Swap Stability in Schelling Games on Graphs

Aishwarya Agarwal, Edith Elkind, Jiarui Gan, Alexandros A. Voudouris

Department of Computer Science, University of Oxford

aishwarya.agarwal@keble.ox.ac.uk, {edith.elkind, jiarui.gan, alexandros.voudouris}@cs.ox.ac.uk

Abstract

We study a recently introduced class of strategic games that is motivated by and generalizes Schelling's well-known residential segregation model. These games are played on undirected graphs, with the set of agents partitioned into multiple types; each agent either occupies a node of the graph and never moves away or aims to maximize the fraction of her neighbors who are of her own type. We consider a variant of this model that we call swap Schelling games, where the number of agents is equal to the number of nodes of the graph, and agents may *swap* positions with other agents to increase their utility. We study the existence, computational complexity and quality of equilibrium assignments in these games, both from a social welfare perspective and from a diversity perspective.

1   Introduction

Segregation is observed in many communities; people tend to group together on the basis of politics, religion, or socioeconomic status. This phenomenon has been extensively documented in residential metropolitan areas, where people may select where to live based on the racial composition of the neighborhoods. To formalize and study how the motives of individuals lead to residential segregation, Thomas Schelling (1969; 1971) proposed the following simple, yet 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 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 hopes of improving their current situation. Schelling described a dynamic process where at each step unhappy agents jump to random open positions or swap positions with other randomly selected agents, and showed that it can lead to a completely segregated placement, even if the agents themselves are tolerant of mixed neighborhoods ($\tau < 1/2$).

Over the years, Schelling's work became very popular among researchers in Sociology and Economics, who proposed and studied numerous variants of his model, mainly via agent-based simulations; see the paper of Clark and Fossett (2008) and references therein for examples of this ap-

proach. Variants of the model have also been rigorously analyzed in a series of papers (Young 2001; Zhang 2004a; Brandt et al. 2012; Barmpalias, Elwes, and Lewis-Pye 2014; Bhakta, Miracle, and Randall 2014; Barmpalias, Elwes, and Lewis-Pye 2015; Immorlica et al. 2017), which showed that the random behavior of the agents leads with high probability to the formation of large monochromatic regions.

While all these papers focused on settings where the agents' behavior is random, it is more realistic to assume instead that the agents are *strategic* and move only when they have an opportunity to improve their situation. So far, only a few papers have followed such a game-theoretic approach. In particular, Zhang (2004b) considered a game where the agents optimize a single-peaked utility function, and very recently, Chauhan, Lenzner, and Molitor (2018), Elkind et al. (2019) and Echzell et al. (2019) studied strategic settings that are closer to the original model of Schelling, but capture more than two agent types and richer graph topologies.

In particular, Chauhan, Lenzner, and Molitor (2018) study a 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. identify 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. Echzell et al. (2019) strengthen these results and extend them to more than two agent types, as well as study the complexity of computing assignments that maximize the number of happy agents.

Elkind et al. (2019) consider a similar model with k types; however, they treat agents' location preferences differently from Chauhan et al. Namely, in their model each agent is either stubborn (i.e., has a preferred location and is unwilling to move) or strategic (i.e., aims to maximize the fraction of her neighbors that are of her own type; this corresponds to setting $\tau = 1$ in the model of Chauhan et. al.). They focus on jump games, i.e., games where agents may jump to empty positions, and analyze the existence and complexity

of computing Nash equilibria, as well as prove bounds on the price of anarchy (Koutsoupias and Papadimitriou 1999) and the price of stability (Anshelevich et al. 2008).

Our Contribution

We combine the two approaches by considering swap games in the model of Elkind et al. (2019). That is, we assume that the number of agents is equal to the number of nodes in the topology, and two agents can swap locations if each of them prefers the other agent's location to her own. We begin by studying the existence of equilibrium assignments. While such assignments exist for highly structured topologies, we show that they may fail to exist in general, even for simple topologies such as trees. Moreover, we show that deciding whether an equilibrium exists is NP-complete. We also study the quality of assignments in terms of their social welfare: we prove bounds on the price of anarchy and the price of stability for many interesting cases, and show that computing an assignment with high social welfare is NP-complete; the latter result complements the result of Elkind et al. in that it applies to the case where the number of agents equals the number of nodes in the topology.

Given that the goal of Schelling's work was to study integration, it is natural to ask what level of integration can be achieved at equilibrium. There is a number of integration indices that have been proposed for this purpose (see, e.g., the survey of Massey and Denton (1988)). However, many of the indices defined in the literature are formulated for settings where the topology is highly regular and there are only two agent types, and it is not immediately clear how to adapt them to our model. We therefore focus on a simple index, which we call the *degree of integration*, that 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. We then study the price of anarchy and the price of stability with respect to this index: that is, we compare the value of our index in the best and worst equilibrium of our game to the optimal value of this index that can be achieved for a given instance. We note that, to the best of our knowledge, this is the first result of this type in the context of Schelling games: the previous work on integration in the Schelling model typically focused on evaluating a given integration index after some number of steps of the underlying dynamical process, and did not ask what level of integration can be achieved if the agents were non-strategic. We obtain strong negative results: it turns out that even the best equilibria are often much less diverse than the maximally diverse assignments. however, when the topology is a line, the price of stability with respect to our index can be bounded by a small constant. We also show that maximizing diversity is computationally hard.

Further Related Work

As mentioned above, Schelling's model has been studied extensively both empirically and theoretically. For an introduction to the model and a survey of its many variants, we refer the reader to the book of Easley and Kleinberg

(2010), and the papers by Brandt et al. (2012) and Immorlica et al. (2017). Besides the closely related papers by Chauhan, Lenzner, and Molitor (2018), Elkind et al. (2019) and Echzell et al. (2019), another work that is similar in spirit is a recent paper by Massand and Simon (2019), who study swap stability in games where a set of items is to be allocated among agents who are connected via a social network, so that each agent gets one item, and her utility depends on the items she and her neighbors in the network get; however, their results are not directly applicable to our setting. Also, Schelling games share a number of properties with hedonic games (Drèze and Greenberg 1980; Bogomolnaia and Jackson 2002), and in particular, with fractional hedonic games (Aziz et al. 2019) and hedonic diversity games (Bredereck, Elkind, and Igarashi 2019). However, a fundamental difference between hedonic games and Schelling games is that in the former agents form pairwise disjoint coalitions, while in the latter the neighborhoods of different nodes of the topology may overlap.

2 Preliminaries

A *k-swap game* is given by a set N of $n \geq 2$ *agents* partitioned into $k \geq 2$ pairwise disjoint *types* $T_1, \ldots, T_k$, and an undirected simple connected graph $G = (V, E)$ with $|V| = n$, called the *topology*. We often identify types with colors: e.g., in a 2-swap game each agent is either red (T_1) or blue (T_2). 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 $R \cup S = N$. Stubborn agents never move away from the nodes they occupy, while a strategic agent aims to maximize her personal utility, and is willing to *swap* positions with other agents to achieve this.

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 $F_i = T_\ell \setminus \{i\}$. Each agent i occupies some node $v_i \in V$ of the topology G so that $v_i \neq v_j$ for every pair of agents $i \neq j$. The vector $\mathbf{v} = (v_1, \ldots, v_n)$ that lists the locations of all agents is called an *assignment*. Given an assignment $\mathbf{v}$, we denote by $\pi_v(\mathbf{v})$ the agent that occupies node $v \in V$, that is, $\pi_{v_i}(\mathbf{v}) = i$.

Given an assignment $\mathbf{v}$, let $N_i(\mathbf{v}) = \{j \neq i : \{v_i, v_j\} \in E\}$ be the set of neighbors of agent i. The utility u_i of a stubborn agent $i \in S$ is independent of the assignment; e.g., we can set $u_i(\mathbf{v}) = 0$ for each $i \in S$. The utility of a strategic agent $i \in R$ for assignment $\mathbf{v}$ is

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

Observe that, since $|V| = n$, every node is occupied by some agent. As G is a connected graph, this implies that $N_i(\mathbf{v}) \neq \varnothing$ for every $i \in N$.

For every assignment $\mathbf{v}$, let $\mathbf{v}^{i \leftrightarrow j}$ be the assignment that is obtained from $\mathbf{v}$ by swapping the positions of 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 if and 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 an *equilibrium* if no pair of agents i, j want to swap positions. That is, $\mathbf{v}$ is an equilibrium if and only if for every $i, j \in R$ we have $u_i(\mathbf{v}) \geq u_i(\mathbf{v}^{i \leftrightarrow j})$ or $u_j(\mathbf{v}) \geq u_j(\mathbf{v}^{i \leftrightarrow j})$. We denote the set of all equilibrium assignments of the k-swap game $\mathcal{G}$ by $\mathrm{EQ}(\mathcal{G})$.

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

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

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

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

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

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

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

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

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 language for the price of stability.

Due to space constraints, some proofs have been omitted, and can be found in the full version (Agarwal et al. 2019).

3 Existence of Equilibria

We begin by discussing the existence of equilibria in swap games. The work of Echzell et al. (2019) implies that when the topology is a regular graph, at least one equilibrium assignment is guaranteed to exist. Furthermore, using a potential function similar to the one proposed by Elkind et al. (2019) for jump games, we can show that equilibria always exist when the topology is a graph of maximum degree 2; we omit the details due to space constraints.

Our first result is a proof of non-existence of equilibria for every $k \geq 2$ for general topologies.

Theorem 1. *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.*
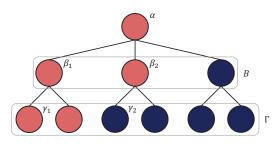

Figure 1: The topology of the 2-swap game considered in the proof of Theorem 1, and an assignment that corresponds to the last case in the analysis; it is not an equilibrium since the red agent at node α and the blue agent at node γ_2 would like to swap.

Proof. We prove the theorem only for the case $k = 2$; the proof for $k \geq 3$ can be found in full version. 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 1. 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 1). 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. $\square$

The topology used in the proof of Theorem 1 is utilized as a subgraph in the proof of the following theorem, to show that the problem of deciding whether an equilibrium exists is computationally hard.

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

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, we give a reduction from the CLIQUE problem, which in known to be NP-complete. An instance of CLIQUE consists of an undirected connected graph $H = (X, Y)$ and an integer λ; it is a yes-instance if H contains a complete subgraph of size λ. Without loss of generality, we assume that $\lambda > 5$.

Given an instance $\langle H, \lambda \rangle$ of CLIQUE with $H = (X, Y)$, we will construct a 2-swap game as follows (the reduction can be extended to any $k > 2$ by adding stubborn 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 λ 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 are defined as follows:
  - To define $G_1 = (V_1, E_1)$, let W_v be a set of $2d_H - d_v + 2\lambda - 3$ nodes for each $v \in X$. Then, $V_1 = X \bigcup_{v \in V} W_v$ 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 $d_H - d_v + 2\lambda - 3$ nodes are occupied by stubborn blue agents. Observe that every node of G_1 not occupied by a stubborn agent has degree $d_1 = 2d_H + 2\lambda - 3$.
  - The subgraph $G_2 = (A \cup B, E_2)$ is a complete bipartite graph with $|A| = \lambda - 5$ and $|B| = 4d_1$. Out of the $4d_1$ nodes of B, $2d_1 + 1$ nodes are occupied by stubborn red agents, while the remaining $2d_1 - 1$ nodes are occupied by stubborn blue agents.
  Hence, a strategic red agent occupying a node of A has utility $\chi_r = \frac{2d_1 + 1}{4d_1} = \frac{1}{2} + \frac{1}{4d_1}$. Similarly, a strategic blue agent has utility $\chi_b = \frac{2d_1 - 1}{4d_1} = \frac{1}{2} - \frac{1}{4d_1}$.
  - To define $G_3 = (V_3, E_3)$, let (V_3', E_3') be the graph used in the proof of Theorem 1, for which there is no equilibrium assignment; see Figure 1. For every node $v \in V_3'$ of degree 3, let Z_v be a set of $100d_1$ nodes such that $50d_1$ of these nodes are occupied by stubborn red agents, while the remaining $50d_1$ nodes are occupied by stubborn blue agents. For every $v \in V_3'$ of degree 1, let Z_v be a set of $10d_1$ nodes such that $5d_1$ of these nodes are occupied by stubborn red agents, while the remaining $5d_1$ nodes are occupied by stubborn blue agents. Then, $V_3 = V_3' \bigcup_{v \in V_3'} Z_v$ and $E_3 = E_3' \cup \{\{v, w\} : v \in V_3', w \in Z_v\}$.
  One can easily verify that the utility of a strategic agent (red or blue) occupying a node of G_3 is at least $\psi_0 = \frac{5d_1 - 1}{10d_1 + 1} > \frac{1}{2} - \frac{1}{4d_1}$ and at most $\psi_1 = \frac{5d_1 + 1}{10d_1 + 1} < \frac{1}{2} + \frac{1}{4d_1}$.

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

$$u = \frac{\lambda - 1 + d_H}{d_1} = \frac{d_H + \lambda - 1.5 + 0.5}{2d_H + 2\lambda - 3} \geq \frac{1}{2} + \frac{1}{2d_1}.$$

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. By swapping positions with a blue agent within G_1, a strategic red agent would still have at most $\lambda - 1 + d_H$ friends, and since every node in G_1 has the same

degree, her utility cannot be improved. Hence, no strategic red agent has a profitable deviation, and $\mathbf{v}$ is an equilibrium.

Conversely, assume that H does not contain a clique of size λ, 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 $\lambda - 1$ other strategic red agents, as this would mean that the nodes they occupy form a clique of size λ. Hence, at least one of these agents, say agent i, is adjacent to at most $\lambda - 2$ strategic red agents. Since every node of G_1 has degree d_1 and every node is adjacent to d_H stubborn red agents, the utility of i is

$$u_i \leq \frac{d_H + \lambda - 2}{d_1} = \frac{d_H + \lambda - 1.5 - 0.5}{2d_H + 2\lambda - 3} = \frac{1}{2} - \frac{1}{2d_1}.$$

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 can obtain by swapping positions with i is

$$u_j = 1 - u_i \geq \frac{1}{2} + \frac{1}{2d_1}.$$

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 $\chi_b < \psi_0$, which implies 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. Consequently, there are 5 strategic red agents and 5 strategic blue agents in G_3. However, similarly to the proof of Theorem 1, we can argue that there is no equilibrium assignment for these agents in G_3; we omit the details here. Since we have exhausted all possibilities, it follows that if H does not have a clique of size λ, then there is no equilibrium assignment. $\square$

4 Social Welfare

We will now consider the efficiency of equilibrium assignments in terms of social welfare, and bound the social price of anarchy and stability for many interesting cases. We restrict our attention to games with strategic agents only and with at least two agents per type. Swap games with stubborn agents or types that consist of a single strategic agent can be easily seen to have unbounded social price of anarchy.[1]

We start with the social price of anarchy of 2-swap games, and consider the general case (given the above restrictions) and the case where each type consists of the same number of agents.

---

[1]For any $k \geq 2$, consider a k-swap game with star topology and k types of agents such that there is a unique red agent r, who is strategic, while all other types contain at least two strategic agents as well as some stubborn agents located at peripheral nodes. The assignment where r occupies the center node is an equilibrium with 0 social welfare, while every assignment where r is not in the center has positive social welfare.
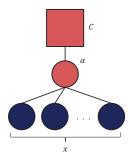
Figure 2: The topology and the equilibrium assignment of the lower bound instance in the proof of part (i) of Theorem 3. The big red square represents a clique whose nodes are occupied by red agents only.

Theorem 3. *The social price of anarchy of 2-swap games with strategic agents only is*

- $\Theta(n)$ *if there are at least two agents of each type, and*
- *between* $921/448 \approx 2.05864$ *and* 4 *if each type consists of the same number of agents.*

Proof. Due to space constraints, we only prove the second claim. For the lower bound, consider a 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 $x - 1$. There is an equilibrium $\mathbf{v}$ where α is occupied by a red agent r, all leaf nodes are occupied by blue agents, and all nodes of C are occupied by red agents; see Figure 2. Hence, $\mathrm{SW}(\mathbf{v}) = x - 1 + \frac{1}{x+1} = \frac{x^2}{x+1}$. On the other hand, 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 $\frac{2x^3 - x^2 - 5x + 2}{x^2(x-1)}$, an expression that takes it maximum value $667/324 \approx 2.05864$ at $x = 9$.

For the upper bound, consider a 2-swap game with $n = 2x$ agents such that there are $x \geq 2$ red and x blue agents. First, assume that some agents get zero utility in the equilibrium assignment $\mathbf{v}$. Observe that it cannot be the case that there exist agents of both types who have zero utility in $\mathbf{v}$. Indeed, if this was true for a non-adjacent red-blue pair, then these agents would have an incentive to swap and increase their utility from zero to 1. On the other hand, suppose this is true for an adjacent red-blue pair (r, b). If both of r and b have other neighbors (besides r and b), then by swapping they can increase their utility from 0 to positive. Hence, suppose that r occupies a leaf node and is connected only to b. In this case, since the graph is connected, there exists a blue agent $b' \neq b$ whose utility is strictly less than 1; then r and b' have an incentive to swap, as r's utility would become positive and b''s utility would become 1.

Thus, assume that at least one blue agent has zero utility and all red agents have positive utility. We denote by B_0 the set of blue agents with zero utility, by R_1 the set of red agents with utility 1, and by $R_<$ the set of red agents whose utility is strictly less than 1. We have $|R_1| + |R_<| = x$, and each agent in B_0 is connected to all agents in $R_<$; otherwise

a non-adjacent pair of agents $i \in R_<$, $j \in B_0$ have an incentive to swap. Since the optimal social welfare is at most $n = 2x$, to show that the price of anarchy is at most 4, it suffices to establish that the sum of red agents' utilities is at least $x/2$. We consider the following subcases.

$|R_<| = 1$. In this case $\mathrm{SW}(\mathbf{v}) \geq |R_1| = x - 1$; since $x \geq 2$, we have $x - 1 \geq x/2$.

$|R_<| \geq 2, |B_0| \geq 2$. In this case any agent in B_0 would like to swap with any agent in $R_<$ to get positive utility (since each agent in $R_<$ is connected to all agents in B_0). Thus, in $\mathbf{v}$ no agent $i \in R_<$ wants to swap with any agent in B_0. Since each agent in B_0 is connected to all agents in $R_<$ and no other agent, the utility that agent i would get by agreeing to swap is $u_i(\mathbf{v}) = (|R_<| - 1)/|R_<| \geq 1/2$, yielding $\mathrm{SW}(\mathbf{v}) \geq |R_1| + |R_>|/2 \geq x/2$.

$|R_<| \geq 2, |B_0| = 1$. Again, we will show that $u_i(\mathbf{v}) \geq 1/2$ for each $i \in R_<$. Indeed, if i is connected to a blue agent not in B_0, then the blue agent in B_0 would like to swap with i; as i does not want to swap, it follows that her utility is at least $1/2$. Otherwise, i is only connected to red agents and the unique agent in B_0, so again $u_i(\mathbf{v}) \geq 1/2$.

Next, we assume that all agents have positive utility. Since $\mathbf{v}$ is an equilibrium, for every red-blue pair of agents it holds that at least one of them has no incentive to swap. Let (i, j) be a red-blue pair of agents, and assume that i does not want to swap. If i and j are *not* neighbors in $\mathbf{v}$, then it must be that $u_i(\mathbf{v}) \geq 1 - u_j(\mathbf{v})$ and hence $u_i(\mathbf{v}) + u_j(\mathbf{v}) \geq 1$. Otherwise, i and j are neighbors in $\mathbf{v}$. If $u_i(\mathbf{v}) + u_j(\mathbf{v}) < 1$, we have $u_i(\mathbf{v}) \in (0, 1)$, $u_j(\mathbf{v}) \in (0, 1)$. Assume that the blue agent j has $x_r \geq 0$ red neighbors besides i, and $x_b \geq 1$ blue neighbors. Then, $u_j(\mathbf{v}) = \frac{x_b}{x_r + x_b + 1}$, and

$$u_i(\mathbf{v}) \geq \frac{x_r}{x_r + x_b + 1} = 1 - u_j(\mathbf{v}) - \frac{1}{x_r + x_b + 1}$$
$$\geq \frac{1}{2} - u_j(\mathbf{v}) \qquad (1)$$

and hence $u_i(\mathbf{v}) + u_j(\mathbf{v}) \geq \frac{1}{2}$, where inequality (1) follows since $x_b \geq 1$. Therefore, we have $u_i(\mathbf{v}) + u_j(\mathbf{v}) \geq \frac{1}{2}$ for every red-blue pair of agents i and j. Since there are x^2 distinct red-blue pairs, and each agent participates in exactly x such pairs, by summing over all these inequalities, we obtain $x \cdot \mathrm{SW}(\mathbf{v}) \geq \frac{1}{2} x^2$ and therefore $\mathrm{SW}(\mathbf{v}) \geq \frac{1}{2} x$. The bound then follows since the maximum possible social welfare is $n = 2x$. □

We remark that even though the upper bound of 4 for the case where each type contains the same number of agents is probably not tight, one cannot expect to improve it using the same technique, as (1) can be seen to be tight.

Next, we show that, surprisingly, if there are at least three types, the social price of anarchy can be unbounded, even for the special case of equal number of agents per type.

Theorem 4. *For every* $k \geq 3$*, the social price of anarchy of k-swap games can be unbounded, even when there is an equal number of strategic agents per type.*
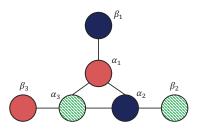
Figure 3: The topology and the equilibrium assignment of the k-swap game considered in the proof of Theorem 4 for $k = 3$. Here, $T_1 =$ red, $T_2 =$ blue and $T_3 =$ green.

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 $T_1, \dots, T_k$. The topology G consists of k nodes $\{\alpha_1, \dots, \alpha_k\}$ that form a cycle, and each node α_ℓ, $\ell \in [k]$, is also connected to an auxiliary node β_ℓ; see Figure 3 for the topology and the equilibrium discussed in the following for $k = 3$.

Let $\mathbf{v}$ be the assignment in which node α_ℓ is occupied by an agent of type T_ℓ, while node β_ℓ is occupied by an agent of type $T_{\ell+1}$, where the subscripts are specified modulo ℓ. This assignment is clearly an equilibrium since there exists no pair of agents that would like to swap positions, and every agent has zero utility. On the other hand, consider the assignment $\mathbf{v}^*$ in which nodes α_ℓ and β_ℓ are occupied by the two agents of type T_ℓ, for every $\ell \in [k]$. Since every agent now has positive utility, we have $\text{SW}(\mathbf{v}^*) > 0$, and hence the social price of anarchy is unbounded. □

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

Theorem 5. *The social price of stability in 2-swap games is at least* $4/3$.

For k-swap games with topology that is a δ-regular graph (in which all nodes have degree δ), we show an upper bound of 1 on the social price of stability. To this end we employ a potential function that is similar to the one defined by Chauhan, Lenzner, and Molitor (2018) and Echzell et al. (2019) to show the existence of equilibria in such games.

Theorem 6. *The social price of stability in k-swap games with topology that is a δ-regular graph is* 1.

Proof. Echzell et al. (2019) showed that for k-swap games with a δ-regular topology, $\Phi(\mathbf{v}) = \sum_{i \in R} |N_i(\mathbf{v}) \setminus F_i|$ is a potential minimization function. Using similar arguments, we can show that the complement, $\overline{\Phi}(\mathbf{v}) = \sum_{i \in R} |N_i(\mathbf{v}) \cap F_i|$, is a potential maximization function. Now, observe that by the definition of the utility of each strategic agent and the fact that the topology is δ-regular, for every assignment $\mathbf{v}$ we have

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

Let $\mathbf{v}^*$ be an optimal assignment. If $\mathbf{v}^*$ is an equilibrium, then the social price of stability is 1. Otherwise, we let the
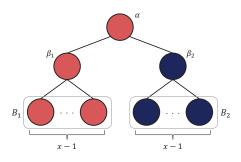

Figure 4: The topology and the only possible equilibrium assignment for the 2-swap game considered in the proofs of Theorems 8 and 9. For k-swap games (Theorem 8) there are k identical subtrees, and in the worst equilibrium each subtree is filled by agents of a different type.

strategic agents swap positions until they reach an equilibrium $\mathbf{v}$. Since $\overline{\Phi}$ is a potential maximization function, we have $\overline{\Phi}(\mathbf{v}) \geq \overline{\Phi}(\mathbf{v}^*)$. Thus, by (2) we obtain

$$\text{SW}(\mathbf{v}) = \frac{1}{\delta} \cdot \overline{\Phi}(\mathbf{v}) \geq \frac{1}{\delta} \cdot \overline{\Phi}(\mathbf{v}^*) = \text{SW}(\mathbf{v}^*),$$

and the bound follows by rearranging terms. □

Next, we focus on the problem of computing assignments with high social welfare. Observe that the complexity of this problem does not depend of the set of deviations available to the agents, i.e., on whether we consider swap games or jump games, with one exception: in swap games, we assume that no node is empty, and in jump games we assume that at least one node is empty. Thus, the hardness result for social welfare maximization in jump Schelling games established by Elkind et al. (2019) does not cover our case. The proof of our next theorem is very different from that of Elkind et al. (2019), and only covers the case $k \geq 3$; for $k = 2$, we were unable to prove the hardness of the problem.

Theorem 7. *For every $k \geq 3$, given a rational number ξ, it is NP-complete to decide whether there exists an allocation that has social welfare at least ξ.*

5 Degree of Integration

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 measure counts the number of agents who are exposed, i.e., have at least one neighbor of a different type. As in the previous section, we again focus on games with strategic agents only.

We start by showing that the integration price of anarchy of k-swap games is n/k, i.e., it scales linearly with the number of agents. This indicates that in the worst case agents of different types are highly segregated, but, as the number of types increases, equilibria become more diverse and the price of anarchy decreases.

Theorem 8. *For any $k \geq 2$, the integration price of anarchy of k-swap games with strategic agents only is at most n/k, and this bound is tight.*

Proof. For the upper bound, consider a k-swap game with n agents. By definition, the optimal degree of integration is at most n. Since the topology is a connected graph, 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 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 4 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 $\beta_\ell, \ell \in [k] \setminus \{1\}$, are exposed, yielding degree of integration $\text{DI}(\mathbf{v}) = k$, and the bound follows. $\square$

Next, we consider the integration price of stability. Using the same instance as in the proof of Theorem 8, 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 the previous theorem, and indicates that we cannot avoid ending up with equilibrium assignments in which the types are highly segregated, even in the best case.

Theorem 9. *The integration price of stability of 2-swap games with strategic agents only is at least $n/2$.*

To develop better intuition for the integration price of anarchy and stability, we also consider the special case where the topology is a line. In this case, while the integration price of anarchy remains linear in n/k, the integration price of stability can be bounded by a small constant.

Theorem 10. *Consider a k-swap game with strategic agents only, at least two agents per type, and a line topology. The integration price of anarchy is at most $\frac{n}{2k-2}$, while the integration price of stability is at most $\frac{9}{4}$. Moreover, if the number of agents of each type grows with n, the integration price of stability is at most $\frac{3}{2} + o(1)$. All these bounds are tight.*

Proof. We prove the upper bound for the price of stability here; see the full version of the paper for the remaining proofs. Let the topology be a line, with nodes $1, \ldots, n$ connected in this order. We partition the agents of each type into blocks of size 2 and 3, with at most one block of size 3 per type (this block is created if the number of agents of that type is odd). Observe that any assignment where the agents of each block are placed contiguously on the line is an equilibrium. Indeed, under any such assignment each agent has at least one neighbor of the same type, and no agent can move to a position where she would have two neighbors of her type; also, the agents at nodes 1 or n are unwilling to swap (as they get utility 1).

It remains to explain how to place these blocks on the line to maximize integration. We do this greedily, from left to right. That is, we first pick some $i \in [k]$ such that $|T_i| \geq$

$|T_j|$ for all $j \in [k]$, and place some block $B \subseteq T_i$ first. Now, suppose that ℓ blocks have been placed, so that the last occupied node is node r, and the agent there is of type T_x. For each $j \in [k]$, let t_j be the number of agents of type T_j who have not yet been placed. If $t_j = 0$ for all $j \in [k] \setminus \{x\}$, we complete the assignment by simply placing all the remaining agents of type T_x on the line. Otherwise, we pick an $i \in [k] \setminus \{x\}$ such that $t_i \geq t_j$ for all $j \in [k] \setminus \{x\}$, and place some block $B \subseteq T_i$ in positions $r + 1, \ldots, r + |B|$.

Let us say that a type T_i is *dominant* if $|T_i| > n/2$. An easy inductive argument shows that if no type is dominant, then under this assignment we never place two blocks of the same type next to each other; the key observation is that if no type is dominant after ℓ blocks have been placed, this remains to be the case after $\ell + 2$ blocks have been placed, and hence if we still have at least two blocks to place, we have at least two types to choose from. In this case, the only agents who are not exposed are agents at nodes 1 and n as well as agents located at the middle of a block of size 3, i.e., at most $k + 2$ agents. Thus, the integration price of stability is at most $\frac{n}{n-k-2}$ in this case. Now, suppose that some type (say, type T_1) is dominant. If there are λ blocks of types $T_2, \ldots, T_k$, then under our procedure we will first alternate between blocks of type T_1 and blocks of other types, and then place the remaining blocks of type T_1 (if any). Then at least 4λ agents will be exposed. On the other hand, at most $k - 1$ of these λ blocks are of size 3, so we have at most $2\lambda + k - 1$ agents of types $T_2, \ldots, T_k$, and hence in any assignment at most $2(2\lambda + k - 1)$ agents of type T_1 can be exposed. Thus, the integration price of stability in this case is at most $\frac{3(2\lambda+k-1)}{4\lambda} = \frac{3}{2} + \frac{3k-3}{4\lambda}$. Since $\lambda \geq k - 1$, this quantity is at most $\frac{9}{4}$. Further, if we assume that the number of agents of each type grows with n, we have $\frac{3k-3}{4\lambda} = o(1)$, and the bound becomes $\frac{3}{2} + o(1)$. $\square$

Hence, for games with simple line topologies, integration can be achieved in equilibrium. However, when left alone, the agents may end up in a very segregated configuration.

We conclude by studying the complexity of computing assignments (not necessarily equilibria) with high degree of integration. Unfortunately, it turns out that even this task is computationally intractable.

Theorem 11. *Given a k-swap game, computing an assignment in which every agent is exposed is NP-complete, even if $k = 2$ and all agents are strategic.*

6 Conclusions and Open Problems

We have studied Schelling games on graphs in which pairs of agents can deviate by swapping their locations. We considered questions related to the existence and the efficiency of equilibrium assignments, both from a social welfare perspective and from a diversity perspective.

While equilibria are known to exist in instances where the topology is highly structured, we showed that their existence is not guaranteed in general, and deciding whether a given swap game admits an equilibrium assignment is NP-complete. 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 Theorem 1 for $k = 2$ holds for any $\tau \in (2/3, 1)$. A challenging open question is to completely characterize the topologies and the threshold values for which equilibria are guaranteed to exist, and also design efficient algorithms to compute equilibria when they exist.

We have introduced a new index for measuring the diversity of a given assignment, which we called the degree of integration. It would be interesting to explore the tradeoffs between diversity and social welfare: can we compute (equilibrium) assignments with a given degree of integration that maximize the social welfare? 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 ambitious diversity indices, e.g., by considering, for each agent, the number of other types she is exposed to.

Acknowledgments

References

Agarwal, A.; Elkind, E.; Gan, J.; and Voudouris, A. A. 2019. Swap stability in Schelling games on graphs. *CoRR* abs/1909.02421.

Anshelevich, E.; Dasgupta, A.; Kleinberg, J. M.; Tardos, É.; Wexler, T.; and Roughgarden, T. 2008. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing* 38(4):1602–1623.

Aziz, H.; Brandl, F.; Brandt, F.; Harrenstein, P.; Olsen, M.; and Peters, D. 2019. Fractional hedonic games. *ACM Transactions on Economics and Computation* 7(2):6:1–6:29.

Barmpalias, G.; Elwes, R.; and Lewis-Pye, A. 2014. Digital morphogenesis via Schelling segregation. In *Proceedings of the 55th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 156–165.

Barmpalias, G.; Elwes, R.; and Lewis-Pye, A. 2015. From randomness to order: unperturbed Schelling segregation in two or three dimensions. *CoRR* abs/1504.03809.

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

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

Brandt, C.; Immorlica, N.; Kamath, G.; and Kleinberg, R. 2012. An analysis of one-dimensional Schelling segregation. In *Proceedings of the 44th Symposium on Theory of Computing Conference (STOC)*, 789–804.

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

Chauhan, A.; Lenzner, P.; and Molitor, L. 2018. Schelling segregation with strategic agents. In *Proceedings of the 11th International Symposium on Algorithmic Game Theory (SAGT)*, 137–149.

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

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

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

Echzell, H.; Friedrich, T.; Lenzner, P.; Molitor, L.; Pappik, M.; Schöne, F.; Sommer, F.; and Stangl, D. 2019. Convergence and hardness of strategic Schelling segregation. *CoRR* abs/1907.07513.

Elkind, E.; Gan, J.; Igarashi, A.; Suksompong, W.; and Voudouris, A. A. 2019. Schelling games on graphs. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 266–272.

Immorlica, N.; Kleinberg, R.; Lucier, B.; and Zadimoghaddam, M. 2017. Exponential segregation in a two-dimensional Schelling model with tolerant individuals. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 984–993.

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

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

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

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

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

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

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

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

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