is $2k$. However, there is also an equilibrium assignment in which the strategic agents occupy the nodes α_0 and $\alpha_{k+1}, \dots, \alpha_{n-1}$, with node α_n remaining empty. In this equilibrium the social welfare is $1 + \frac{2}{3k-1} = \frac{3k+1}{3k-1}$. Hence, the price of anarchy is at least $\frac{3k-1}{3k+1} \cdot 2k$.

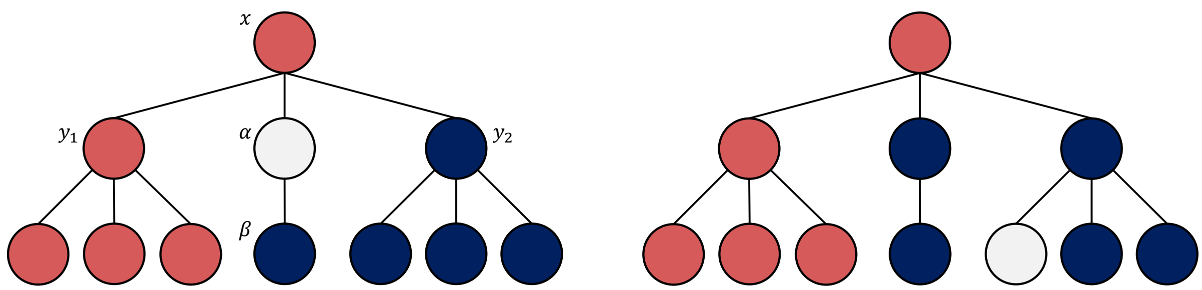

Figure 5: The topology used in the proof of Theorem 5.6. The assignment on the left is the unique equilibrium of this 2-jump game (up to symmetry), while the assignment on the right is the optimal one.

For the upper bound, consider an arbitrary strongly-balanced k-jump game with t strategic and $\frac{n}{k} - t$ stubborn agents per type, for some integer $t \geq 2$. We will show that the social welfare of any equilibrium assignment is at least $t-1$. Since the utility of every strategic agent is at most 1, the optimal social welfare is at most kt, so we can upper-bound the price of anarchy as $\frac{kt}{t-1} \leq 2k$.

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 nodes containing agents of type T_i, for $i \in [k]$, and x_i^R of them are strategic. Also, let $s = \sum_{i\in[k]} x_i \geq 1$. Now, consider a strategic agent of type T_i. A deviation to v would give her utility $\frac{x_i}{s}$ if her current location is not connected to v, and utility $\frac{x_i - 1}{s-1}$ otherwise; by convention, $\frac{0}{0} = 0$. Since at equilibrium no strategic 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

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

where the second inequality holds since

$$\frac{1}{s}\left(tx_i - x_i^R x_i\right) + x_i^R \cdot \frac{x_i - 1}{s-1} \geq \frac{1}{s}\left(tx_i - x_i^R\right) \quad \text{for all } i \in [k],$$

and the last inequality follows since $x_i^R \leq x_i$ for each $i \in [k]$. $\square$

We conclude our analysis of jump games by showing that even the best equilibrium need not be socially optimal. Specifically, we describe a fully-strategic balanced 2-jump game in which the ratio between the optimal social welfare and the maximum social welfare in an equilibrium assignment is at least $34/33$. This bound was very recently improved to approximately 2 by Kanellopoulos et al. [2021b].

Theorem 5.6. *The social price of stability of fully-strategic balanced* 2*-jump games is at least* $34/33$.

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

Let $\mathbf{v}'$ be the following assignment: node x, node y_1 and all of the children of y_1 are occupied by red agents, while node α, node β, node y_2 and two of the children of y_2 are occupied by blue agents. The social welfare of this assignment is $\mathrm{SW}(\mathbf{v}') = 34/4$, offering an improvement over $\mathbf{v}$. However, 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 node (a child of y_2) in order to increase her utility to 1.

To complete the proof, we need to argue that $\mathbf{v}$ is a social welfare maximizing equilibrium. To this end, we establish some properties of equilibrium assignments.

- *In an equilibrium node x must be occupied.* Indeed, suppose that x is left empty. Suppose first that nodes y_1, α and y_2 are occupied by agents of the same type (say, red). Then at least one of them will be connected to a node occupied by a blue agent and hence will have an incentive to deviate to x, so as to be connected to red agents only. On the other hand, suppose that y_1, α and y_2 are occupied by two agents of one type (say, red) and one agent of the other type (say, blue). Then there exists a blue agent occupying a leaf node whose only neighbor is a red agent; this blue agent can then increase her utility from 0 to $1/3$ by deviating to x.

- *In an equilibrium nodes y_1 and y_2 must be occupied.* Indeed, suppose that one of these nodes, say y_1, is left empty. We can assume without loss of generality that node x is occupied by a red agent. If there is a child of y_1 occupied by a red agent then this agent can increase her utility from 0 to positive by deviating to y_1. Thus, the children of y_1 can only be occupied by blue agents; as at least two children of y_1 are occupied, an agent in one of these nodes can increase her utility from 0 to positive by deviating to y_1.

- *In an equilibrium all leaf nodes are occupied.* Suppose that z is an unoccupied leaf node. We can assume without loss of generality that z is not a child of y_1, and that the parent of z is red. Suppose first that y_1 is occupied by a red agent. Then it cannot be the case that all four neighbors of y_1 are occupied by red agents, so the agent at y_1 has an incentive to deviate to z to increase her utility to 1. Thus, y_1 must be occupied by a blue agent. Now, if one of the four neighbors of y_1 is occupied by a red agent, this agent has an incentive to deviate to z; thus, both x and all children of y_1 must be occupied by blue agents. Consider, then, the node in $\{\alpha, y_2\}$ that is not the parent of z. This node must be occupied by a red agent, and it is adjacent to x, which is occupied by a blue agent; hence, this red agent has an incentive to deviate to z.

We can now conclude that in an equilibrium α must be empty. Moreover, the agent at x and the agent at β cannot belong to the same type, since otherwise the agent at β could increase her utility by deviating to α. Thus, any agent could achieve a positive utility (namely, $1/2$) by moving to α. It follows that the agents at the leaves of y_1 must have the same type as the agent at y_1, and the agents at the leaves of y_2 must have the same type as the agent at y_2: otherwise, there will be an agent at a leaf who has zero utility and can therefore benefit from moving to α. It now follows by a counting argument that agents at y_1 and y_2 cannot belong to the same type. Hence, any equilibrium is essentially equivalent to $\mathbf{v}$: the agents at the leaves of y_1 and y_2 have utility 1, the agent at x has utility $1/2$, the agent at β has utility 0, and finally, the agents at y_1 and y_2 contribute $1 + \frac{3}{4}$ to the social welfare. Hence, the social welfare at any equilibrium of this game is $\frac{33}{4}$. The lower bound on the price of stability follows. □

5.2 Swap Games

We now turn our attention to swap games. We start by presenting bounds on the price of anarchy of fully-strategic 2-swap games, both for the general and for the balanced case. Throughout this section, we assume that the topology is a connected graph $G = (V, E)$ with $|V| = n$, and that there are at least two agents per type.

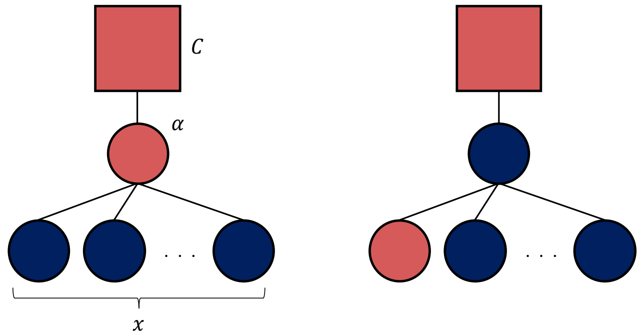

Figure 6: The topology of the lower bound instance used in the proof of Theorem 5.8. The big square represents a clique of nodes. The assignment on the left is an equilibrium, while the assignment on the right is optimal. In both cases, all nodes of the clique are occupied by red agents.

Theorem 5.7. *The social price of anarchy of fully-strategic* 2*-swap games with at least two strategic agents per type and connected topology is* $\Theta(n)$*.*

Proof. For the lower bound, consider a fully-strategic 2-swap game with star topology, $n-2$ red agents, and 2 blue agents. There is an equilibrium where one of the blue agents occupies the center of the star. The social welfare in this equilibrium $1+\frac{1}{n-1}=\frac{n}{n-1}$. In contrast, in an optimal assignment the central node is occupied by a red agent, and the social welfare is $n-3+\frac{n-3}{n-1}=\frac{n}{n-1}\cdot(n-3)$. Therefore, the price of anarchy is at least $n-3$.

For the upper bound, consider a fully-strategic 2-swap game with n agents: $n_r\geq 2$ red agents and $n_b\geq 2$ blue agents. Let $\mathbf{v}$ be an equilibrium assignment of this game.

Suppose first that each agent has positive utility in $\mathbf{v}$. Then the utility of each agent is at least $1/n$, so the social welfare in $\mathbf{v}$ is at least 1; the bound then follows since the optimal social welfare is at most n.

Now, suppose there is some agent, say ℓ, whose utility is zero; we can assume without loss of generality that ℓ is red, and hence all agents connected to ℓ are blue. If ℓ is connected to all n_b blue agents, then, since the topology is connected, at least one of these blue agents, say t, is connected to another red agent. If t is connected to x blue agents, her utility is at most $x/(x+2)\leq(n_b-1)/(n_b+1)$, where the inequality holds since $x\leq n_b-1$ and the function $z/(z+2)$ is increasing in z. But then ℓ and t have an incentive to swap: a swap would increase ℓ's utility from 0 to at least $1/n$ and it would increase t's utility from at most $(n_b-1)/(n_b+1)$ to $(n_b-1)/n_b$.

Thus, there must exist a blue agent that is not connected to ℓ; let t' be some such agent. Now, if t' is connected to some red agent, her utility is less than 1, and ℓ and t' have an incentive to swap: a swap would increase ℓ's utility from 0 to at least $1/n$ and it would increase t''s utility to 1. We conclude that t is not connected to any red agents; since the topology is connected and the number of nodes equals the number of agents, it follows that t''s utility in $\mathbf{v}$ is at least 1. Hence, $\mathrm{SW}(\mathbf{v})\geq 1$; this implies that the price of anarchy is at most n. ☐

In contrast, when the fully-strategic 2-swap game is balanced, the social price of anarchy is bounded by 4. Later, in Theorem 5.11), we will show that the bound holds even in the presence of stubborn agents. We should note that after the conference version of our paper, Bilò et al. [2020] improved the bound to 3 for the fully-strategic case.

Theorem 5.8. *The social price of anarchy of fully-strategic balanced* 2*-swap games with at least two agents per type and connected topology is between* $667/324\approx 2.05864$ *and* 4*.*

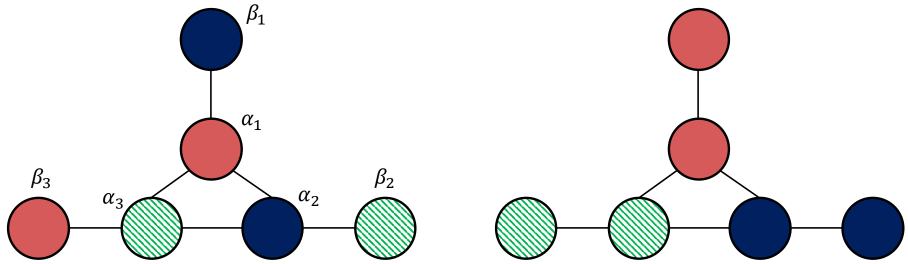

Figure 7: The topology of the k-swap game considered in the proof of Theorem 5.9 for $k = 3$. Here, T_1 = red, T_2 = blue and T_3 = green. The assignment on the left is an equilibrium with zero social welfare, while the assignment on the right is an optimal one, in which all agents have a positive utility.

Proof. The upper bound follows from Theorem 5.11, which will be proved later in the paper. For the lower bound, consider a fully-strategic balanced 2-swap game with the following topology: there is a node α of degree $x + 1$ that is connected to x leaf nodes and to one node in a clique C of size $x - 1$, where x is to be chosen later. There are x agents of each type. There is an equilibrium $\mathbf{v}$ where α and all nodes of C are occupied by red agents, and all leaf nodes are occupied by blue agents, see Figure 6. Hence,

$$\mathrm{SW}(\mathbf{v}) = x - 1 + \frac{1}{x+1} = \frac{x^2}{x+1}.$$

Let r be the red agent that is located in α. For the assignment $\mathbf{v}^*$ obtained from $\mathbf{v}$ by swapping r with one of the blue agents we have

$$\mathrm{SW}(\mathbf{v}^*) = 2x - 3 + \frac{x-2}{x-1} + \frac{x-1}{x+1}.$$

Hence, the price of anarchy is at least

$$\frac{2x^3 - x^2 - 5x + 2}{x^2(x-1)};$$

this expression attains its maximum value $667/324 \approx 2.05864$ over $\mathbb{N}$ at $x = 9$. □

In contrast, for three or more types, the social price of anarchy can be unbounded even in fully-strategic balanced swap games.

Theorem 5.9. *The social price of anarchy of fully-strategic balanced k-swap games is unbounded for every $k \geq 3$.*

Proof. Consider a k-swap game with $n = 2k$ agents such that there are exactly two agents of each of the $k \geq 3$ types $T_1, \ldots, T_k$. The topology G consists of $2k$ nodes $\alpha_1, \beta_1, \ldots, \alpha_k, \beta_k$, so that $\alpha_1, \ldots, \alpha_k$ form a cycle and for each $\ell \in [k]$ the nodes α_ℓ and β_ℓ are connected; see Figure 7 for the topology and the equilibrium assignment discussed in the following for $k = 3$. Throughout the proof, all subscripts are interpreted modulo k, i.e., we set $T_0 := T_k$, $\alpha_0 := \alpha_k$, $\beta_0 := \beta_k$.

Consider an assignment $\mathbf{v}$ where for each $\ell \in [k]$ the agents of type T_ℓ occupy the nodes α_ℓ and $\beta_{\ell-1}$. This assignment is an equilibrium. Indeed, the only way the agent of type T_ℓ located at α_ℓ can increase her utility is by moving to node $\alpha_{\ell-1}$, so as to connect with the agent of type T_ℓ located at $\beta_{\ell-1}$. However, in $\mathbf{v}$ the node $\alpha_{\ell-1}$ is occupied by an agent of type $T_{\ell-1}$, who is not interested in moving to α_ℓ. Similarly, the agent of type T_ℓ located at $\beta_{\ell-1}$ cannot swap with another agent, because the only agent who would like to move to $\beta_{\ell-1}$ is the agent of type $T_{\ell-1}$ located at $\beta_{\ell-2}$, and the agent of type T_ℓ does not benefit from moving to $\beta_{\ell-2}$. Note that $\mathrm{SW}(\mathbf{v}) = 0$.

On the other hand, consider the assignment $\mathbf{v}^*$ in which α_ℓ and β_ℓ are occupied by the two agents of type T_ℓ, for every $\ell \in [k]$. We have $\mathrm{SW}(\mathbf{v}^*) > 0$, so the social price of anarchy is unbounded. □

Theorem 5.9 can be extended to games with stubborn agents. To see this, observe that we can modify the construction in the proof of this theorem by assuming that one agent of each type T_ℓ is a stubborn agent placed at α_ℓ, for $\ell \in [k]$. Then, the same arguments show that an optimal assignment guarantees a positive utility to every strategic agent, while there exists an equilibrium in which all strategic agents have utility 0. In fact, in the presence of stubborn agents, the price of anarchy remains unbounded even for strongly-balanced 2-swap games.

Proposition 5.10. *The social price of anarchy of strongly-balanced* 2*-swap games is unbounded.*

Proof. Consider a strongly-balanced 2-swap game with four agents: r_1 is a stubborn red agent, r_2 is a strategic red agent, b_1 is a stubborn blue agent, and b_2 is a strategic blue agent. The topology is a path of length 4. Suppose that r_1 occupies the first node and b_1 occupies the second node. In the optimal assignment, the third node is occupied by b_2, who obtains positive utility. However, the assignment where the third node is occupied by r_2 and the fourth node is occupied by b_2 (so that the utility of each agent is 0) is an equilibrium as well, as r_2 does not benefit from swapping with b_2. Hence, the price of anarchy is unbounded. □

Similarly to jump games (Theorem 5.5), to obtain a meaningful bound on the social price of anarchy for general 2-swap games, we need to restrict our attention to the class of strongly-balanced games with at least two strategic agents per type.

Theorem 5.11. *The social price of anarchy of strongly-balanced* 2*-swap games with at least two strategic agents per type and connected topology is between* $8/3 \approx 2.667$ *and* 4*.*

Proof. For the lower bound, consider a 2-swap game such that each type consists of two strategic agents and one stubborn agent. The topology is a path with five nodes, say, $\alpha_1, \dots, \alpha_5$, with the node α_4 connected to another node β. The stubborn red agent is positioned at node α_2 and the stubborn blue agent at node α_3. Now, consider the equilibrium assignment $\mathbf{v}$ according to which α_1 and α_4 are occupied by the strategic red agents, and α_5 and β by the strategic blue agents. Then, the strategic red agent at α_1 has utility 1 and all other strategic agents have utility 0, yielding $\mathrm{SW}(\mathbf{v}) = 1$. On the other hand, consider the assignment $\mathbf{v}^*$ obtained from $\mathbf{v}$ by swapping the strategic red agent at α_4 with the strategic blue agent at α_5. Then, the strategic red agent at α_5 has utility 0, the strategic red agent at α_1 continues to get utility 1, the strategic blue agent at β now has utility 1, and the strategic blue agent at α_4 has utility $2/3$, yielding $\mathrm{SW}(\mathbf{v}^*) = 8/3$. Consequently, the price of anarchy is at least $8/3$.

For the upper bound, consider a strongly-balanced 2-swap game with $t \geq 2$ strategic and $\frac{n}{2} - t$ stubborn agents per type (red and blue). Fix an equilibrium assignment $\mathbf{v}$. We will prove that the social welfare in $\mathbf{v}$ is at least $t/2$; since the optimal social welfare is at most $2t$, this gives an upper bound of 4 on the price of anarchy. Let B_0 denote the set of strategic blue agents with utility 0 in $\mathbf{v}$ and let R_0 denote the set of strategic red agents with utility 0 in $\mathbf{v}$. We consider three cases.

$\boldsymbol{B_0 \neq \varnothing, R_0 \neq \varnothing}$. We claim that in this case $t-1$ strategic agents of one type must all get utility 1 at equilibrium. We make the following observations:

- All agents in R_0 must be connected to all agents in B_0. Indeed, if an agent $r \in R_0$ is not connected to an agent $b \in B_0$, then r and b would prefer to swap and increase their utility from 0 to 1, by getting connected to agents of their own type only.

- It holds that $|R_0| = 1$ or $|B_0| = 1$. Indeed, suppose that $|R_0| \geq 2$, $|B_0| \geq 2$. But then an agent $r \in R_0$ and an agent $b \in B_0$ can increase their utility by swapping. From now on, we will assume that R_0 is a singleton, $R_0 = \{r\}$.

- If $|B_0| > 1$ then at least $t-1$ red agents obtain utility 1. Indeed, let b_1, b_2 be two distinct agents in B_0. Note that they are both connected to $r \in R_0$. If neither b_1 nor b_2 occupy a leaf node, then both of them are connected to red agents other than r, so b_1 and r would prefer to swap, a contradiction. Hence at least one of them (say, b_1) occupies a leaf node; note that the only agent it is connected to is r. But then the utility of any strategic red agent $r' \neq r$ is equal to 1: otherwise, r' and b_1 can both benefit from swapping. From now on, we will assume that B_0 is a singleton, $B_0 = \{b\}$.

- At least one of the agents in $\{r, b\}$ occupies a leaf node. Otherwise, if both nodes occupied by these agents have degree at least two, the agents would prefer to swap with each other. From now on, we will assume that b occupies a leaf node; note that the parent of this node is occupied by r.

Let $R_{>0}$ denote the set of strategic red agents with positive utility. By the above observations we have $|R_{>0}| = t-1$. Consider an agent $r' \in R_{>0}$; note that r' is not connected to b. It follows that r''s utility is exactly 1, as otherwise r' and b would prefer to swap. Hence, we have identified $t-1$ agents with utility 1, as claimed. In particular, since $t \geq 2$, the social welfare is at least $t/2$.

$\boldsymbol{B_0 \neq \varnothing, R_0 = \varnothing}$ (the case $\boldsymbol{B_0 = \varnothing, R_0 \neq \varnothing}$ can be handled using symmetric arguments). We denote by R_1 the set of strategic red agents with utility 1, and by $R_<$ the set of strategic red agents whose utility is strictly less than 1; note that $|R_1| + |R_<| = t$.

We have $\mathrm{SW}(\mathbf{v}) \geq |R_1|$, so if $|R_<| \leq 1$, then we have $\mathrm{SW}(\mathbf{v}) \geq t-1$. Now, suppose that $|R_<| \geq 2$. We will prove that in this case the utility of each agent in $R_<$ is at least $1/2$ and hence

$$\mathrm{SW}(\mathbf{v}) \geq |R_1| + \frac{|R_>|}{2} \geq \frac{t}{2}.$$

Each agent in B_0 has to be connected to all agents in $R_<$, as otherwise a non-adjacent pair of agents $r \in R_<, b \in B_0$ would like to swap. Therefore, if $|B_0| \geq 2$, any agent in B_0 would like to swap with any agent in $R_<$ to get a positive utility. Thus, in $\mathbf{v}$ no agent $r \in R_<$ wants to swap with any agent in B_0. Since each agent in B_0 is connected to all agents in $R_<$ and no blue agents (and possibly to some stubborn red agents), the utility that agent r would get by swapping with an agent in B_0 is at least $\frac{|R_<|-1}{|R_<|} \geq 1/2$, so $u_r(\mathbf{v}) \geq 1/2$. On the other hand, suppose that $|B_0| = 1$, i.e., $B_0 = \{b\}$ for some strategic blue agent b. If an agent $r \in R_<$ is connected to a (strategic or stubborn) blue agent other than b, then b wants to swap with r, and hence r does not want to swap with b, which means that her utility is at least $1/2$. Otherwise, r is connected to at least one red agent and at most one blue agent (i.e., b), so $u_r(\mathbf{v}) \geq 1/2$.

$\boldsymbol{B_0 = \varnothing, R_0 = \varnothing}$. Let us pair up strategic red and strategic blue agents arbitrarily, creating t red-blue pairs. We will show that for every such pair (r, b) we have $u_r(\mathbf{v}) + u_b(\mathbf{v}) \geq 1/2$; by summing over all pairs, we then obtain $\mathrm{SW}(\mathbf{v}) \geq t/2$.

Consider a pair (r, b). Since $\mathbf{v}$ is an equilibrium, at least one of these agents has no incentive to swap positions. Assume that r does not want to swap. If r and b are *not* neighbors in $\mathbf{v}$, then it must be that $u_r(\mathbf{v}) \geq 1 - u_b(\mathbf{v})$ and hence $u_r(\mathbf{v}) + u_b(\mathbf{v}) \geq 1$. Now, suppose that r and b are neighbors in $\mathbf{v}$. Assume

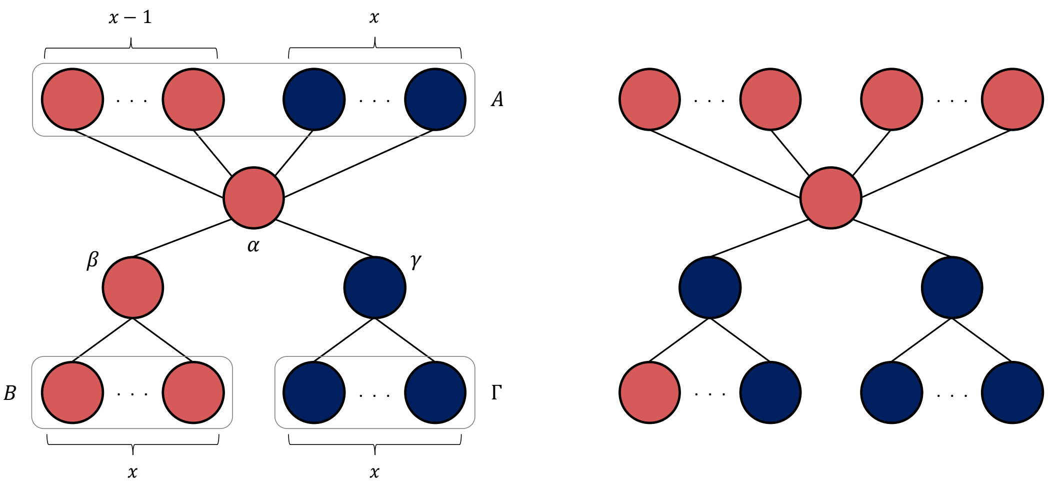

Figure 8: The topology of the 2-swap games considered in the proof of Theorem 5.12. The assignment on the left is the best equilibrium in terms of social welfare, while the assignment on the right is the optimal one.

that b has x_r red neighbors besides r, and x_b blue neighbors. Since b has positive utility, it has at least one blue neighbor, so $x_r + x_b \geq 1$. Then, $u_b(\mathbf{v}) = \frac{x_b}{x_r+x_b+1}$, and, since r does not want to swap,

$$u_r(\mathbf{v}) \geq \frac{x_r}{x_r + x_b + 1} = 1 - u_b(\mathbf{v}) - \frac{1}{x_r + x_b + 1} \geq \frac{1}{2} - u_b(\mathbf{v}),$$

or, equivalently, $u_r(\mathbf{v}) + u_b(\mathbf{v}) \geq \frac{1}{2}$, and hence $\mathrm{SW}(\mathbf{v}) \geq t/2$. The bound now follows since the maximum possible social welfare is $2t$. ☐

Next, we turn our attention to the social price of stability and show a lower bound for fully-strategic balanced 2-swap games.

Theorem 5.12. *The social price of stability of fully-strategic balanced* 2*-swap games is at least* $4/3$*.*

Proof. Let $x \geq 3$ be a parameter, and consider a 2-swap game with $2x+1$ red and $2x+1$ blue agents. The topology is a tree with a root node α, which is connected to two nodes β and γ, as well as to a set A of $2x-1$ leaf nodes. Moreover, node β is connected to a set B of x leaf nodes, and node γ is connected to a set Γ of x leaf nodes. The topology and the best equilibrium assignment (which we discuss below) are depicted in Figure 8.

We will now establish some properties of equilibria of this swap game. Without loss of generality, we assume that the root node α is occupied by a red agent. Let s be the number of blue agents that are located in A. We consider the following cases:

- $\boldsymbol{s \geq x+1}$. Let b be a blue agent occupying a node of A. If $s \geq x+1$, there are at most $x-2$ red agents in A. Thus, at least $x+2$ red agents occupy nodes of the subtrees rooted at β and γ. Since each of these subtrees has $x+1$ nodes, at least one of these red agents, say agent r, is connected to at least one blue agent. But then this assignment is not an equilibrium, since agent r can swap positions with b to increase their utilities from strictly smaller than 1 and 0 to 1 and positive, respectively.

- $\boldsymbol{s = x}$. Let b be a blue agent occupying a node of A. If $s = x$, the remaining $x-1$ nodes of A are occupied by red agents. We claim that the only equilibrium assignment $\mathbf{v}_1$ (up to symmetries) is such that all nodes of the β-subtree are occupied by red agents, and all nodes of the γ-subtree are occupied by blue agents. Indeed, if either of these subtrees contains agents of both types, then in particular it contains a red agent connected to a blue agent, who has a mutually beneficial swap with b. The social welfare of $\mathbf{v}_1$ is

$$\mathrm{SW}(\mathbf{v}_1) = 3x + \frac{x}{x+1} + \frac{x}{2x+1} \leq 3x + 2.$$

- $\boldsymbol{1 \leq s \leq x - 1}$. Let b be a blue agent occupying a node of A. Consider the subtrees rooted at β and γ. There are at least 2 and at most x red agents that occupy nodes of these subtrees. As each of these subtrees has $x + 1$ nodes, at least one of them has to contain a red agent connected to a blue agent; this red agent then has a mutually beneficial swap with b.
- $\boldsymbol{s = 0}$. In this case, all nodes of A and α are occupied by $2x$ red agents; denote the (unique) remaining red agent by r. There is no equilibrium where r is located in a leaf node: if r is placed in a node of B, she has a mutually beneficial swap with the blue agent at γ, and if r is placed in a node of Γ, she has a mutually beneficial swap with the blue agent at β. Thus, in an equilibrium, agent r occupies either node β or node γ. The social welfare of any such equilibrium $\mathbf{v}_2$ is

$$\mathrm{SW}(\mathbf{v}_2) = 3x - 1 + \frac{2x}{2x+1} + \frac{1}{x+1} + \frac{x}{x+1} \leq 3x + 1.$$

Now consider the assignment $\mathbf{v}^*$ in which the red agents occupy node α, all nodes of A, and one node of B, while all other nodes are occupied by blue agents (see Figure 8). The social welfare of this assignment is

$$\mathrm{SW}(\mathbf{v}^*) = 4x - 2 + \frac{2x-1}{2x+1} + \frac{x-1}{x+1} + \frac{x}{x+1} \geq 4x - 2.$$

Therefore, the social price of anarchy is at least $\frac{4x-2}{3x+2}$, which tends to $4/3$ as x tends to infinity. □

In contrast, in fully-strategic k-swap games in which the topology is a δ-regular graph (i.e., the degree of each node equals to δ), there is always a Nash equilibrium that is socially optimal, i.e., the social price of stability is 1. We show this by exploiting a potential function that is similar to the one used by Chauhan et al. [2018] and Echzell et al. [2019] to show the existence of equilibria in such games.

Theorem 5.13. *The social price of stability in fully-strategic k-swap games with topology that is a δ-regular graph is* 1.

Proof. We use a potential function argument inspired by Echzell et al. [2019]. For an assignment $\mathbf{v}$, let $\overline{\Phi}(\mathbf{v})$ be the number of edges connecting agents of the same type. Consider a pair of agents (i, j) such that i is of type T_x and j is of type T_y, with $y \neq x$. Since i and j want to swap positions if and only if they can both increase their utility, and since $|N_i(\mathbf{v})| = |N_j(\mathbf{v})| = \delta$ for any assignment $\mathbf{v}$, it follows that if i and j want to swap positions, then each of them is adjacent to a higher number of friends after the swap. Thus, each swap increases the potential function $\overline{\Phi}$. Consequently, an assignment $\mathbf{v}$ that maximizes $\overline{\Phi}$ must be an equilibrium. We will now argue that every such assignment also maximizes the social welfare. Recall that F_i denotes the set of friends of agent i, and R is the set of strategic agents. Given an assignment $\mathbf{v}$, we can relate its social welfare and the value of $\overline{\Phi}(\mathbf{v})$ as follows:

$$\mathrm{SW}(\mathbf{v}) = \sum_{i \in R} u_i(\mathbf{v}) = \sum_{i \in R} \frac{|N_i(\mathbf{v}) \cap F_i|}{|N_i(\mathbf{v})|} = \frac{1}{\delta} \cdot \sum_{i \in R} |N_i(\mathbf{v}) \cap F_i| = \frac{2}{\delta} \cdot \overline{\Phi}(\mathbf{v}).$$

Thus, if an assignment $\mathbf{v}^*$ is such that $\overline{\Phi}(\mathbf{v}^*) \geq \overline{\Phi}(\mathbf{v})$ for any assignment $\mathbf{v}$, then $\mathbf{v}^*$ is an equilibrium and $\mathrm{SW}(\mathbf{v}^*) \geq \mathrm{SW}(\mathbf{v})$ for any assignment $\mathbf{v}$, which establishes our claim. □

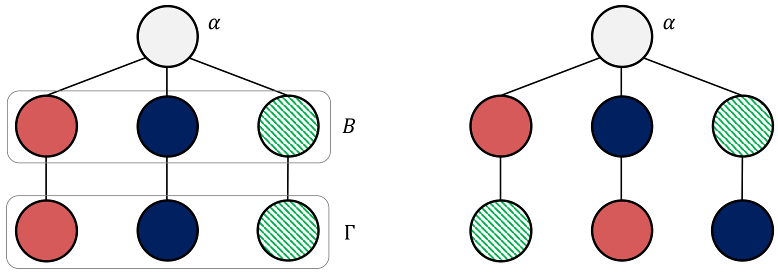

Figure 9: The topology used in the proof of Theorem 6.1 for $k = 3$. The assignment on the left is the unique equilibrium in which the degree of integration is 0 (no agent is exposed), while the assignment on the right is an optimal one in which all agents are exposed.

6 Degree of Integration at Equilibrium

We now investigate whether equilibrium assignments can be diverse, by bounding the price of anarchy and stability in terms of the degree of integration; recall that this benchmark counts the number of agents who are exposed, that is, they have at least one neighbor of a different type. In this section we focus entirely on fully-strategic games.

For jump games, it turns out that the integration price of stability (and thus the price of anarchy) is unbounded for every $k \geq 2$. In the proof of the following theorem we present a balanced fully-strategic k-jump game in which the unique equilibrium assignment is such that the different types are completely segregated, leading to zero degree of integration.

Theorem 6.1. *For any $k \geq 2$, the integration price of stability of fully-strategic balanced k-jump games is unbounded.*

Proof. Consider a fully-strategic balanced k-jump game with two agents per type; hence $n = 2k$. The topology is a tree with $2k + 1$ nodes in total: the root node α has k children (set B), each of which has one child (set Γ); see Figure 9 for an example of the topology and the assignments discussed in the rest of the proof for $k = 3$.

To reason about the structure of equilibrium assignments, we proceed by case analysis. We consider all possibilities for the (unique) empty node.

- Node α is empty. Then, all nodes in B must be occupied by agent of different types. Indeed, if there are two agents of type T_i located in nodes of B, either of them would want to jump to α in order to connect with their (only) friend. Consequently, the remaining agents that occupy the nodes of Γ must be assigned so that they are connected to friends only, as otherwise they would have an incentive to move to α. This leads to an equilibrium in which all types are segregated and the degree of integration is 0; an example of this assignment for $k = 3$ is depicted in the left part of Figure 9.

- Some node $\beta \in B$ is empty. Let $\gamma \in \Gamma$ be the child of β, and assume that γ is occupied by an agent of type T_i. This means that the second agent of type T_i has utility 0 and hence has an incentive to move to β. Thus, such an assignment cannot be an equilibrium.

- Some node $\gamma \in \Gamma$ is empty. Let $\beta \in B$ be the parent of γ. Assume that β is occupied by an agent of type T_i. The second agent of type T_i is then connected to at least one agent of type other than

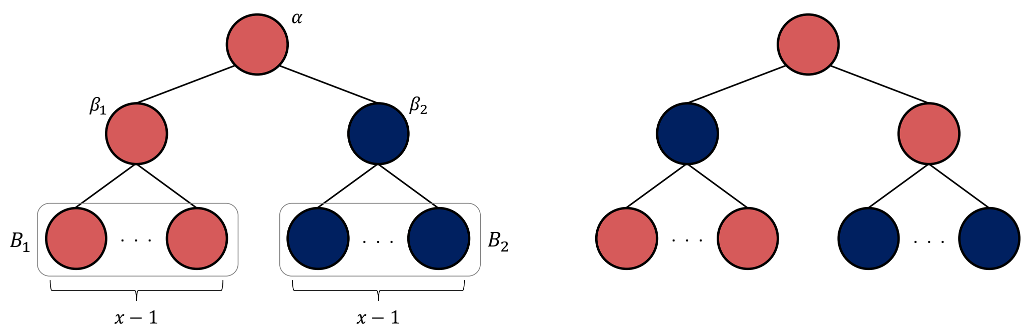

Figure 10: The topology of the k-swap game considered in the proofs of Theorems 6.2 and 6.3 for $k = 2$. The assignment on the left is the unique equilibrium (for $k = 2$) in which only two agents are exposed, while the assignment on the right is an optimal one in terms of the degree of integration.

T_i, and therefore has an incentive to move to γ. Consequently, such an assignment cannot be an equilibrium.

By the above discussion, we conclude that the unique (up to symmetries) assignment is such that α is empty, and each connected pair $(\beta, \gamma) \in B \times \Gamma$ is occupied by agents of the same type. Since we can assign the agents to the nodes of the topology so that everyone is exposed (e.g., by assigning the agents of type T_i to the i-th node in B and the $(i-1)$-th node in Γ, where indices are taken modulo k; see the right part of Figure 9), the integration price of stability is unbounded. □

For swap games, the picture becomes more interesting. We start by showing a tight bound on the price of anarchy, which indicates that in the worst case, agents of different types are highly segregated. However, as the number of types increases, equilibria become more diverse and the price of anarchy decreases.

Theorem 6.2. *For any $k \geq 2$, the integration price of anarchy of fully-strategic k-swap games is n/k.*

Proof. For the upper bound, consider an arbitrary fully-strategic k-swap game with n agents. By definition, the degree of integration is at most n. Since the topology is a connected graph and the number of agents is equal to the number of nodes, in any assignment $\mathbf{v}$ at least one agent per type must be exposed. Hence, $\text{DI}(\mathbf{v}) \geq k$, and the integration price of anarchy is at most n/k.

For the lower bound, consider a fully-strategic k-swap game with $n = kx + 1$ agents such that there are $x + 1$ agents of type T_1 and x agents of type T_ℓ for every $\ell \in [k] \setminus \{1\}$. The topology is a tree with root node α that has k children nodes $\beta_1, \ldots, \beta_k$, each of which has $x - 1$ children leaf nodes of its own; see Figure 10 for an example of this topology for $k = 2$.

One can assign the agents to the nodes of the topology so that each agent is exposed; thus the maximum possible degree of integration is n. However, there is an equilibrium assignment $\mathbf{v}$ in which α is occupied by an agent of type T_1 and for each $\ell \in [k]$ all nodes of the β_ℓ-subtree are occupied by agents of type T_ℓ. In $\mathbf{v}$ only the agent in α and the agents in nodes β_ℓ, $2 \leq \ell \leq k$, are exposed, yielding degree of integration $\text{DI}(\mathbf{v}) = k$, and the bound follows. □

Next, we consider the integration price of stability for swap games. Using the same instance as in the proof of Theorem 6.2, we show a lower bound that depends linearly on the number of agents for the fundamental case of two agent types. This bound is tight by Theorem 6.2, and indicates that, in some instances, we cannot avoid high levels of segregation at equilibrium even in the best case.

Theorem 6.3. *The integration price of stability of fully-strategic* 2*-swap games is* $n/2$.

Proof. By Theorem 6.2, it suffices to show the lower bound. Consider a 2-swap game with $x+1$ red agents and x blue agents, for a total of $n = 2x+1$ agents. The topology is the same as in Theorem 6.2: a tree consisting of a root node α with two children nodes β_1 and β_2, each of which has $x-1$ children leaf nodes of its own (sets B_1 and B_2); see Figure 10. The optimal degree of integration is n. We will now argue that in each equilibrium assignment the blue agents occupy the nodes of the subtree rooted at β_i for some $i = 1, 2$, so that α and the nodes of the subtree rooted at β_{3-i} are occupied by red agents. The degree of integration of any such assignment is exactly 2, so the theorem follows.

Consider an assignment $\mathbf{v}$ and suppose for the sake of contradiction that agent $\pi_\alpha(\mathbf{v})$ is blue. We distinguish the following cases with regard to the agents occupying nodes β_1 and β_2.

- Both $\pi_{\beta_1}(\mathbf{v})$ and $\pi_{\beta_2}(\mathbf{v})$ are of the same type. Assume first that both of these agents are blue; as there are $x+1$ red agents, there must be a red agent r in B_1 and another red agent r' in B_2. But then the blue agent in β_2 and r can swap to increase their utility from strictly less than 1 and zero to 1 and positive, respectively. Hence, it must be the case that $\pi_{\beta_1}(\mathbf{v})$ and $\pi_{\beta_2}(\mathbf{v})$ are both red. Then some leaf node is occupied by a blue agent; assume without loss of generality that there is a blue agent b in B_1. But then b and the red agent in β_2 have an incentive to swap.

- $\pi_{\beta_1}(\mathbf{v})$ and $\pi_{\beta_2}(\mathbf{v})$ are of different types. Assume without loss of generality that β_1 is occupied by a red agent and β_2 is occupied by a blue agent. Since there are x red agents remaining, at least one of them must be in B_2. But then such an agent can swap positions with the blue agent in α so that they increase their utility from zero and $1/2$ to $1/2$ and 1, respectively.

Therefore, the agent in α must be red. Similarly to the previous case, we observe that if the agents in β_1 and β_2 are both blue, there must be red agents in both B_1 and B_2, and if the agents in β_1 and β_2 are both red, there must be blue agents in both B_1 and B_2, which means that the agent in β_1 and some agent in a node of B_2 would have an incentive to swap. Thus, the agents in β_1 and β_2 must be of different types. Assume without loss of generality that the agent in β_1 is red and the agent in β_2 is blue. Then, if there is a blue agent in B_1, by a counting argument there is also a red agent in B_2, and these two agents would have an incentive to swap. Thus, all agents in B_1 must be red and all agents in B_2 must be blue. □

7 Variants and Extensions

Throughout this paper, we have so far 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; to simplify our discussion in what follows, we will refer to this class of (jump or swap) games as *k-typed*. In this section, we introduce three modifications of this model and briefly discuss some preliminary results; a more thorough investigation of these alternative models is left for future work.

7.1 Schelling Games with Social Networks

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

By definition, k-typed games form a subclass of social games in which the social network consists of k independent cliques. Hence, our next theorem implies Theorem 3.1 in Section 3.

Theorem 7.1. *Every social (jump or swap) game with topology that 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 an arbitrary social game with topology G and set of agents $N = R \cup S$ who are connected via a social network $\mathcal{N}$.

Suppose that G is a star with center v. If v is occupied by a stubborn agent, all strategic agents are indifferent among the leaves, so they have no incentive to deviate. If v is not occupied by a stubborn agent, consider an assignment $\mathbf{v}$ that places some strategic agent i in v. All strategic agents are indifferent among the leaves, so no agent in $R \setminus \{i\}$ can benefit from a jump to another leaf, or a swap with another agent located at a leaf. Moreover, agent i cannot benefit from a jump to an empty leaf, as that would reduce her utility to 0. Finally, for agent i to benefit from a swap, she has to swap positions with a friend of hers, and all friends of i have utility 1 under $\mathbf{v}$ and therefore do not want to deviate. Thus $\mathbf{v}$ 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 the *potential of* e as

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

Let $\Phi(\mathbf{v}) = \sum_{e \in E} \phi(\mathbf{v}, e)$. We claim that Φ is an ordinal potential function for our setting, that is, for any beneficial deviation, the potential function increases. Since the function Φ takes values in the set $\{\frac{\ell}{3} : \ell = 0, \dots, 3|E|\}$, where $|E|$ is the number of edges of the topology, this implies that any best response dynamics starting from an arbitrary initial configuration converges to an equilibrium in $O(|E|)$ steps. Thus, it remains to prove our claim that Φ is an ordinal potential function.

Given an edge $e \in E$ and two assignments $\mathbf{v}$ and $\mathbf{v}'$, let

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

Also, for $z \in V$, let

$$\Delta(\mathbf{v}, \mathbf{v}', z) = \sum_{e: z \in e} \Delta(\mathbf{v}, \mathbf{v}', e).$$

We omit $\mathbf{v}$ and $\mathbf{v}'$ from the notation when they are clear from the context.

In our analysis, we distinguish between jump and swap social games.

Jump games. 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 $\mathbf{v}'$. Observe that i's move only changes the potential of edges incident to v and w. We will now prove that $\Delta(v) + \Delta(w) > 0$. This will prove our claim for the case where v and w are not adjacent, as in this case we have $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = \Delta(v) + \Delta(w)$. Subsequently, 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 the case that $\deg(w) > 0$.
- 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 $\phi(\mathbf{v}, e_w) = \frac{1}{3}$. Since agent i benefits from moving to w, we have $\phi(\mathbf{v}', e_w) = 1$. 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 $\phi(\mathbf{v}, e_{w,1}) = \phi(\mathbf{v}, e_{w,2}) = \frac{1}{3}$. Since agent i benefits from moving to w, we have $\phi(\mathbf{v}', e_{w,1}) + \phi(\mathbf{v}', e_{w,2}) \geq 1$. 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 $\phi(\mathbf{v}, e_v) \leq \frac{1}{3}$. Since v is left empty in $\mathbf{v}'$, we have $\phi(\mathbf{v}', e_v) = \frac{1}{3}$ 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 $\phi(\mathbf{v}', e_{v,1}) = \phi(\mathbf{v}', e_{v,2}) = \frac{1}{3}$. Since agent i benefits from moving away from v, we have $\phi(\mathbf{v}, e_{v,1}) + \phi(\mathbf{v}, e_{v,2}) \leq 1$. Thus, $\Delta(v) \geq -\frac{1}{3}$.

The observations above show 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 non-friend, and $\Delta(w) = \frac{1}{3}$ only if in $\mathbf{v}'$ agent i is adjacent to one friend and one non-friend; but in such a case, agent i would not have an incentive to move, a contradiction. This completes the analysis for the case $\{v, w\} \notin E$.

Now, suppose that v and w are adjacent. In this case $\Delta(v) + \Delta(w)$ double-counts the contribution of the edge $\{v, w\}$ to the potential, so we have $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = \Delta(v) + \Delta(w) - \Delta(\{v, w\})$. However, since w is empty in $\mathbf{v}$ and v is empty in $\mathbf{v}'$, it holds that $\Delta(\{v, w\}) = \frac{1}{3} - \frac{1}{3} = 0$, and hence $\Delta(v) + \Delta(w) > 0$ implies $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) > 0$ in this case as well.

Swap games. In swap games, all nodes of the topology are occupied by agents, and thus $\phi(\mathbf{v}, e)$ takes values in $\{0, 1\}$. Consider an assignment $\mathbf{v}$ and a pair of agents (i, j) with $v_i = v$ and $v_j = w$ who swap their positions, which results in the assignment $\mathbf{v}' = \mathbf{v}^{i \leftrightarrow j}$. The swap of i and j only affects the potential of edges incident to v and w; moreover, if i and j are adjacent in $\mathbf{v}$ then they remain adjacent in $\mathbf{v}'$, and thus the potential of the edge connecting them does not change. Thus, $\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = \Delta(v) + \Delta(w)$. Hence, to prove that Φ is an ordinal potential function, it suffices to show that $\Delta(v) + \Delta(w) > 0$.

We make the following two observations for $z \in \{v, w\}$:

- If $\deg(z) = 1$, let $e \in E$ be the edge that is adjacent to z. Since the agent occupying z in $\mathbf{v}$ has an incentive to leave, we have $\phi(\mathbf{v}, e) = 0$. Since the agent occupying z in $\mathbf{v}'$ has an incentive to move there, we have $\phi(\mathbf{v}', e) = 1$. Hence, $\Delta(e) = 1$ and, consequently, $\Delta(z) = 1$.
- If $\deg(z) = 2$, let e_1 and e_2 be the two edges incident to z. Since the agent occupying z in $\mathbf{v}$ has an incentive to leave, we have $\phi(\mathbf{v}, e_1) + \phi(\mathbf{v}, e_2) \leq 1$. Since the agent occupying z in $\mathbf{v}'$ has an incentive to move there, we have $\phi(\mathbf{v}', e_1) + \phi(\mathbf{v}', e_2) \geq 1$. Consequently, $\Delta(z) \geq 0$.

By the above observations, it follows that $\Delta(v) + \Delta(w) > 0$ unless $\Delta(v) = 0$ and $\Delta(w) = 0$. However, this is impossible. Indeed, if $\Delta(v) = 0$ then i and j are connected to the same number of friends at v, and this number can only be 1: if it is 0, j would not want to move to v, and if it is 2, i would not want to move away. But then it has to be the case that i is connected to two friends at w and j is connected to zero friends at w, so $\Delta(w) = 2$. This completes the proof. □

Since social games generalize k-typed games, all of our non-existence results, hardness results, and lower bounds on the price of anarchy and stability presented in the previous sections apply to social games as well. In fact, for social Schelling games it is easy to prove that maximizing the social welfare is NP-hard even if all agents are strategic. Moreover, this hardness result holds even when the topology is a graph of maximum degree 2; in other words, social welfare maximization may be hard even when finding an equilibrium is easy.

Theorem 7.2. *Given a social Schelling game and a rational value ξ, it is* NP*-complete to decide whether the game admits an assignment with social welfare at least ξ. The hardness result holds even when all agents are strategic and the topology 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 $\{x_{|X|}, x_1\} \in Y$ and for each $i \in [|X| - 1]$ it holds that $\{x_i, x_{i+1}\} \in Y$.

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 $\{p_z : \{z, v\} \in Y\}$.
- The topology $G = (V, E)$ is a cycle consisting of $|X|$ nodes.

By construction, a social welfare of $\xi = |X|$ 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 Sections 5 and 6 only apply to k-typed games with further restrictions on the structure of each type, so they cannot be directly extended to the social setting.

7.2 Schelling Games with Enemy Aversion

In our model, if an agent is not adjacent to any friends, it does not matter how many non-friends she is adjacent to—her utility is always zero. This is also the case in unweighted modified fractional hedonic games: agents are indifferent between being alone and being in coalitions consisting of their non-friends only. This assumption makes sense when the non-friends of an agent are simply agents that do not contribute to her welfare. However, it may be the case that if two agents are not friends, they are enemies, and an agent may prefer being alone to being in a group full of enemies. To take into account such preferences, we can modify the definition of the utility function of a strategic agent so that the agent herself is also included in the set of her friends. In this case, the agent's utility becomes $\frac{f+1}{f+e+1}$, where f is the number of friends and e is the number of enemies that the agent is connected to.

Many of our results extend to this definition of the utility function. For example, for 2-jump games we can construct instances without equilibria, using ideas similar to those in the reduction of Theorem 3.3. Further, for k-jump 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 3.4. As we mentioned in Section 1.2, Kanellopoulos et al. [2021a] recently conducted a more thorough analysis of Schelling games with enemy aversion. However, it still remains an open question if instances with no stubborn agents always admit an equilibrium in this model.

7.3 Schelling Games with Linear Utilities

Throughout the paper we have assumed 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 (i.e., non-friends) 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*; see, e.g., the survey by 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 7.3. *Consider a modification of our 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 given $\alpha, \beta \geq 0$. Then, both for swap games and for jump games, and even if the friendship relation is given by a social network, 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 $\mathcal{N}$. 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.

Fix an assignment $\mathbf{v}$. For each $i \in N$, let

$$\phi_i(\mathbf{v}) = \alpha f_i(\mathbf{v}) - \beta e_i(\mathbf{v}) \qquad \text{and} \qquad \Phi(\mathbf{v}) = \sum_{i \in N} \phi_i(\mathbf{v}).$$

We will argue that Φ is an ordinal potential function for our game. This implies that if the strategic agents follow the best response dynamics starting from any initial configuration, they will converge to an equilibrium. Also, observe that if all agents are strategic, $\Phi(\mathbf{v})$ is equal to the social welfare of $\mathbf{v}$, and thus when $N = R$, the social welfare is maximized at equilibrium. (However, in general this is not the case: intuitively, Φ ascribes "strategic" utilities to the stubborn agents.) Finally, observe that Φ takes values in the set $\{\alpha i - \beta j : 0 \leq i, j \leq n\}$, where n is the number of agents. Thus, any best response dynamics converges in $O(n^2)$ iterations.

We will now argue that Φ is a potential function. We start with jump games. Consider 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 $\mathbf{v}'$. Assume that agent i has f friends and e enemies in $\mathbf{v}$, and f' friends and e' enemies in $\mathbf{v}'$. 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 a neighbor of i in neither $\mathbf{v}$ nor $\mathbf{v}'$, then $\phi_j(\mathbf{v}) = \phi_j(\mathbf{v}')$. Now, suppose that j is connected 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 connected 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

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

For swap games, the proof is similar. Let i and j be a pair of strategic agents who swap their positions $v_i = v$ and $v_j = w$, leading to the assignment $\mathbf{v}' = \mathbf{v}^{i \leftrightarrow j}$. For $z \in \{v, w\}$, we define the following quantities:

- $f_i(z)$ is the number of agents connected to z that are friends of i and enemies of j;
- $e_i(z)$ is the number of agents connected to z that are enemies of i and friends of j;
- $f_{ij}(z)$ is the number of agents connected to z that are friends of both i and j;
- $e_{ij}(z)$ is the number of agents connected to z that are enemies of both i and j.

Since the swap is profitable for both agents, it holds that

$$\phi_i(\mathbf{v}') - \phi_i(\mathbf{v}) = \alpha(f_i(w) + f_{ij}(w) - f_i(v) - f_{ij}(v)) - \beta(e_i(w) + e_{ij}(w) - e_i(v) - e_{ij}(v)) > 0$$

and

$$\phi_j(\mathbf{v}') - \phi_j(\mathbf{v}) = \alpha(f_j(v) + f_{ij}(v) - f_j(w) - f_{ij}(w)) - \beta(e_j(v) + e_{ij}(v) - e_j(w) - e_{ij}(w)) > 0.$$

By summing these two expressions, we can bound $\phi_i(\mathbf{v}') - \phi_i(\mathbf{v}) + \phi_j(\mathbf{v}') - \phi_j(\mathbf{v})$ as

$$\alpha(f_i(w) - f_i(v)) - \beta(e_i(w) - e_i(v)) + \alpha(f_j(v) - f_j(w)) - \beta(e_j(v) - e_j(w)) > 0.$$

Now, observe that there is no change in the utility of the common friends and the common enemies of i and j. Consequently, by considering agents that are friends of i and enemies of j, or vice versa, just as in the jump case above, we can conclude that

$$\begin{aligned}\Phi(\mathbf{v}') - \Phi(\mathbf{v}) = 2\bigg(&\alpha(f_i(w) - f_i(v)) - \beta(e_i(w) - e_i(v))\bigg) \\ &+ 2\bigg(\alpha(f_j(v) - f_j(w)) - \beta(e_j(v) - e_j(w))\bigg) > 0,\end{aligned}$$

as desired. □

8 Conclusions and Open Problems

In this paper, we have extensively studied games inspired by Schelling's seminal segregation model, in which the agents are partitioned into multiple types, occupy nodes of a graph topology, and can increase their utility either by jumping to empty locations or by swapping locations with other agents. We considered questions related to the existence and the efficiency of equilibrium assignments, as well as the price of anarchy and price of stability, both from a social welfare and from a diversity standpoint. We also proposed several variants and extensions for further study.

Concerning equilibrium existence, while we showed that an equilibrium always exists for simple topologies such as stars and graphs of maximum degree 2, we gave an example demonstrating that it may fail to exist even if the topology does not contain cycles. Furthermore, deciding whether a given game admits an equilibrium assignment is NP-complete in general. Even though we have implicitly assumed that the tolerance threshold of every agent is 1, and thus she is never truly happy unless she is connected to friends only, our proofs extend to other threshold values as well. For instance, one can verify that for $k = 2$, Theorem 3.2 holds for any $\tau \in (3/4, 1)$ and Theorem 3.5 holds for any $\tau \in (2/3, 1)$. A challenging open question is to completely characterize the topologies and threshold values for which equilibria are guaranteed to exist, and also to design efficient algorithms to compute equilibria when they exist. Alternatively, one could aim to design parameterized algorithms for deciding the existence of equilibria; for example, can the result of Theorem 3.4 for tree topologies be improved to an FPT algorithm with respect to the number of types k?

For welfare and integration maximization, a natural question is whether one can efficiently compute assignments with nearly optimal social welfare. We note that our NP-hardness reductions in Theorems 4.1 and 4.2 are not approximation preserving, and thus they do not rule out this possibility. Recently, Bullinger et al. [2021] showed that an assignment with at least half of the optimal social welfare can be computed in polynomial time; whether this factor can be improved or a PTAS exists remains open. In addition, it would be interesting to explore the tradeoffs between diversity and social welfare: given two parameters p and q, can we compute an (equilibrium) assignment whose degree of integration is at least p and whose social welfare is at least q? While our results indicate that this problem is hard for general topologies, one could hope to obtain approximate or parameterized algorithms, or focus on simple topologies. One can also investigate more fine-grained diversity indices, for example by considering the number of other types that each agent is exposed to.

Acknowledgments

This work has been supported by the European Research Council (ERC) under grant number 639945 (ACCORD), by the EPSRC International Doctoral Scholars Grant EP/N509711/1, by the KAKENHI Grant-in-Aid for JSPS Fellows number 18J00997, by JST, ACT-X, and by an NUS Start-up Grant. We are grateful to the reviewers of IJCAI 2019, AAAI 2020 and AIJ for their valuable comments.

References

Aishwarya Agarwal, Edith Elkind, Jiarui Gan, and Alexandros A. Voudouris. Swap stability in Schelling games on graphs. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1758–1765, 2020.

Richard D. Alba and John R. Logan. Minority proximity to whites in suburbs: An individual-level analysis of segregation. *American Journal of Sociology*, 98(6):1388–1427, 1993.

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

Haris Aziz and Rahul Savani. Hedonic games. In *Handbook of Computational Social Choice*, pages 356–376. 2016.

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

George Barmpalias, Richard Elwes, and Andrew 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.

George Barmpalias, Richard Elwes, and Andrew Lewis-Pye. Unperturbed Schelling segregation in two or three dimensions. *Journal of Statistical Physics*, 164:1460–1487, 2015.

Stephen Benard and Robb Willer. A wealth and status-based model of residential segregation. *Journal of Mathematical Sociology*, 31(2):149–174, 2007.

Itzhak Benenson, Erez Hatna, and Ehud Or. From Schelling to spatially explicit modeling of urban ethnic and economic residential dynamics. *Sociological Methods and Research*, 37(4):463–497, 2009.

Prateek Bhakta, Sarah Miracle, and Dana Randall. Clustering and mixing times for segregation models on $\mathbb{Z}^2$. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 327–340, 2014.

Davide Bilò, Vittorio Bilò, Pascal Lenzner, and Louise Molitor. Topological influence and locality in swap Schelling games. In *Proceedings of the 45th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 15:1–15:15, 2020.

Niclas Boehmer and Edith Elkind. Individual-based stability in hedonic diversity games. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1822–1829, 2020.

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

Christina Brandt, Nicole Immorlica, Gautam Kamath, and Robert Kleinberg. An analysis of one-dimensional Schelling segregation. In *Proceedings of the 44th Symposium on Theory of Computing Conference (STOC)*, pages 789–804, 2012.

Robert Bredereck, Edith Elkind, and Ayumi Igarashi. Hedonic diversity games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 565–573, 2019.

Martin Bullinger, Warut Suksompong, and Alexandros A. Voudouris. Welfare guarantees in Schelling segregation. *Journal of Artificial Intelligence Research*, 71:143–174, 2021.

Hau Chan, Mohammad T. Irfan, and Cuong Viet Than. Schelling models with localized social influence: a game-theoretic framework. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 240–248, 2020.

Ankit Chauhan, Pascal Lenzner, and Louise Molitor. Schelling segregation with strategic agents. In *Proceedings of the 11th International Symposium on Algorithmic Game Theory (SAGT)*, pages 137–149, 2018.

William Clark and Mark Fossett. Understanding the social context of the Schelling segregation model. *Proceedings of the National Academy of Sciences*, 105(11):4109–4114, 2008.

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

David A. Easley and Jon M. Kleinberg. *Networks, Crowds, and Markets – Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

Hagen Echzell, Tobias Friedrich, Pascal Lenzner, Louise Molitor, Marcus Pappik, Friedrich Schöne, Fabian Sommer, and David Stangl. Convergence and hardness of strategic Schelling segregation. In *Proceedings of the 15th International Conference on Web and Internet Economics (WINE)*, pages 156–170, 2019.

Edith Elkind, Jiarui Gan, Ayumi Igarashi, Warut Suksompong, and Alexandros A. Voudouris. Schelling games on graphs. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 266–272, 2019.

Edith Elkind, Angelo Fanelli, and Michele Flammini. Price of Pareto optimality in hedonic games. *Artificial Intelligence*, 288:103357, 2020a.

Edith Elkind, Neel Patel, Alan Tsang, and Yair Zick. Keeping your friends close: Land allocation with friends. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 318–324, 2020b.

Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, New York, 1979.

Michael R. Garey, David S. Johnson, and Larry Stockmeyer. Some simplified NP-complete problems. In *Proceedings of the 6th annual ACM Symposium on Theory of Computing (STOC)*, pages 47–63, 1974.

Nicole Immorlica, Robert Kleinberg, Brendan Lucier, and Morteza Zadimoghaddam. 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.

Panagiotis Kanellopoulos, Maria Kyropoulou, and Alexandros A. Voudouris. Modified Schelling games. *Theoretical Computer Science*, 880:1–19, 2021a.

Panagiotis Kanellopoulos, Maria Kyropoulou, and Alexandros A. Voudouris. Not all strangers are the same: The impact of tolerance in schelling games. *CoRR*, abs/2105.02699, 2021b.

Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 404–413, 1999.

Luca Kreisel, Niclas Boehmer, Vincent Froese, and Rolf Niedermeier. Equilibria in Schelling games: Computational complexity and robustness. arXiv 2105.06561, 2021.

Stanley Lieberson and Donna K. Carter. A model for inferring the voluntary and involuntary causes of residential segregation. *Demography*, 19(4):511–526, 1982.

Sagar Massand and Sunil Simon. Graphical one-sided markets. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 492–498, 2019.

Douglas S. Massey and Nancy A. Denton. The dimensions of residential segregation. *Social Forces*, 67(2): 281–315, 1988.

Gianpiero Monaco, Luca Moscardelli, and Yllka Velaj. Stable outcomes in modified fractional hedonic games. *Autonomous Agents and Multi-Agent Systems*, 34(1):4:1–4:29, 2019.

Martin Olsen. On defining and computing communities. In *Proceedings of the 18th Computing: Australasian Theory Symposium (CATS)*, pages 97–102, 2012.

Romans Pancs and Nicolaas J. Vriend. Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics*, 91(1–2):1–24, 2007.

Thomas C. Schelling. Models of segregation. *American Economic Review*, 59(2):488–493, 1969.

Thomas C. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2):143–186, 1971.

H. Peyton Young. *Individual Strategy and Social Structure: an Evolutionary Theory of Institutions.* Princeton University Press, 2001.

Junfu Zhang. A dynamic model of residential segregation. *Journal of Mathematical Sociology*, 28(3):147–170, 2004a.

Junfu Zhang. Residential segregation in an all-integrationist world. *Journal of Economic Behavior and Organization*, 54(4):533–550, 2004b.

A Proof of Theorem 3.4

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 $\mathcal{U} = \{i/j : i \in [n], j \in [n], i \le j\} \cup \{0\}$; note that $|\mathcal{U}| \le n^2$.

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

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

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* of *false*. Specifically, $\tau_v(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \textit{true}$ 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 $C = \textit{empty}$, then node v is empty. Otherwise, the node 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 the 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. Note that for conditions 4 and 5, if there is no agent with the specified property, the corresponding condition is vacuously true. Similarly, for conditions 6–8, if there is no empty node to deviate to, the corresponding condition vacuously holds.

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 $n_B = |T_B|$, $n_R = |T_R|$, $\tau_r(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) =$ *true* for the root node r of G, and, moreover,

- if $C =$ *blue*, then
$$\frac{k_B}{k_B + k_R} \geq \hat{u}_{top};$$
- if $C =$ *red*, then
$$\frac{k_R}{k_B + k_R} \geq \hat{u}_{top};$$
- if $C =$ *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 polynomial time.

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

$$T_v(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \begin{cases} \textit{true}, & \text{if } \mathbf{n} = (\mathbb{1}_B(C), \mathbb{1}_R(C)) \text{ and } \mathbf{k} = (0,0) \\ \textit{false}, & \text{otherwise.} \end{cases} \tag{4}$$

Suppose now that for a node w we have constructed the table τ_v for each $v \in \textit{child}(w)$. We will construct τ_w using these tables as follows. Let $\textit{child}(w) = \{v_1, \dots, v_L\}$. 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 $\textit{tree}_\ell(w)$ obtained from $\textit{tree}(w)$ 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 (4). 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: $\theta_w^\ell(C, \mathbf{n}, \mathbf{k}, \check{\mathbf{u}}, \hat{\mathbf{u}}) = \textit{true}$ if and only if there exists 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 $\theta_w^{\ell-1}(C', \mathbf{n}', \mathbf{k}', \check{\mathbf{u}}', \hat{\mathbf{u}}') = \tau_{v_\ell}(C'', \mathbf{n}'', \mathbf{k}'', \check{\mathbf{u}}'', \hat{\mathbf{u}}'') = \textit{true}$ 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 $C'' = \textit{blue}$, then
$$\frac{k''_B + \mathbb{1}_B(C')}{k''_B + k''_R + (1 - \mathbb{1}_E(C'))} \geq \check{u}_{B^\dagger}$$
and, if $C'' = \textit{red}$, 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 $\textit{tree}_\ell(w) \setminus (\textit{child}(w) \cup \{w\})$ have utility at least $\check{u}_X$.

6. For each $X \in \{B, R\}$,
$$\hat{u}'_X \leq \hat{u}_X, \quad \hat{u}''_X \leq \hat{u}_X$$
and, if $C'' = \textit{empty}$, then
$$\frac{k''_X + \mathbb{1}_X(C')}{k''_B + k''_R + (1 - \mathbb{1}_E(C'))} \leq \hat{u}_X.$$

This sets upper bounds for the utilities that agents that do not occupy nodes of $\mathit{tree}_\ell(w)$ can obtain by deviating to a node in the first $\ell-1$ branches, a node other than v_ℓ in the ℓ-th branch, and node v_ℓ, respectively.

7. If $C' = \mathit{blue}$, then
$$\hat{u}'_{\mathit{top}} \le \hat{u}_{\mathit{top}}, \quad \hat{u}''_B \le \hat{u}_{\mathit{top}}$$
and, if also $C'' = \mathit{empty}$, then
$$\frac{k''_B}{k''_B + k''_R} \le \hat{u}_{\mathit{top}},$$
so that the blue agent occupying node w has utility at most $\hat{u}_{\mathit{top}}$ 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 $C' = \mathit{red}$, then
$$\hat{u}'_{\mathit{top}} \le \hat{u}_{\mathit{top}}, \quad \hat{u}''_R \le \hat{u}_{\mathit{top}}$$
and, if also $C'' = \mathit{empty}$,
$$\frac{k''_R}{k''_B + k''_R} \le \hat{u}_{\mathit{top}}.$$

8. If $n''_B > 0$, then $\check{u}''_B, \check{u}''_{B\dagger} \ge \hat{u}'_B$ 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. If also $C'' = \mathit{empty}$ and $n''_B > k''_B$, then
$$\check{u}''_B \ge \frac{k''_B + \mathbb{1}_B(C')}{k''_B + k''_R + \mathbb{1}_B(C')},$$
while if also $C'' = \mathit{empty}$ and $k''_B > 0$ then
$$\check{u}''_{B\dagger} \ge \frac{k''_B + \mathbb{1}_B(C') - 1}{k''_B + k''_R + \mathbb{1}_B(C') - 1},$$
so that blue agents occupying nodes other than v_ℓ in the ℓ-th branch have no incentive to deviate to v_ℓ. Since $\tau_{v_\ell}(C'', \mathbf{n}'', \mathbf{k}'', \check{\mathbf{u}}'', \hat{\mathbf{u}}'') = \mathit{true}$ 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 $\mathit{tree}_\ell(w) \setminus \{w\}$. Further, if $C'' = \mathit{blue}$, then
$$\frac{k''_B + \mathbb{1}_B(C')}{k''_B + k''_R + \mathbb{1}_B(C')} \ge \hat{u}'_B, \hat{u}''_{\mathit{top}},$$
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. If $n'_B > k'_B + \mathbb{1}_B(C')$, then $\check{u}'_B \ge \hat{u}''_B$ so that blue agents occupying nodes in the first $\ell-1$ branches other than $v_1, \ldots, v_{\ell-1}$ have no incentive to deviate to nodes in the ℓ-th branch (excluding node v_ℓ). If also $C'' = \mathit{empty}$, then
$$\check{u}'_B \ge \frac{k''_B + \mathbb{1}_B(C')}{k''_B + k''_R},$$
so that these blue agents have no incentive to deviate to v_ℓ if it is empty. Likewise, if $k'_B > 0$, then analogous constraints must hold with $\check{u}'_{B\dagger}$ taking the role of $\check{u}'_B$, so that any blue agent in $v_1, \ldots, v_{\ell-1}$ has no incentive to deviate to nodes in the ℓ-th branch. 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. $\square$