

Network Creation with Homophilic Agents

Martin Bullinger¹, Pascal Lenzner² and Anna Melnichenko²

¹ Technical University of Munich

² Hasso Plattner Institute, University of Potsdam

bullinge@in.tum.de, {pascal.lenzner, anna.melnichenko}@hpi.de

Abstract

Network Creation Games are an important framework for understanding the formation of real-world networks. These games usually assume a set of indistinguishable agents strategically buying edges at a uniform price leading to a network among them. However, in real life, agents are heterogeneous and their relationships often display a bias towards similar agents, say of the same ethnic group. This homophilic behavior on the agent level can then lead to the emergent global phenomenon of social segregation. We study Network Creation Games with multiple types of homophilic agents and non-uniform edge cost, introducing two models focusing on the perception of same-type and different-type neighboring agents, respectively. Despite their different initial conditions, both our theoretical and experimental analysis show that both the composition and segregation strength of the resulting stable networks are almost identical, indicating a robust structure of social networks under homophily.

1 Introduction

Networks play an eminent role in today’s world. They are crucial for our energy supply (power grid networks), our information exchange (the Internet and the World Wide Web), and our social relationships (friendship networks, email exchange, or co-author networks). There exists an abundance of approaches to provide formal frameworks for modeling networks, see, for example, the books by Jackson [2010] and Newman [2018]. In many of these models, the nodes of the network correspond to agents that strategically create connections which is particularly suitable for our main focus of modeling social networks. One such stream of research considers variants of the *Network Creation Game (NCG)* as proposed by Fabrikant *et al.* [2003]. There, selfish agents create edges to form a network among themselves. Forming edges is costly and hence agents try to create only the most useful edges. On the other hand, the force that causes agents to form edges at all is well-connectivity within the network, captured by a desire to occupy a central position.

The NCG is a stylized model of social interaction, providing valuable insight to agents’ decision processes when in-

teracting with each other. However, it is important to refine the basic model to spotlight specific details of this decision making. In this sense, we study network creation under the additional assumption that agents are separated into various *types* that model ethnic groups or affiliations.

Our goal is to contribute a new perspective on the simple causes that lead to the segregation of a society, similar to Schelling’s checkerboard model for residential segregation [Schelling, 1969; Schelling, 1971]. Therefore, our agents’ cost functions have a bias towards the creation of relationships with agents of the same type. Specifically, we study two models based on two seemingly orthogonal treatments of other agents. In the first model, agents incur a fixed cost for every created edge and a variable cost that only depends on the number of edges towards same-type agents. In the second model, edges towards different-type agents are initially more expensive but their cost drops with an inverse linear decay. Both models give a different point-of-view on the same underlying principle, namely homophily of agents, i.e., the tendency to form connections with like-minded people. This is often summarized with the proverb “birds of a feather flock together”, a dominant intrinsic force repeatedly observed in the creation of social networks, see McPherson *et al.* [2001] for a survey on the extensive sociological research on homophily in social networks. While our first model expresses homophily explicitly by an increasing comfort among friends, the second model incorporates homophily indirectly by accounting for a decreasing effort of integration once first contact is established. The latter paradigm is closely related to the well-known effect in social sciences called the “contact hypothesis” which states that stereotypes and prejudices between ethnic groups can be weakened by intensified contact [Allport *et al.*, 1954; Amir, 1969; Dovidio *et al.*, 2003].

We measure the desirability of networks by means of stability. Since we consider social networks, we assume a bilateral model where two agents have to cooperate to connect. Consequently, we use pairwise stability [Jackson and Wolinsky, 1996] as solution concept, rather than, for instance, Nash stability which is more appropriate for unilateral models.

Interestingly, we find an almost identical structure of stable networks for both models. This hints at a robust structure of networks created under homophily incentives. Naturally, a very small edge cost causes extremely high connectivity.

For moderately small edge cost, we provide characterizations of stable networks which are all highly segregated. We interpret this as identifying a sweet spot of high sensitivity towards agent types. For larger edge cost, stability causes a large spectrum of networks to form with respect to segregation strength. We accompany this theoretical limitation with an average-case analysis by detailed simulations of a simple distributed dynamics, where agents perform improvements towards stable networks. It would be plausible if a generally high edge cost causes less distinction of agent types. While this is sometimes confirmed, we also identify contrasting tendencies towards extreme segregation. An important driver for the different behavior is the initial segregation level, indicating that segregation can be avoided by a high initial effort without constant further interaction.

2 Related Work

In the original NCG the cost of every edge is α , where α is a parameter of the game that allows adjusting the tradeoff between the agents' cost for creating edges and their cost for the centrality in the network, e.g., the sum of distances to all other nodes. Stable networks always exist, in particular, for $\alpha < 1$, only cliques are stable, whereas for $1 \leq \alpha < n$ stars, other trees and also non-tree networks can be stable [Mamagishvili *et al.*, 2015]. Bilateral NCGs with uniform edge price have been introduced by [Corbo and Parkes, 2005]. Also variants of the NCG with non-uniform edge cost have been studied: a version where edges of differing quality can be bought [Cord-Landwehr *et al.*, 2014], and NCGs where the edge cost depends on the node degrees [Chauhan *et al.*, 2017], on the length of the edges in a geometric setting [Bilò *et al.*, 2019], or on the hop-distance of the endpoints [Bilò *et al.*, 2021]. The latter is motivated by social networks, and bilateral edge formation with pairwise stability as a solution concept is considered. The NCG variant by Meirum *et al.* [2014] features different types of agents and different but fixed edge costs for each agent type.

Closest to our work is the model proposed by Martí and Zenou [2017] that is a variant of the connections model [Jackson and Wolinsky, 1996] with different types of agents. Similar to our model, the cost for maintaining an other-type connection depends on the homogeneity of the neighborhoods of the involved agents. In contrast to us, the cost for same-type edges is fixed and the distance cost is defined differently. The authors study the existence and structure of equilibria but do not investigate segregation quantitatively. The latter has been done by Henry *et al.* [2011] using a stochastic process that starts with a randomly drawn network with nodes of different types. Then edges are randomly rewired with a built-in bias toward favoring same-type edges. As the main result, the authors show that the network strongly segregates over time, even if the built-in bias is very low.

Residential segregation has recently received a lot of attention by a stream of research developing a game-theoretic framework based on Schelling's checkerboard model [Chauhan *et al.*, 2018; Agarwal *et al.*, 2021; Echzell *et al.*, 2019; Bilò *et al.*, 2020; Kanellopoulos *et al.*, 2021; Bullinger *et al.*, 2021]. There, agents of several types strate-

gically select positions on a given *fixed* network and they individually aim for having at least a τ -fraction of same-type neighbors, for some $0 < \tau \leq 1$. Also hedonic diversity games [Bredereck *et al.*, 2019; Boehmer and Elkind, 2020] are similar. These are coalition formation games where an agent's utility depends on the type distribution of her coalition.

3 Preliminaries and Model

We consider a set $V = \{1, \dots, n\}$ of n agents partitioned into $k \geq 2$ disjoint *types*. The set of types is denoted by \mathcal{T} , and for every type $T \in \mathcal{T}$, let V_T be the set of agents of type T , i.e., $V = \bigcup_{T \in \mathcal{T}} V_T$ and $V_T \cap V_{T'} = \emptyset$ for $T, T' \in \mathcal{T}$, with $T \neq T'$. For an agent $u \in V$, we denote by $\mathcal{T}(u)$ her type, i.e., $u \in V_{\mathcal{T}(u)}$. Given a type $T \in \mathcal{T}$, then $n_T = |V_T|$ denotes the number of agents of type T . We identify types with colors and we assume that there are specific types B and R of *blue* and *red* agents, respectively, which are associated with an agent type having the smallest and largest number of agents, respectively. Thus, for every type $T \in \mathcal{T}$, we have $n_B \leq n_T \leq n_R$. In particular, with exactly two agent types we have precisely a blue minority and a red majority type.

In a network creation game, agents will buy edges to eventually form a network, which is an undirected graph $G = (V, E)$. Therefore, it is useful to introduce some common concepts and notation from graph theory. Consider an undirected graph $G = (V, E)$ together with vertices $u, v \in V$. We denote the (potential) edge between u and v by uv (whether it is present or not). For two agents $u, v \in V$, the edge uv is called *monochromatic* if u and v are of the same type, and *bichromatic*, otherwise. If $uv \in E$, we use the notation $G - uv = (V, E \setminus \{uv\})$, otherwise we use $G + uv = (V, E \cup \{uv\})$. Further, let $N_G(u) = \{v \in V : uv \in E\}$ denote the *neighborhood* of u in G , let $\deg_G(u) = |N_G(u)|$ be the *degree* of u in G , i.e., the size of its neighborhood, and let $d_G(u, v)$ be the *distance* from u to v in G , i.e., the length of a shortest path from u to v in G . The *diameter* of G is defined as $\text{diam}(G) = \max_{u, v \in V} d_G(u, v)$, i.e., the maximum length of any shortest path in G . Finally, a useful measure for the centrality of a vertex in a network is its distance to a set of vertices. Given a subset $V' \subseteq V$ of vertices, let $d_G(u, V') = \sum_{v \in V'} d_G(u, v)$ denote the sum of distances from u to all vertices in V' . Also, given a subset of agents $C \subseteq V$, we denote by $G[C]$ the subgraph of G *induced* by C , i.e., $G[C] = (C, F)$, where $F = \{uv \in E : u, v, \in C\}$.

Before formally defining our network creation model, we introduce some relevant special types of graphs. The graph $\mathbf{K}_n = (V, E)$ is called *complete* if $E = \{uv : u, v, \in V\}$, i.e., all possible edges are present. Further, $\mathbf{S}_n = (V, E)$ is called *star* if there exists $u \in V$ such that $E = \{uv : v \in V \setminus \{u\}\}$. We also define networks for the special case of two types. Given two agents $u \in V_B$ and $v \in V_R$, the network $\mathbf{DS}_n = (V, E)$ is called *double star* if $E = uv \cup \{uw : w \in V_B\} \cup \{vw : w \in V_R\}$ and $\mathbf{DSX}_n = (V, E)$ is called *double star with switched centers* if $E = uv \cup \{uw : w \in V_R\} \cup \{vw : w \in V_B\}$. An undirected graph G is called *complete*, *star*, *double star*, or *double star with exchanged centers* if it is isomorphic to \mathbf{K}_n , \mathbf{S}_n , \mathbf{DS}_n , or \mathbf{DSX}_n , respectively (where

isomorphisms have to preserve agent types).

Network Creation Games with Homophilic Agents. We study network creation within a cost-oriented bilateral model à la Corbo and Parkes [2005], where the agent cost is separated into a neighborhood cost encompassing the cost of sponsoring edges and a distance cost encompassing the cost of the agents' centrality. In both of our models, a created network G has a *distance cost* for agent u of $d_G(u) := d_G(u, V)$, i.e., the sum of agent u 's distances to all other agents. The neighborhood cost is different in our two models and will be specified in the definition of our network creation games.

To model the cost dependency on the types of neighbors, we define the set of same-type agents in the neighborhood of agent u as $F_G(u) = V_T \cap N_G(u)$, if $u \in V_T$. The set of other-type neighbors is defined as $E_G(u) = N_G(u) \setminus F_G(u)$. We denote the cardinalities of these sets by $f_G(u) = |F_G(u)|$ and $e_G(u) = |E_G(u)|$, respectively.

Now we define our network creation games. A *network creation game with increasing comfort among friends* (ICF-NCG) with cost parameter $\alpha > 0$ is a network creation game where the neighborhood cost is given by

$$a_G^{ICF}(u) = \deg_G(u) \cdot \alpha \left(1 + \frac{1}{f_G(u) + 1} \right),$$

i.e., there is a fixed cost of α for every edge and an additional cost that decreases with an increasing number of friends.

A *network creation game with decreasing effort of integration* (DEI-NCG) with cost parameter $\alpha > 0$ is a network creation game where the neighborhood cost is given by

$$a_G^{DEI}(u) = \alpha \left(\deg_G(u) + \sum_{k=1}^{e_G(u)} \frac{1}{k} \right).$$

Hence, there is a fixed edge cost of α for every edge to an agent in the neighborhood together with a harmonically decreasing additional cost for edges towards other-type agents. Note that the sum is empty for $e_G(u) = 0$, and therefore, the game is identical to the single-type bilateral network creation game by Corbo and Parkes [2005] if $k = 1$.

For the neighborhood cost, we omit the superscript indicating the type of network creation game, whenever this is clear from the context. Also, for both of our models, we define the cost of an agent u as $c_G(u) = a_G(u) + d_G(u)$.

The cost functions mimic the two effects that we want to model, namely a general homophilic behavior via the ICF-NCG and diminishing prejudices with intensified contact via the DEI-NCG. In both models, edge costs have a similar decay structure and identical range of $[\alpha, 2\alpha]$. In the ICF-NCG, the cost of edges is 2α for each edge if an agent has no friends, and the edge cost is approaching α when the number of neighboring friends is growing. In the DEI-NCG, the cost of edges to friends is always α and the variable cost only affects other-type agents, where we approach α with a harmonic decay starting at a cost of 2α for the first other-type agent.

Measures for Desirable Networks. We analyze networks by the incentives of agents to maintain the network in terms of stability and by the diversity of their neighborhood with

respect to other agent types. Following Jackson and Wolinsky [1996], a network $G = (V, E)$ is called *pairwise stable* if the following two properties hold:

- (i) for all agents $u \in V$ and neighbors $v \in N_G(u)$, it holds that $c_G(u) \leq c_{G-uv}(u)$, i.e., no agent can benefit from unilaterally severing an edge, and
- (ii) for all agents $u \in V$ and non-neighbors $v \notin N_G(u)$, it holds that $c_G(u) \leq c_{G+uv}(u)$ or $c_G(v) \leq c_{G+uv}(v)$, i.e., no pair of agents can bilaterally create an edge such that the individual cost for both agents decreases.

Connectivity is an important aspect in network analysis. With multiple agent types, the internal connectivity per type deserves special consideration. Formally, a network $G = (V, E)$ is called *fully intra-connected* if, for every pair $u, v \in V$ of same-type agents, it holds that $uv \in E$. Further, G is *fully connected* if G is complete.

For the evaluation of diversity, we consider two segregation measures. Given a network $G = (V, E)$, its *local segregation*, denoted by $LS(G)$, is defined as the average fraction of agents of the same type, i.e., $LS(G) = \frac{1}{|V|} \sum_{u \in V} \frac{f_G(u)}{\deg_G(u)}$. The *global segregation*, called $GS(G)$, is the proportion of monochromatic edges, i.e., $GS(G) = \frac{\sum_{u \in V} f_G(u)}{2|E|}$. Note that $\frac{1}{2} \sum_{u \in V} f_G(u)$ is the number of monochromatic edges.¹

Finally, the minimum willingness to integrate of an agent can be evaluated by checking if she entertains any bichromatic edge. Therefore, we call an agent *curious* if she is part of a bichromatic edge. Similarly, a type of agents is called *curious* if it solely consists of curious agents. Note that this concept is related to the degree of integration, which is identical to the number of curious agents and has been studied in game-theoretic segregation models [Agarwal *et al.*, 2021].

4 Increasing Comfort among Friends

In this section we perform our theoretical analysis of the ICF-NCG. Unless explicitly stated otherwise, all statements hold for an arbitrary number of types. All missing proofs can be found in the full version of the paper [Bullinger *et al.*, 2022].

We start by gathering some statements concerning structural properties and simple pairwise stable networks. Their proof follows by a careful analysis of the cost difference after the creation and deletion of edges.

Proposition 4.1. *For the ICF-NCG the following hold:*

1. If $\alpha < \frac{6}{7}$, then every pairwise stable network is fully intra-connected.
2. If $\alpha < \frac{4}{3}$, then $\text{diam}(G) \leq 2$ for every pairwise stable network G . In particular, G contains a curious type.
3. Let $\alpha < 1$, G a pairwise stable network, and $C \subseteq V$ such that every agent in C is curious and $C \subseteq V_T$ for some type $T \in \mathcal{T}$. Then, $G[C]$ is a clique. In particular, every curious type of agents is fully intra-connected.
4. If $\alpha \leq \frac{n_B}{n_B+1}$, then the complete network \mathbf{K}_n is pairwise stable. Moreover for $\alpha < \min\{\frac{6}{7}, \frac{n_B}{n_B+1}\}$, \mathbf{K}_n is the

¹ LS and GS are (related to) standard measures in social sciences to capture the agents' *exposure* [Massey and Denton, 1988]. LS is used by [Paolillo and Lorenz, 2018] and GS is used in the simulation framework Netlogo [Wilensky, 1997] and by [Zhang, 2011].

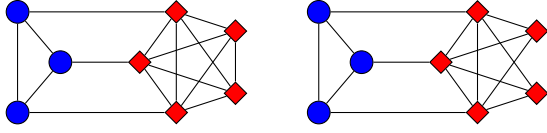


Figure 1: Pairwise stable networks for $\frac{n_B}{n_B+1} \leq \alpha < \tau$ (left) and $\tau \leq \alpha < 1$ (right).

unique pairwise stable network.

5. If $\alpha \geq 1$, then the star \mathbf{S}_n is pairwise stable.

The uniqueness in Proposition 4.1(4) excludes the parameter range $\frac{6}{7} \leq \alpha \leq \frac{n_B}{n_B+1}$, which can only happen for sufficiently many blue agents. In fact, there the uniqueness ceases to hold, as we show in the full version [Bullinger *et al.*, 2022].

For the existence of stable networks, we still have to consider the intermediate parameter range $\frac{n_B}{n_B+1} < \alpha < 1$. We provide the construction for two agent types. The general case is covered in the full version [Bullinger *et al.*, 2022].

Proposition 4.2. *In the ICF-NCG, there exists a pairwise stable network for every $\frac{n_B}{n_B+1} \leq \alpha < 1$.*

Construction for two agent types. Consider an instance of the ICF-NCG and let $\frac{n_B}{n_B+1} \leq \alpha < 1$. We will define a stable network for α dependent on the threshold $\tau = \frac{n_B(n_B+1)}{n_B(n_B+1)+1}$. Note that $\frac{n_B}{n_B+1} < \tau < 1$, as $n_B(n_B+1) > n_B$.

We assume $V_B = \{b_1, \dots, b_{n_B}\}$ and $V_R = \{r_1, \dots, r_{n_R}\}$ and define the edge set of the graph $G = (V, E)$ as follows:

- $\{x_i, x_j\} \in E$, for $x \in \{b, r\}$, $i, j \in \{1, \dots, n_B\}$,
- $\{r_i, b_i\} \in E$, for $i \in \{1, \dots, n_B\}$,
- $\{r_i, r_j\} \in E$, for $i \in \{1, \dots, n_B\}$ and $j \in \{n_B+1, \dots, n_R\}$,
- if $\alpha < \tau$, then $\{r_i, r_j\} \in E$, for $i, j \in \{n_B+1, \dots, n_R\}$, and no further edges are in E ;
otherwise, no further edges are in E .

The two cases for the network G are illustrated in Figure 1. They can be shown to be pairwise stable for their respective parameter range. \square

Interestingly, the stable networks constructed in the previous proof give an almost full characterization of stable networks for the considered range of edge costs when $k = 2$.

Theorem 4.3. *Consider the ICF-NCG with parameter α and $k = 2$ agent types. Let $\frac{n_R}{n_R+1} < \alpha < 1$ and assume that G is pairwise stable. Then, the blue agents are fully intra-connected, the bichromatic edges form a matching of size n_B , and curious red agents are connected to all other red agents.*

Proof sketch. Let $\frac{n_R}{n_R+1} < \alpha < 1$ and assume that G is pairwise stable network in the ICF-NCG with cost parameter α . By Proposition 4.1(2), the diameter of G is bounded by 2 and there exists a curious type of agents. By Proposition 4.1(3), the curious type of agents forms a clique C and the curious agents of the other type form a clique as well.

Now, it can be shown that the bichromatic edges form a matching by proving that any agent incident to two such edges can sever one of them. Therefore, only a minority type can be a curious type and we can conclude that the blue agents

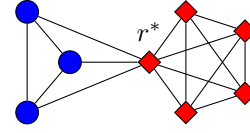


Figure 2: Pairwise stable network for $\frac{n_B}{n_B+1} \leq \alpha \leq \frac{n_R}{n_R+1}$.

are fully intra-connected and that the matching of bichromatic edges is of size n_B . Showing that curious red agents maintain edges with non-curious red agents follows from another application of Proposition 4.1(2). \square

Example 4.4. The characterization encountered in Theorem 4.3 does not cover the whole range of Proposition 4.2. In fact, it does not hold for $\frac{n_B}{n_B+1} \leq \alpha \leq \frac{n_R}{n_R+1}$. Hence, further pairwise stable networks exist. Assume that $n_R \geq 2$ and let $r^* \in V_R$. Consider the network $G = (V, E)$, where $E = \{\{v, w\} : v, w \in V_R\} \cup \{\{v, w\} : v, w \in V_B\} \cup \{\{v, r^*\} : v \in V_B\}$, i.e., the network is fully intra-connected and there is a special agent r^* to which all blue agents are connected. The structure of this network is depicted in Figure 2. It is straightforward to check that the network is pairwise stable. \triangleleft

Until now, we set our focus on the existence of pairwise stable networks. In the remainder of the section, we want to consider the segregation of pairwise stable networks. First, Theorem 4.3 yields very high segregation for $\frac{n_R}{n_R+1} < \alpha < 1$.

Corollary 4.5. *Consider the ICF-NCG with parameter α and $k = 2$ agent types. Let $\frac{n_R}{n_R+1} < \alpha < 1$ and assume that G is pairwise stable. Then, $GS(G) \geq 1 - \frac{1}{n}$ and $LS(G) \geq 1 - \frac{2}{n}$.*

We know that segregation is low for sufficiently low parameter α , where cliques are (uniquely) pairwise stable. Then, there is a transition at $\alpha = \frac{n_R}{n_R+1}$, where segregation is provably high regardless of further parameters like the distribution of agents into types. Once, the cost parameter increases to $\alpha \geq 1$, the picture becomes less clear. Stars can have very high and very low segregation.

Proposition 4.6. *Consider the ICF-NCG with parameter $\alpha \geq 1$. Then, for every $n \geq 2$, there exist pairwise stable networks G and G' on n nodes such that $GS(G) = LS(G) = 1$ and $GS(G') = LS(G') = \frac{1}{n-1}$.*

The networks in the previous proposition have the drawback that we need to fix the exact numbers of agents of each type to obtain the desired segregation. By contrast, for $\alpha \geq \frac{4}{3}$, the double star is always highly segregated.

Proposition 4.7. *Consider the ICF-NCG with $\alpha \geq \frac{4}{3}$. Then, the double star \mathbf{DS}_n is a pairwise stable network with $GS(\mathbf{DS}_n) = 1 - \frac{1}{n-1}$ and $LS(\mathbf{DS}_n) \geq 1 - \frac{2}{n}$.*

5 Decreasing Effort of Integration

We consider the DEI-NCG. See the full version [Bullinger *et al.*, 2022] for all omitted details. We start by collecting some results determining simple stable networks for sufficiently small and large values of α , respectively. Note that we implicitly assume the restriction to two agent types when considering the networks \mathbf{DS}_n and \mathbf{DSX}_n . All other statements hold for an arbitrary number of agent types.

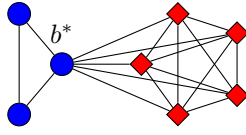


Figure 3: Pairwise stable network for $\frac{1}{2} \leq \alpha \leq \frac{n_R}{n_R+1}$.

Proposition 5.1. *For the DEI-NCG the following holds:*

1. If $\alpha < \frac{1}{2}$, then \mathbf{K}_n is the unique pairwise stable network.
2. If $\alpha < 1$, then every pairwise stable network is fully intra-connected.
3. If $\alpha < 1$, then every pairwise stable network G satisfies $\text{diam}(G) \leq 2$.
4. The network \mathbf{K}_n is pairwise stable if $\alpha \leq \frac{n-n_R}{n-n_R+1}$.
5. If $\alpha \geq 1$, then \mathbf{S}_n and \mathbf{DS}_n are pairwise stable networks.
6. If $\alpha \geq \frac{4}{3}$, then \mathbf{DSX}_n is a pairwise stable network.

Proposition 5.1(2) and Proposition 5.1(3) imply that, for $\alpha < 1$, every pairwise stable network consists of two monochromatic cliques and one type of agents is curious. Still, there are highly segregated pairwise stable networks. Also, note that the highly integrated clique investigated in Proposition 5.1(4) is not the unique stable network for $\alpha \geq \frac{1}{2}$, as the next example shows for the case $k = 2$.

Example 5.2. Assume $k = 2$ and $\frac{1}{2} \leq \alpha \leq \frac{n_R}{n_R+1}$. Recall that n_R is the size of the majority type of agents. In particular, this covers the case $\alpha \leq \frac{n_B}{n_B+1} = \frac{n-n_R}{n-n_R+1}$. Assume that $n_B \geq 2$ and let b^* be some fixed blue agent, i.e., an agent from the minority type. Consider the network $G = (V, E)$ with $E = \{vw : v, w \in R\} \cup \{vw : v, w \in B\} \cup \{vb^* : v \in R\}$, i.e., the network is fully intra-connected and there is a special blue agent b^* to which all red agents are connected. There are no further bichromatic edges. For an illustration of the network, see Figure 3. Pairwise stability of this network follows from straightforward computations. \triangleleft

In the previous example, it was still possible to simultaneously have full intra-connectivity while there are agents entertaining several bichromatic edges. This is not possible anymore if we further increase α . In fact, we can even characterize all pairwise stable networks for $\frac{n_R}{n_R+1} < \alpha < 1$ and $k = 2$.

Theorem 5.3. *Let $k = 2$ in the DEI-NCG. Assume that $\frac{n_R}{n_R+1} < \alpha < 1$ and consider a network G . Then, G is pairwise stable if and only if it is fully intra-connected and its bichromatic edges form a matching covering V_B .*

Proof sketch. Clearly, if $k = 2$ and $n_R = 1$, then the unique stable network consists of a neighboring blue and red agent and the assertion is true. Thus, we may assume that $n_R \geq 2$. We show that the given conditions imply pairwise stability.

Therefore, assume that G is a fully intra-connected network such that the bichromatic edges form a matching covering one type of agents. Then, no edge can be severed because monochromatic edges only decrease the neighborhood cost by $\alpha < 1$ while increasing the distance cost by 1. Also, bichromatic edges decrease the neighborhood cost by $2\alpha < 2$

while increasing the distance cost by 2. Finally, it is impossible to create another bichromatic edge. This edge would be the second bichromatic edge incident to its endpoint from the minority type of agents. This agent would only decrease her distance cost by 1 while increasing her neighborhood cost by $\frac{3}{2}\alpha \geq \frac{3}{2} \frac{n_R}{n_R+1} \geq 1$, where we use $n_R \geq 2$ in the last step. \square

The second part of the above proof shows that the networks characterized in the theorem are even stable for $\frac{2}{3} \leq \alpha < 1$. Putting together Proposition 5.1, Example 5.2, and Theorem 5.3, we have proved the existence of pairwise stable networks for almost every DEI-NCG if $k = 2$ (except a limit case when $n_B = 1$). By generalizing the encountered networks, we can show their existence for an arbitrary number of types. The generalization of the network in Example 5.2 is straightforward, maintaining the property that there exists one specific agent entertaining all bichromatic edges. However, the generalization of the network in Theorem 5.3 is a bit disguised. We define the network by providing an efficient algorithm. This algorithm initially considers a fully intra-connected network and adds edges by having agents create bichromatic edges via specific better responses. In the special case of $k = 2$, this results precisely in the matchings encountered in Theorem 5.3 (see the full version for details).

Theorem 5.4. *In the DEI-NCG pairwise stable networks always exist.*

Finally, we want to consider the segregation of pairwise stable networks in the DEI-NCG. Clearly, segregation only depends on the networks, not on the type of NCG. Hence, we transfer from the investigation of ICF-NCGs that cliques provide low segregation for small α , stars provide high or low segregation for high α , but require a specific distribution of agents into types. Independently of this distribution, double stars provide high segregation and it is clear that $GS(\mathbf{DSX}_n) = LS(\mathbf{DSX}_n) = 0$. Finally, for an intermediate range of α , high segregation is inevitable.

Corollary 5.5. *Let $k = 2$ and $\frac{n_R}{n_R+1} < \alpha < 1$. Then, every pairwise stable network G in the DEI-NCG with parameter α satisfies $GS(G) \geq 1 - \frac{2}{n}$ and $LS(G) \geq 1 - \frac{2}{n}$.*

6 Experimental Analysis

While our theoretical results indicate a clear structure of stable networks for $\alpha \leq 1$, there is a broad range of possibilities for larger α . Therefore, we support the theoretical findings for $\alpha > 1$ by a detailed experimental analysis². To this end, we simulate a simple dynamic process based on distributed and strategic edge creation and deletion over time, incentivized by optimizing the cost functions of our two models.

The dynamics start with sparse initial networks (spanning tree or grid) and distribute agents of two equally-sized types such that the segregation of the initial network is very low or high. In each step, one agent is activated uniformly at random and can either create or delete an edge, performing a best response with respect to the cost function under consideration. In particular, we study also an *add-only* variant of the model,

²The source code for the experiments can be found at <https://github.com/melnan/HomophilicNCG>

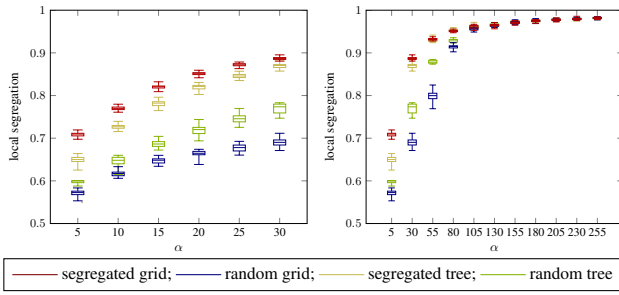


Figure 4: Local segregation of 1.01-approximate networks in the DEI-NGC obtained by iterative best response moves for $n = 1000$ over 50 runs starting from a random or segregated tree and grid.

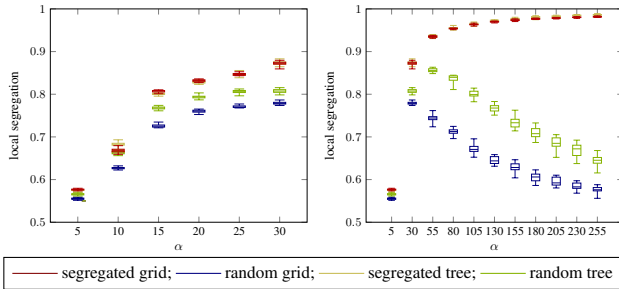


Figure 5: Local segregation of pairwise stable networks in the DEI-NGC obtained by iterative best add-only moves for $n = 1000$ over 50 runs starting from a random or segregated tree and grid.

where agents can only create edges. This dynamics is particularly natural when modeling social networks, as confirmed by the observation that many real-world social networks get denser over time [Leskovec *et al.*, 2005]. The dynamics proceed until the consideration of no agent changes the network.

See [Bullinger *et al.*, 2022] for a detailed discussion of our experimental setup and further results. An exemplary consideration of the dynamics based on the cost function of the DEI-NGC can be found in Figure 4 and Figure 5 for the general and add-only version, respectively.³ Interestingly, the results for the ICF-NGC are qualitatively the same, regardless of measuring segregation with the local or global segregation measure. The experiments indicate that the segregation strength is proportional to α , with low segregation for low α , despite the theoretical necessity of high segregation for α close to 1.⁴ Moreover, except for high α , the initial agent distribution influences the segregation, with more observed segregation for segregated initial states. The structure of the initial network seems less important for the qualitative behavior. Interestingly, the add-only version displays a similar

³As discussed in the full version, for computational efficiency, we consider convergence to 1.01-approximate pairwise stable states which is qualitatively similar to pairwise stability. Also, note that “random” means that the initial agent distribution is chosen uniformly at random, which implies low initial segregation.

⁴The provably high segregation for $\alpha < 1$ close to 1 is not contradicting the experimental results. Just before we reach a cost parameter of α , we hit the sweet spot where buying monochromatic edges is desirable while buying bichromatic edges is not.

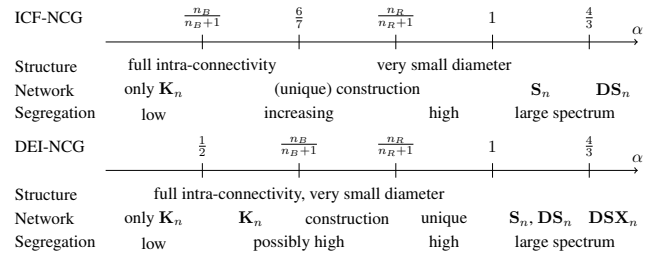


Figure 6: Overview of our theoretical results. We display structural properties of pairwise stable networks, explicit pairwise stable networks and findings about the segregation of pairwise stable networks. The two models behave surprisingly similar.

behavior for low α , but the behavior changes drastically for moderately high α . Instead of high segregation, we find that initially integrated networks converge to only moderately segregated states, whereas this is not true for initially segregated networks, suggesting an escape route from segregation.

7 Conclusion

We have investigated two network creation games that consider heterogeneous edge creation of agents acting according to homophily. Our main goal was to analyze segregation within reasonable networks measured by pairwise stability. Our results are summarized in Figure 6. Even though our two game models feature two seemingly orthogonal perspectives based on a direct and an indirect consideration of homophily, their qualitative behavior is surprisingly similar.

Clearly, stable networks are highly integrated for a very small edge cost, when agents can afford to buy all available edges. Once our cost parameter reaches the sweet spot where agents need to balance neighborhood and distance cost, there is provably high segregation, following from characterizations of stable networks. For slightly larger edge cost, our theoretical results cannot give a clear tendency of the segregation strength. In principle, both low and high segregation can be achieved by stable networks. Therefore, we performed an average-case analysis by running extensive simulation experiments. These experiments provide general tendencies about segregation contrasting the large theoretical spectrum for $\alpha \geq 1$. Most importantly, we observe low segregation under integrated initial conditions if edges cannot be deleted. This yields an escape route from segregation by a high initial investment establishing permanent integration.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under grants BR 2312/11-2, BR 2312/12-1, and LE 3445/3-1.

References

[Agarwal *et al.*, 2021] Aishwarya Agarwal, Edith Elkind, Jiarui Gan, Ayumi Igarashi, Warut Suksompong, and Alexandros A. Voudouris. Schelling games on graphs. *Artificial Intelligence*, 301:103576, 2021.

- [Allport *et al.*, 1954] Gordon W. Allport, Kenneth Clark, and Thomas Pettigrew. *The nature of prejudice*. Addison-wesley Reading, MA, 1954.
- [Amir, 1969] Yehuda Amir. Contact hypothesis in ethnic relations. *Psychological bulletin*, 71(5):319, 1969.
- [Bilò *et al.*, 2019] Davide Bilò, Tobias Friedrich, Pascal Lenzner, and Anna Melnichenko. Geometric network creation games. In *SPAA 2019*, pages 323–332, 2019.
- [Bilò *et al.*, 2020] Davide Bilò, Vittorio Bilò, Pascal Lenzner, and Louise Molitor. Topological influence and locality in swap schelling games. In *MFCS 2020*, pages 15:1–15:15, 2020.
- [Bilò *et al.*, 2021] Davide Bilò, Tobias Friedrich, Pascal Lenzner, Stefanie Lowski, and Anna Melnichenko. Selfish creation of social networks. In *AAAI 2021*, pages 5185–5193, 2021.
- [Boehmer and Elkind, 2020] Niclas Boehmer and Edith Elkind. Individual-based stability in hedonic diversity games. In *AAAI 2020*, pages 1822–1829, 2020.
- [Bredereck *et al.*, 2019] Robert Bredereck, Edith Elkind, and Ayumi Igarashi. Hedonic diversity games. In *AAMAS 2019*, pages 565–573, 2019.
- [Bullinger *et al.*, 2021] Martin Bullinger, Warut Suksompong, and Alexandros A. Voudouris. Welfare guarantees in schelling segregation. *Journal of Artificial Intelligence Research*, 71:143–174, 2021.
- [Bullinger *et al.*, 2022] Martin Bullinger, Pascal Lenzner, and Anna Melnichenko. Network creation with homophilic agents. *CoRR*, abs/2204.13757, 2022.
- [Chauhan *et al.*, 2017] Ankit Chauhan, Pascal Lenzner, Anna Melnichenko, and Louise Molitor. Selfish network creation with non-uniform edge cost. In *SAGT 2017*, pages 160–172, 2017.
- [Chauhan *et al.*, 2018] Ankit Chauhan, Pascal Lenzner, and Louise Molitor. Schelling segregation with strategic agents. In *SAGT 2018*, pages 137–149, 2018.
- [Corbo and Parkes, 2005] Jacomo Corbo and David C. Parkes. The price of selfish behavior in bilateral network formation. In *PODC 2005*, pages 99–107, 2005.
- [Cord-Landwehr *et al.*, 2014] Andreas Cord-Landwehr, Alexander Mäcker, and Friedhelm Meyer auf der Heide. Quality of service in network creation games. In *WINE 2014*, pages 423–428, 2014.
- [Dovidio *et al.*, 2003] John F. Dovidio, Samuel L. Gaertner, and Kerry Kawakami. Intergroup contact: The past, present, and the future. *Group processes & intergroup relations*, 6(1):5–21, 2003.
- [Echzell *et al.*, 2019] 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 *WINE 2019*, pages 156–170, 2019.
- [Fabrikant *et al.*, 2003] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H. Papadimitriou, and Scott Shenker. On a network creation game. In *PODC 2003*, pages 347–351, 2003.
- [Henry *et al.*, 2011] Adam D. Henry, Paweł Prałat, and Cun-Quan Zhang. Emergence of segregation in evolving social networks. *PNAS*, 108(21):8605–8610, 2011.
- [Jackson and Wolinsky, 1996] Matthew O. Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, 1996.
- [Jackson, 2010] Matthew O. Jackson. *Social and economic networks*. Princeton university press, 2010.
- [Kanellopoulos *et al.*, 2021] Panagiotis Kanellopoulos, Maria Kyropoulou, and Alexandros A. Voudouris. Modified schelling games. *Theoretical Computer Science*, 880:1–19, 2021.
- [Leskovec *et al.*, 2005] Jure Leskovec, Jon M. Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *SIGKDD 2005*, pages 177–187, 2005.
- [Mamageishvili *et al.*, 2015] Akaki Mamageishvili, Matús Mihalák, and Dominik Müller. Tree nash equilibria in the network creation game. *Internet Mathematics*, 11(4-5):472–486, 2015.
- [Martí and Zenou, 2017] Joan De Martí and Yves Zenou. Segregation in friendship networks. *The Scand. Journal of Economics*, 119(3):656–708, 2017.
- [Massey and Denton, 1988] Douglas S. Massey and Nancy A. Denton. The dimensions of residential segregation. *Social forces*, 67(2):281–315, 1988.
- [McPherson *et al.*, 2001] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.
- [Meirom *et al.*, 2014] Eli A. Meirom, Shie Mannor, and Ariel Orda. Network formation games with heterogeneous players and the internet structure. In *EC 2014*, pages 735–752, 2014.
- [Newman, 2018] Mark Newman. *Networks*. Oxford university press, 2018.
- [Paolillo and Lorenz, 2018] Rocco Paolillo and Jan Lorenz. How different homophily preferences mitigate and spur ethnic and value segregation: Schelling’s model extended. *Advances in Complex Systems*, 21(06n07):1850026, 2018.
- [Schelling, 1969] Thomas C. Schelling. Models of segregation. *American Economic Review*, 59(2):488–493, 1969.
- [Schelling, 1971] Thomas C. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186, 1971.
- [Wilensky, 1997] Uri Wilensky. Netlogo segregation model. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.*, 1997.
- [Zhang, 2011] Junfu Zhang. Tipping and residential segregation: a unified schelling model. *Journal of Regional Science*, 51(1):167–193, 2011.