# Segregation in Networks under Stochastic Dynamics

Jens Fischer, Timo Kötzing, Martin S. Krejca

University of Potsdam, Potsdam, Germany

**Abstract.** Segregation based on race, political and religious views or social norms is a reoccurring phenomenon throughout history. In the past it was openly visible due to geographic separation of the segregated groups. But with the rise of online social networks there is a new structure prone to segregation, which cannot be analyzed directly from a geographic perspective.
The formulation of appropriate models for segregation phenomena in dynamic environments is thus crucial to advance their analysis. In the spirit of Schelling's segregation model we propose a stochastic, agent based model where the agents change their relationships to satisfy their own preferences. We can show that in our model, under suitable conditions on the number of connections $m$ in the network and for binary opinions, the expected time to segregation is in $\mathcal{O}\left(\frac{m}{p}\log m\right)$, where $p \in (0,1)$ is some internal parameter of the dynamics. Finally, we extend this result to networks exhibiting $N$ different opinions and find that the expected time to segregation is in $\mathcal{O}\left(\frac{Nm}{p}\log m\right)$.

**Keywords:** Network Segregation; Markov Chains; Drift Theory; Agent Based Modeling

## 1 Introduction

Online social networks display our personal relationships more accessibly than ever before. With that comes also the public display of breaking up old and creating new relationships. But the dynamics behind the forming and maintenance of relationships in social networks have a long research history both from a modeling and an empirical point of view.

Schelling proposed in [1] and [2] models for the behaviour of interconnected agents depending on their similarity. He put the agents on the nodes of a grid and allowed them to move if their neighbors do not satisfy their personal preferences in terms of similarity. Schelling showed that this model leads to residential segregation of the agents regardless of their aversion or tolerance of each other. In Schelling's model, the agents' dynamic is inherently constrained by the grid and, therefore, by the structure of their environment. For the empirical view see for example [3].
But with the rise of online social networks there emerged social networks, which

are not subject to this constraint. Thus, Schelling's model is not applicable to these networks. Nonetheless, we observe segregation, due to diverging political views or social values, within these networks. Thus there is a demand for appropriate models to get insight into these segregation dynamics. Henry, Prałat and Zhang proposed in [4] a graph based model where the graph's vertices are considered as agents and the edges as their respective relationships. They examine segregation in such networks due to dissimilarities between the agents. To model the similarities the authors assign values in $[-1,1]$ to each agent and describe the similarity of two agents via a distance $d$ on $[-1,1]$. The random dynamics defined on the graph then lead to segregation in the asymptotic limit where the number of agents and the number of relationships tend to infinity. The authors exploit an approximation discussed for example by Wormald in [5], introducing an artificial continuous time and showing convergence for the approximating process. For a more in depth discussion of the method see [5].

In contrast to [4] we approach the network evolution from a fully discrete perspective and fix the size of our network. We use methods from drift theory, see [6] and [7], to obtain upper bounds on the expected time to segregation in $\mathcal{O}\left(\frac{m}{p}\log m\right)$. The parameter $p \in (0,1)$ therein describes the probability that a relationship between to agents is broken up. This parameter can also be found in [4] where the authors call it aversion bias. We exploit the fact that in our model the transition probabilities from relationships between agents of low similarity to relationships between agents of high similarity are proportional to the number of connections between agents of different opinion. Similarity is here to be understood as the difference in the opinions of two agents. Furthermore, we derive explicit lower bounds on these transition probabilities. In contrast to [4], we exclusively consider discrete opinions, while in [4] the authors only discuss opinions as continuous objects. The opinion an agent has is represented by the color, the corresponding vertex in the network has. We begin the discussion with a network evolution process where we allow multiple edges between one pair of vertices and exactly two colors. Subsequently, we constrain us to the more realistic setting where there is at most one link between two agents, represented by at most one edge between two vertices. In this context, we develop a method that allows us to extend our results directly to $k$ different colors and therefore a multitude of opinions in our network. We observe segregation in all discussed cases and obtain bounds on the expected time to segregation in $\mathcal{O}\left(\frac{km}{p}\log m\right)$.

## 2 Preliminaries

The goal of this work is to model segregation within networks as a time dependent process. The segregation will be due to random behaviour of the agents that make up the network. The agents' behaviour is governed by the wish to reduce over time the number of relationships with agents who have a different opinion. Our model is based on the idea to model networks via graphs $G$. The agents within the network are represented by the graphs vertices and the agents' relationships are given by the edges of $G$. We denote the set of vertices by $V$ and the set of edges by $E$. To incorporate the agents' opinions we associate a color to each vertex in the graph. Edges between identically colored vertices are then understood as conflict free relationships between two agents, while edges between differently colored vertices represent conflict-prone relationships. We want to access the different edge types within a graph-theoretical framework and, thus, fix the following notions.

**Definition 1 (Monochromatic and bichromatic edges).** *Consider an $N$-color colored Graph $G = (V, E)$. For $k \in V$ denote by $c_k$ the color of $k$. We call an edge $e = \{i, j\} \in E$ between vertices $i, j \in E$ monochromatic if and only if $c_i = c_j$. Otherwise, we call $\{i, j\}$ bichromatic. Denote the corresponding subsets of $E$ by $E^m$ and $E^b$.*

The model runs over discrete time steps, where in every time step there is a possibility that a randomly chosen bichromatic edge is being erased and replaced with some other edge, which shares one of the end nodes with the erased edge. A bichromatic edge is erased with a fixed probability $p \in (0, 1)$ but might be replaced with a bichromatic edge. In the following discussion, we are focusing on differently structured graphs with varying amounts of distinct colors and show segregation, as well as derive upper bounds on the expected time to segregation. We model the discrete time evolution process of some network via a sequence of colored graphs, which we denote by $(G_t)_{t\in\mathbb{N}}$. At each time step $t \geq 0$ the graph $G_t$ consists of a set of vertices $V$ and a set of edges $E_t$ with $n = |V|$ and $m = |E_t|$. Note that the set of vertices $V$, its size $n$ and the number of edges $m$ do not depend on $t$ and are, therefore, fixed throughout the evolution process. Additionally, the coloring is fixed, i.e., the vertices may not change their color over time. Lastly, we denote an edge between two vertices $i$ and $j$ by $\{i, j\}$.

## 3 Two Color Graph with Multiple Edges

In this section we study the network evolution allowing for multiple edges to exist or form between a fixed pair of vertices. We introduce the basic dynamics and the main tools for our analysis. The tools from this section will reoccur throughout this work.

### 3.1 Initial Conditions

We assign to each vertex $i \in V$ a color $c_i \in \{-1, 1\}$ and denote the set of vertices of color 1 and $-1$ by $V_1$ and $V_{-1}$, respectively. The cardinalities of those two sets are then abbreviated by $n_1 = |V_1|$ and $n_{-1} = |V_{-1}|$. To avoid trivialities, we assume that $n$ is much larger than 4. Finally, we define the following transition between graphs.

At each time step $t \geq 1$ we draw uniformly a random edge $\{i, j\} \in E_{t-1}$. Given that it is bichromatic we split the edge with probability $p \in (0, 1)$ and draw uniformly a vertex $D_{\{i,j\}}$ from $\{I, J\}$. We then draw uniformly another vertex

$K \in V\backslash\{D_{\{i,j\}}\}$ and rewire $D_{\{i,j\}}$ with $K$, i.e., we add the edge $(D_{\{i,j\}}, K)$ to $E_t$.

### 3.2 Result

While our result is on segregation of graphs, we have not yet properly defined what segregation means in our framework. We define segregation in a more general sense to reuse the definition later on.

**Definition 2 (Segregation of colored graphs).** *Consider a $N$-color colored graph $G = (V, E)$. We call $G$ segregated if and only if all edges are monochromatic. We say that a sequence of graphs $(G_t)_{n\in\mathbb{N}}$ segregates if and only there is some $t_0 \geq 0$ such that $G_{t_0}$ is segregated. Furthermore, we define the time to segregation by*

$$T = \inf\{t \in \mathbb{N} | G_t \textit{ is segregated}\}.$$

Drawing from this definition we state our result as follows.

**Theorem 1.** *Let $\mathcal{G}^2(n, m)$ be the set of two color colored graphs $G = (V, E)$ with $n = |V|$ and $m = |E|$. Then the expected time to segregation satisfies*

$$\mathbb{E}[T] = \mathcal{O}\left(\frac{m}{p}\log m\right) \tag{1}$$

*such that the graph process $(G_t)_{t\in\mathbb{N}}$ segregates almost surely in finite time under the previously described dynamics.*

The proof rests mainly on the fact that we consider a discrete time evolution of a graph with fixed number of edges where in each time step $t$ the number of bichromatic edges may at most be reduced by 1. That is only possible, if we draw a bichromatic edge at time $t$. We define the process $Y = (Y_t)_{t\in\mathbb{N}}$ by

$$Y_t = \mathrm{card}(\{\{i, j\} \in E_t | \{i, j\} \text{ bichromatic}\}), \tag{2}$$

which represents the number of bichromatic edges at time $t$ and shows us how far away the network is from segregation. In particular, we have the relation that $Y_t = 0$ if and only if $G_t$ is segregated. Since the probability to draw a bichromatic edge from the set of all edges is proportional to the number of bichromatic edges and for any $t$ the random variable $Y_{t+1} - Y_t$ only takes values in $\{0, 1\}$, the expected value of $Y_{t+1} - Y_t$ is also proportional to the number of bichromatic edges. The following theorem exploits such properties and translates them into an upper bound for the expected value of the random time $T$, as well as a concentration inequality for $T$.

**Theorem 2 (Multiplicative Drift).** *Let $(X_t)_{t\in\mathbb{N}}$ be a sequence of random variables of $\{0, 1\} \cup S$ where $S \subset \mathbb{R}_{>1}$ and denote by $T$ the first hitting time $T = \inf\{t \in \mathbb{N} | X_t < 1\}$. Assume that there is a $\delta > 0$ such that for all $s \in S\backslash\{0\}$ and all $t \in \mathbb{N}$*

$$\mathbb{E}[X_t - X_{t+1} | X_t = s] \geq \delta s.$$

*Then*

$$\mathbb{E}[T] \leq \frac{1 + \ln(\mathbb{E}[X_0])}{\delta}. \tag{3}$$

*Further, for all $k > 0$ and $s \in S$*

$$\mathbb{P}\left[T > \frac{k + \ln(s)}{\delta} \,\middle|\, X_0 = s\right] \leq \mathrm{e}^{-k}. \tag{4}$$

Setting $X_t = Y_t$ in the multiplicative drift theorem then yields the result.

*Proof.* Consider the process $X = (X_t)_{t\in\mathbb{N}}$ as defined previously by

$$X_t = \mathrm{card}(\{\{i,j\} \in E_t | c_i \neq c_j\}).$$

Then $X$ is by definition of the dynamics a monotonously decreasing process bounded from below by 0. Denote by $T$ the first hitting time of $X$ hitting state 0, i.e.,

$$T = \inf\{t \in \mathbb{N}_0 | X_t = 0\}, \tag{5}$$

which is equivalent to the first time that $G_t$ is segregated. For all $s \in \{1, \ldots, N\}$ given $\{Y_t = s\}$ the random variable $Y_t - Y_{t+1}$ is Bernoulli distributed with parameter $p_b(s)$. Since $p_b(s)$ is the probability to make progress towards the state 0, it is the probability of drawing a bichromatic edge, splitting this edge, picking one of the end nodes and rewiring it with a vertex of the same color. Since we draw edges uniformly, we draw a bichromatic one given $\{Y_t = s\}$ with probability $\frac{s}{m}$. It is then split with probability $p$. Then we draw uniformly one of the end nodes. Thus, drawing the end node of one color happens with probability $\frac{1}{2}$. Forming a monochromatic edge happens, if we subsequently draw one of the $\alpha n - 1$ or $(1-\alpha)n - 1$ remaining nodes of the same color from all $n$ vertices. Therefore, we have that

$$\begin{aligned} p_b(s) &= \mathbb{P}[Y_t - Y_{t+1} = 1 | Y_t = s] \\ &= \frac{s}{m}\frac{p}{2}\left(\frac{\alpha n - 1}{n} + \frac{(1-\alpha)n - 1}{n}\right) = \frac{s}{m}\frac{p}{2}\frac{n-2}{n} \geq \frac{np}{4m}. \end{aligned}$$

Furthermore, by the fact that given $\{Y_t = s\}$ the random variable $Y_t - Y_{t+1}$ is Bernoulli distributed with parameter $p_b(s)$ we have for $s \in \mathbb{N}$

$$\mathbb{E}[Y_t - Y_{t+1} | Y_t = s] = p_b(s) \geq \frac{sp}{4m}.$$

By the multiplicative drift theorem we can deduce

$$\mathbb{E}[T] \leq \frac{4m(1 + \ln(\mathbb{E}[Y_0]))}{p} = \mathcal{O}\left(\frac{m}{p}\log m\right). \tag{6}$$

Consequently $\mathbb{P}[T < \infty] = 1$ and hence $\mathbb{P}[Y_\infty = 0] = 1$ such that the graph segregates almost surely in finite time.

□

The problem gets more complex as soon as we assume the graphs to be simple. In the context of social networks simple connections seem sensible, because one can only have at most one link to another person. We investigate segregation in simple graphs under a similar evolution process in the next section.

## 4 Segregation Dynamics for Simple Colored Graphs

Let $G = (V, E)$ be a simple $\mathcal{G}(n, m)$ graph, i.e., a simple graph with $n$ vertices and $m$ edges. The vertex set exhibits $N$ distinct colors $\{c^1, \ldots, c^N\}$, where for some $\alpha_i \in (0, 1)$ for $i \in \{1, \ldots, N\}$ the amount of vertices of color $c^i$ is given by $n_i := \alpha_i n$ where we assume $n_i \in \mathbb{N}$. Since it should hold that

$$n = \sum_{i=1}^{N} n_i,$$

we have

$$\sum_{i=1}^{N} \alpha_i = 1. \tag{7}$$

Note that the definition of monochromatic and bichromatic edges from section 2 is still valid. With this understanding of the colored graph we can now define the dynamics, under which we want to let such graphs evolve.

### 4.1 The Dynamics

The discrete time process $(G_t)_{t \in \mathbb{N}}$ again starts at time $t = 0$ with an initial graph $G_0 = G$ and we define the following transition between graphs in terms of the edges at time $t$ denoted by $E_t$.
At each time step $t \geq 1$, assuming that $G_{t-1}$ is a simple graph, we draw uniformly a random edge $\{i, j\} \in E_{t-1}$. Given that it is bichromatic, we split the edge with probability $p \in (0, 1)$ and draw uniformly a vertex $D_{\{i,j\}}$ from $\{I, J\}$. In contrast to the previous section, we want to ensure that a simple graph stays simple for all times $t$. Thus, we add the following steps. Denote the neighborhood of $D_{\{i,j\}}$ at time $t$ by

$$\mathcal{N}_t^{D_{\{i,j\}}} = \{k \in V | (D_{\{i,j\}}, k) \in E_{t-1}\} \backslash (\{I, J\} \backslash \{D_{\{i,j\}}\}). \tag{8}$$

We rewire $D_{\{i,j\}}$ with another $k \in V \backslash \mathcal{N}_t^{D_{\{i,j\}}}$ drawn uniformly from $V \backslash \mathcal{N}_t^{D_{\{i,j\}}}$.
Note that

$$V \backslash \mathcal{N}_t^{D_{\{i,j\}}} \cap V^c_{c_{D_{\{i,j\}}}} \neq \emptyset$$

for all $t$ since $\{I, J\} \backslash \{D_{\{i,j\}}\} \subset V \backslash \mathcal{N}_t^{D_{\{i,j\}}}$.
In fact, the evolution process erases also in this case over time all bichromatic edges under proper assumptions on $m$. We then call the graph segregated analogously to definition 2. Our results are given in the next section.

### 4.2 Results

Using the notion of segregated graphs from section 3.2, we can state our results as follows.

**Theorem 3.** *Let $\mathcal{G}^2(n,m)$ be the set of two color colored simple graphs $G = (V,E)$ with $n = |V|$ and $m = |E|$. If*

$$m \leq \frac{n^2}{64}$$

*then, when starting in $\mathcal{G}^2(n,m)$, the process $(G_t)_{t\in\mathbb{N}}$ stays almost surely in $\mathcal{G}^2(n,m)$ and*

$$\mathbb{E}[T] = \mathcal{O}\left(\frac{m}{p}\log m\right) \tag{9}$$

*such that $(G_t)_{t\in\mathbb{N}}$ segregates almost surely in finite time.*

We can extend this result to $N \geq 3$ colors as follows.

**Theorem 4.** *Let $\mathcal{G}^N(n,m)$ be the set of $N$-color colored simple graphs $G = (V,E)$ with $n = |V|$ and $m = |E|$ where $N \geq 3$ independent of $n$. Denote by $\alpha_i$ the proportion of the vertex set of color $c^i$ and assume $\alpha_i \geq \frac{1}{2N}$ as well as $\alpha_i \leq \frac{1}{2}$. Then we have the following two complementary results.*

*1) If for all $i \in \{1,\dots,N\}$ it holds $\alpha_i \geq \frac{1}{8}$ and*

$$m \leq \frac{n^2}{256}$$

*then, when starting in $\mathcal{G}^N(n,m)$, the process $(G_t)_{t\in\mathbb{N}}$ stays almost surely in $\mathcal{G}^N(n,m)$ and*

$$\mathbb{E}[T] = \mathcal{O}\left(\frac{m}{p}\log m\right) \tag{10}$$

*such that $(G_t)_{t\in\mathbb{N}}$ segregates almost surely in finite time.*

*2) If there is a $i \in \{1,\dots,N\}$ with $\alpha_i < \frac{1}{8}$ and*

$$m \leq \frac{n^2}{16N^2}$$

*then, when starting in $\mathcal{G}^N(n,m)$, the process $(G_t)_{t\in\mathbb{N}}$ stays almost surely in $\mathcal{G}^N(n,m)$ and*

$$\mathbb{E}[T] = \mathcal{O}\left(\frac{mN}{p}\log m\right) \tag{11}$$

*such that $(G_t)_{t\in\mathbb{N}}$ segregates almost surely in finite time.*

To prove these results we use again the multiplicative drift theorem, see theorem 2. We discuss its connection to our result in the next section.

### 4.3 Methodology

To show convergence we employ again the processes $(Y_t)_{t\in\mathbb{N}}$ defined by

$$Y_t = \mathrm{card}(\{\{i,j\} \in E_t | \{i,j\} \text{ bichromatic}\}) \tag{12}$$

for all $t \in \mathbb{N}$, Again, we employ techniques from drift theory to find an upper bound on $T$. To that end, by the nature of the process as described in section 3.2, we prove multiplicative drift in the sense of theorem 2. Unfortunately, setting $X_t = Y_t$ in the multiplicative drift theorem yields a possibly very small $\delta$ in the estimate

$$\mathbb{E}[X_t - X_{t+1} | X_t = s] \geq \delta s,$$

because it does not account for the local structure of the graph but only the amount of bichromatic edges. This then gives only very large upper bounds on the expected value of $T$, which are not as meaningful. We are, therefore, going to employ a function $h$, which we call a potential function, mapping a graph to a non-negative integer and satisfying for the process $(G_t)_{n\in\mathbb{N}}$

$$h(G_t) = 0 \Leftrightarrow Y_t = 0.$$

We will then use $X_t = h(G_t)$ to show multiplicative drift and, by the previous equivalence, deduce convergence of the process to a segregated graph. The potential function $h$ will take different forms depending on the case we investigate.

### 4.4 Proof of Result for two Colors

Bichromatic edges between two vertices $i \in V_1$ and $j \in V_{-1}$, which are adjacent to $n_1-1$ and $n_{-1}-1$ monochromatic edges, respectively, slow down the dynamics because splitting these edges only results in another bichromatic edge. To access this property we have to argue about the connectivity of a vertex $i$ within the set of vertices of color $c_i$.

**Definition 3.** *Consider a two color colored graph $G = (V, E)$ with $n = |V|$. The sets of identically colored vertices are called $V_1 = \{i \in V | c_i = 1\}$ and $V_{-1} = \{j \in V | c_j = -1\}$ with $n_1 = |V_1|$ and $n_{-1} = |V_{-1}|$. For a vertex $k \in V$ we define*

$$\deg^m(k) = \mathrm{card}(\{\{i,j\} \in E | i = k, j = x, x \in V_{c_k}\})$$

*where $c_k$ represents the color of $k$. We call $\deg^m(k)$ the monochromatic degree of $k$.*

The definition of the monochromatic degree of a vertex gives us a notion to work with edges incident to two vertices $i \in V_1$ and $j \in V_{-1}$, which exhibit more than $\lfloor \beta n_1 \rfloor$ and $\lfloor \beta n_{-1} \rfloor$ incident monochromatic edges. $\beta$ is a positive constant smaller than one which can be seen as connectivity parameter. $\beta$ close to one then allows a vertex to have many connections within the own realm while $\beta$ close to zero enforces many open connections between vertices.

**Definition 4.** *Consider a two color colored graph $G = (V, E)$ with $n = |V|$. Denote by $c_k$ the color of some vertex $k \in V$. The sets of identically colored vertices are called $V_1 = \{i \in V | c_i = 1\}$ and $V_{-1} = \{j \in V | c_j = -1\}$ with $n_1 = |V_1|$ and $n_{-1} = |V_{-1}|$. We call a vertex $k \in V$ $\beta$-saturated if and only if the number of monochromatic edges $(k, x)$ incident to $k$ is greater or equal to $\lfloor \beta n_{c_k} \rfloor$, i.e.,*

$$\deg^m(k) \geq \lfloor \beta n_{c_k} \rfloor$$

*for a fixed $\beta \in (0, 1)$. Assuming that $n_1 = \max\{n_1, n_{-1}\}$ we call a bichromatic edge $e = \{i, j\} \in E$ saturated iff $i \in V_1$ is saturated. Otherwise, the edge is called non-saturated.*

In what follows we set $\beta = \frac{1}{2}$ and call $\frac{1}{2}$-saturated vertices saturated.

We extend our definitions by

$$E_t^{b,n} := \{\{i, j\} \in E_t | \{i, j\} \text{ bichromatic, not saturated}\}$$
$$E_t^{b,s} := \{\{i, j\} \in E_t | \{i, j\} \text{ bichromatic, saturated}\}.$$

We now have the necessary tools to show the result in the case of a two color colored graph $G = (V, E)$ with $n = |V|$ and $m = |E|$ where $n_1 := \lceil \alpha n \rceil$ and $n_{-1} := \lfloor (1 - \alpha) n \rfloor$. Without loss of generality we assume that

$$n_1 = \max\{n_1, n_{-1}\}, \tag{13}$$

i.e., $\alpha \in (\frac{1}{2}, 1)$. Additionally, we assume that $\frac{n_1}{2}$ is even[1]. By assumption on $m$ we have

$$m \leq \frac{n^2}{64} < \frac{(\lceil \alpha n \rceil)^2}{16}$$

which implies that at most $\frac{n_1}{4}$ vertices in $n_1$ may be saturated in the sense of definition 4 setting $\beta = \frac{1}{2}$. This is due to the fact that by assuming that $\frac{n_1}{2}$ is even we can form a $\frac{n_1}{2}$-regular graph $G_1^r$ of color 1 with $\frac{n_1}{4}$ vertices. Therefore, all vertices $G_1^r$ are saturated and $G_1^r$ is the graph spanned by the minimal number of edges to saturate $\frac{n_1}{4}$. But since $G_1^r := (V_1^r, E_1^r)$ is a $\frac{n_1}{2}$-regular graph with $\frac{n_1}{4}$ vertices it satisfies

$$|E_1^r| = \frac{1}{2} \frac{n_1}{4} \frac{n_1}{2} = \frac{(\lceil \alpha n \rceil)^2}{16}$$

we obtain by assuming

$$m \leq \frac{n^2}{64} < \frac{(\lceil \alpha n \rceil)^2}{16} = |E_1^r|$$

that at most $\frac{n_1}{4}$ vertices in $n_1$ may be saturated. Throughout the following calculations we use excessively that $n_1 \geq \frac{n}{2}$, which will not be mentioned explicitly at every instance.

To access the transition probabilities we can define the following amounts.

[1] If $\frac{n_1}{2}$ is odd then $\frac{n_1}{2} - 1$ is even and the discussion that follows is made with $\frac{n_1}{2} - 1$. This assumption is only made to avoid the $-1$ and, thus, increase readability.

- $l_m = \#$ monochromatic edges
- $l_{b,s} = \#$ bichromatic saturated edges
- $l_{b,n} = \#$ bichromatic non-saturated edges

Evidently, $m = l_m + l_{b,s} + l_{b,n}$. The process may then be seen as a replacement procedure that, in case of convergence, leads to $m = l_m$ for the graph $G_\infty$. The replacement procedure may be depicted as the transition graph in figure 1. This

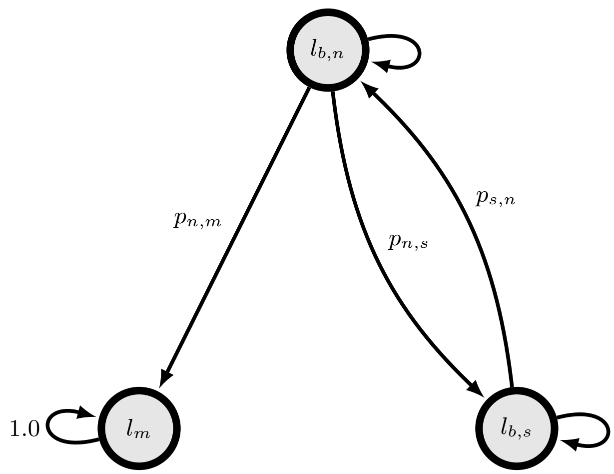


**Fig. 1.** The transition graph that displays the probabilities to replace an edge of some kind with another one of same or different kind. Only probabilities leading to a real transition carry names.

perspective leads to the following decomposition of the potential.

$$h(G_t) = 16 l_{b,n} + 17 l_{b,s}.$$

By the assumption on the number of saturated vertices we have the following property. When drawing an edge of type $l_{b,n}$ the end node in $V_1$ is not saturated and therefore has more than $\frac{n_1}{2}$ possibilities to form a monochromatic edge when being chosen. Therefore,

$$p_{n,m} \geq \frac{1}{2}\frac{n_1}{2n} \geq \frac{1}{2}\frac{n}{4n} = \frac{1}{8} \tag{14}$$

and by definition $p_{n,s} \leq 1$. Consequently,

$$\begin{aligned}
\mathbb{E}[h(G_t) - h(G_{t+1}) | G_t = G] &= p\left(\frac{l_{b,n}}{m}(16 p_{n,m} - p_{n,s}) + \frac{l_{b,s}}{m} p_{s,n}\right) \\
&= \frac{p}{m}\left(l_{b,n}(16 p_{n,m} - p_{n,s}) + l_{b,s} p_{s,n}\right) \\
&\geq \frac{p}{m}\left(l_{b,n}\left(16\frac{1}{8} - 1\right) + l_{b,s} p_{s,n}\right) = \frac{p}{m}\left(l_{b,n} + l_{b,s} p_{s,n}\right).
\end{aligned}$$

To employ the theorem 2 we need to find a lower bound of $\frac{p}{m}\left(l_{b,n} + l_{b,s}p_{s,n}\right)$ of the form $h(G)\delta$. To that end, we go over the different distinct cases for $l_{b,n}$ and $l_{b,s}$ that may arise.

If $l_{b,s} = 0$ set $\delta' = \frac{1}{16}$. Then we have

$$\mathbb{E}[h(G_t) - h(G_{t+1})|G_t = G] = \frac{p}{m} l_{b,n} = \frac{p\delta'}{m}\left(16 l_{b,n} + 17 l_{b,s}\right) = h(G)\frac{p\delta'}{m}.$$

The second case to consider is given by $l_{b,s} > 0$. Thus, it is possible to draw a saturated edge. In this case, it is only possible to make progress if a transition from a saturated edge to a non-saturated edge takes place. That is again only possible, if we rewire the node contained in the smaller subset $V_{-1}$ of $V$ with a non-saturated vertex in $V_1$. Unfortunately, the end node contained in $V_{-1}$ may have many connections with vertices in $V_1$, such that the transition probability becomes very small. To examine this situation we define the sets

$$\begin{aligned} A &:= \left\{ e \text{ saturated} | \mathcal{D}_s^{-e_{-1}} \geq \frac{n_1}{2} \right\}, \\ B &:= \left\{ e \text{ saturated} | \mathcal{D}_s^{-e_{-1}} < \frac{n_1}{2} \right\} \end{aligned}$$

where we set for $i \in V_{-1}$ the parameter $\mathcal{D}_s^{-i}$ as the number of unsaturated vertices of color 1 to which $i$ is adjacent. $A$ is thereby the set of edges, which exhibit end nodes in $V_{-1}$ with many connections to vertices in $V_1$. Note that $l_{b,s} = |A \cup B| = |A \sqcup B| = |A| + |B|$. Thus,

$$\begin{aligned} p_{s,n} \geq \frac{1}{l_{b,s}} \sum_{e \in B} \mathbb{P}[&\text{transition from } e \text{ to semi-saturated edge} \\ &|G_t = G, e \text{ has been split}]. \end{aligned} \tag{15}$$

For $e \in B$ we find the probability to transition from saturated to non-saturated by first choosing with probability $\frac{1}{2}$ the endpoint $e_{-1}$ which has less than $\frac{n_1}{2}$ adjacent unsaturated nodes of opposite color. Secondly, from $e_{-1}$ we can rewire with at least $\frac{n_1}{4}$ vertices to obtain a semi-saturated edge, such that

$$\begin{aligned} &\mathbb{P}[\text{transition from } e \text{ to semi-saturated edge} | G_t = G, e \text{ has been split}] \\ &\qquad \geq \frac{1}{2}\frac{n_1}{4n} \geq \frac{1}{2}\frac{n}{8n} = \frac{1}{16}. \end{aligned} \tag{16}$$

Furthermore, since each $e \in A$ induces at least $\frac{n_1}{2}$ non-saturated edges it holds $l_{b,n} \geq \dfrac{n}{2}$ and

$$m \geq |A|\frac{n_1}{2} \geq \frac{|A|n}{4} \Rightarrow \frac{n^2}{64} \geq \frac{|A|n}{4} \Leftrightarrow |A| \leq \frac{4n^2}{64n} \leq \frac{n}{16}.$$

Therefore, $|A| \leq l_{b,n}$. We conclude with deriving the expected change

$$\begin{aligned}\mathbb{E}[h(G_t) - h(G_{t+1})|G_t = G] &\geq \frac{p}{m}\left(l_{b,n} + l_{b,s}p_{s,n}\right)\\ &\geq \frac{p}{m}\left(l_{b,n} + \frac{|B|}{16}\right) \geq \frac{p}{16m}\left(l_{b,n} + |B|\right)\\ &\geq \frac{p}{16m}\left(\frac{1}{2}l_{b,n} + \frac{1}{2}|A| + |B|\right)\\ &\geq \frac{p}{32m}\left(l_{b,n} + l_{b,s}\right) = \frac{p}{544m}\left(17l_{b,n} + 17l_{b,s}\right)\\ &\geq \frac{p}{544m}\left(16l_{b,n} + 17l_{b,s}\right) = \frac{p}{544m}h(G).\end{aligned}$$

Note that the cases $|A| = 0$ or $|B| = 0$ allow the same estimates, because for $|A| = 0$ the inequality $|A| \leq l_{b,n}$ still holds and on the other hand for $|B| = 0$ we have $l_{b,s}p_{s,n} \geq 0 = \frac{|B|}{16}$. Therefore, we have proven in all possible cases multiplicative drift in sense of theorem 2 with constant

$$\delta = \min\left\{\frac{p}{16m}, \frac{p}{544m}\right\} = \frac{p}{544m}. \tag{17}$$

Therefore, by the drift theorem 4.1, the time to segregation $T$ satisfies

$$\mathbb{E}[T] \leq \frac{1 + \ln \mathbb{E}[h(G_0)]}{\delta} = \mathcal{O}\left(\frac{m}{p}\log m\right).$$

Thus, under the given evolution process we have segregation almost surely in finite time.

□

In fact, we can extend the method to subdivide the set of saturated edges in the sets $A$ and $B$ to argue about segregation in the multicolor case, i.e, where the number of colors is greater or equal to 3.

### 4.5 Proof of Results for $N$ Colors

We can extend the method which we used to prove the multiplicative drift in the preceding section to prove multiplicative drift when we increase the number of different colors from 2 to $N \geq 3$. At this point, we want to emphasize that we could include the case $N = 2$ in this discussion. We exclude it, nonetheless, because the lower bound for the drift $\delta$ we obtained in the previous section is larger and therefore preferable. The difference in the constants is due to the fact that, beforehand, we did not have to account for the variability in $N$. Hence, the following discussion is based on $N \geq 3$. First of all, we have to adjust the definition of saturated vertices and the different kinds of edges.

**Definition 5.** *Consider an $N$-color colored graph $G = (V, E)$ with $n = |V|$. Denote by $c_k$ the color of some vertex $k \in V$. For $i \in V$ we define the set of* *identically colored vertices as $i$ by $V_i = \{k \in V | c_k = c_i\}$ with $n_i = |V_i|$. We call a vertex $k \in V$ $\beta$-saturated, if and only if for a fixed $\beta \in (0,1)$ the number of monochromatic edges $\{k,x\}$ incident to $k$ is greater or equal to $\lfloor \beta n_k \rfloor$, i.e.,*

$$\deg^m(k) \geq \lfloor \beta n_k \rfloor$$

*We call a bichromatic edge $e = \{i,j\} \in E$ between $V_i$ and $V_j$, if both $i \in V_i$ and $j \in V_j$ are saturated. If exactly one of the two end nodes of $e = \{i,j\}$ is saturated we call $e$ semi-saturated. Otherwise, the edge is called fully non-saturated.*

In what follows, we, again, set $\beta = \frac{1}{2}$ and call $\frac{1}{2}$-saturated vertices saturated. Let $G = (V,E)$ be an $N$-color colored graph with $n = |V|$, $m = |E|$ and $3 \leq N$ where $N$ is a constant independent of $n$. By the assumptions in theorem 4, the share of vertices of color $c_i$ is given by $n_i = \alpha_i n$ with $n_i < \frac{n}{2}$, as well as $\alpha_i \geq \frac{1}{2N}$ for all $i \in \{1,\ldots,N\}$. The colors are, therefore, more or less evenly distributed between all vertices and no color has the majority, but there is also no negligible vertex set. Additionally, we assume that $n_i$ is even. This is a convenience assumption to avoid extensive use of the floor function, which would reduce the overall readability.

By assumption in the first case in theorem 4 we have for all $\alpha_i$ that $\alpha_i \geq \frac{1}{8}$ and

$$m \leq \frac{n^2}{256}, \tag{18}$$

which is in particular the case if $N \leq 4$. We can generate $\frac{n}{8}$ saturated vertices by spanning a $\frac{n}{16}$-regular graph within one of the colors. In doing this, we need the least amount of edges $\tilde{m}$ to generate $\frac{n}{8}$ saturated vertices since each edge contributes to saturating 2 vertices. To span this $\frac{n}{16}$-regular graph we need

$$\tilde{m} = \frac{\frac{n}{8}\frac{n}{16}}{2} = \frac{n^2}{256} \geq m$$

edges such that the assumption on $m$ allows at most $\frac{n}{8}$ saturated vertices. Given that at some time point $t$ the graph $G_t = G$, we want to make statements about the transition probabilities for each edge type. To access the transition probabilities we define the following amounts.

- $l_m = \#$ monochromatic edges
- $l_{b,s} = \#$ bichromatic saturated edges
- $l_{b,b} = \#$ bichromatic semi-saturated edges
- $l_{b,n} = \#$ bichromatic fully non-saturated edges

As in the previous seciton we evidently have $m = l_m + l_{b,s} + l_{b,n} + l_{b,b}$. The process may again be seen as a replacement procedure that, in case of convergence to some graph $G_\infty$, leads to $m = l_m$ almost surely for the graph $G_\infty$. The replacement procedure, depicted as a transition graph in figure 2, takes a slightly different form than in the previous section. This perspective leads to the following suitable potential function

$$h(G_t) = 18 l_{b,n} + 19 l_{b,b} + 20 l_{b,s}$$

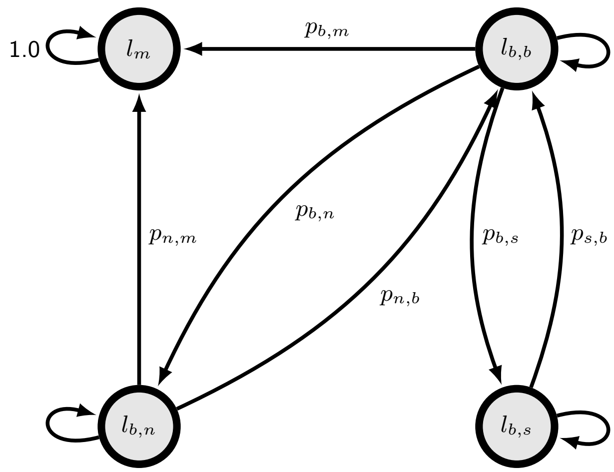


**Fig. 2.** The transition graph that displays the probabilities to replace an edge of some kind with another one of same or different kind. Only probabilities leading to a real transition carry names.

and the expected change

$$\begin{aligned}&\mathbb{E}[h(G_t) - h(G_{t+1})|G_t = G]\\ =& \frac{p}{m}\left(l_{b,n}\left(18p_{n,m} - p_{n,b}\right) + l_{b,b}\left(19p_{b,m} + p_{b,n} - p_{b,s}\right) + l_{b,s}p_{s,b}\right).\end{aligned}$$

By the assumption on the number of saturated vertices we have the following property. When drawing an edge $e$ of type $l_{b,n}$ such that $e = \{i, j\}$ is not saturated and therefore has more than $\frac{n_i}{2}$ or $\frac{n_j}{2}$ possibilities to form a monochromatic edge when being chosen. Therefore,

$$\begin{aligned}p_{n,m} &\geq \frac{1}{l_{b,n}} \sum_{e=\{i,j\}\in l_{b,n}} \frac{1}{2}\left(\frac{\frac{n_i}{2} + \frac{n_j}{2}}{n}\right)\\ &= \frac{1}{l_{b,n}} \sum_{e=\{i,j\}\in l_{b,n}} \frac{\alpha_i + \alpha_j}{4} \geq \frac{1}{16}.\end{aligned} \tag{19}$$

On the other hand

$$p_{n,b} \leq \frac{n}{8n} = \frac{1}{8}, \tag{20}$$

because there are at most $\dfrac{n}{8}$ saturated nodes.

Consider $e \in l_{b,b}, e = \{i, j\}$. Without loss of generality assume $j$ be the saturated node. Denote by $\mathcal{D}_s^{-i}$ the number of unsaturated vertices of color $-c_i$

to which $i$ is adjacent. Set

$$n_{-i} := \left| \bigcup_{l \neq i} V_l \right| \geq \frac{n}{2}.$$

Define the sets

$$A_b := \left\{ e = \{i,j\} \text{ semi-saturated} | i \text{ not saturated}, \mathcal{D}_s^{-i} \geq \frac{n_{-i}}{4} \right\},$$
$$B_b := \left\{ e = \{i,j\} \text{ semi-saturated} | i \text{ not saturated}, \mathcal{D}_s^{-i} < \frac{n_{-i}}{4} \right\},$$

where the assumption $i$ not saturated, is just for notation convenience. Then $l_{b,b} = |A_b \sqcup B_b| = |A_b| + |B_b|$ and since each $e \in A_b$ induces at least $\frac{n_{-i}}{4}$ fully non-saturated edges it holds $l_{b,n} \geq \frac{n}{8}$. Under the assumption

$$m \leq \frac{n^2}{256} \tag{21}$$

we derive

$$m \geq |A_b| \frac{n_{-i}}{4} \geq \frac{|A_b| n}{8} \Rightarrow \frac{n^2}{256} \geq \frac{|A_b| n}{8} \Leftrightarrow |A_b| \leq \frac{8n^2}{256n} = \frac{n}{32} \leq \frac{n}{32} \leq l_{b,n}.$$

Therefore, we obtain $|A_b| \leq l_{b,n}$.

For $e \in B_b$ we find a way to transition from semi-saturated to fully non-saturated by first choosing with probability $\frac{1}{2}$ the endpoint $e_i$ which is not saturated and has less than $\frac{n_{-i}}{4}$ adjacent unsaturated nodes of different color. Since there are at most $\frac{n}{8}$ saturated vertices we can rewire $e_i$ with at least

$$n_{-i} - \frac{n_{-i}}{4} - \frac{n}{8} = \frac{3n_{-i}}{4} - \frac{n}{8} \geq \frac{3n}{8} - \frac{n}{8} = \frac{n}{4}$$

vertices to form a fully non-saturated edge, such that

$$\begin{aligned} &\mathbb{P}[\text{transition from } e \text{ to fully non-saturated} | G_t = G, e \text{ has been split}] \\ &\quad \geq \frac{1}{2} \frac{n}{4n} = \frac{1}{8}. \end{aligned} \tag{22}$$

The transition from $e \in l_{b,b}$ to monochromatic happens if $i$ is drawn and consequently rewired with one of the more than $\frac{n_i}{2}$ which amounts to the probability

$$\begin{aligned} &\mathbb{P}[\text{transition from } e \text{ to monochromatic} | G_t = G, e \text{ has been split}] \\ &\quad \geq \frac{1}{2} \frac{n_i}{2n} = \frac{\alpha_i}{4}. \end{aligned} \tag{23}$$

Thus,

$$p_{b,m} \geq \frac{1}{l_{b,b}} \sum_{e=\{i,j\} \in l_{b,b}} \frac{\alpha_i}{4} \geq \frac{1}{32}. \tag{24}$$

Transforming a semi-saturated edge to a saturated one happens if we pick the saturated end node and rewire it with a saturated node. Since there are at most $\frac{n}{8}$ saturated nodes, we have

$$p_{b,s} \leq \frac{1}{2}\frac{n}{8n} = \frac{1}{16}. \tag{25}$$

Continuing with the saturated edges we arrive at a similar result by using the same method. Define the sets

$$\begin{aligned} A_s &:= \left\{ e = \{i,j\} \text{ saturated} | \mathcal{D}_s^{-i} \geq \frac{n_{-i}}{4} \text{ and } \mathcal{D}_s^{-j} \geq \frac{n_{-j}}{4} \right\}, \\ B_s &:= \left\{ e = \{i,j\} \text{ saturated} | \mathcal{D}_s^{-i} < \frac{n_{-i}}{4} \text{ or } \mathcal{D}_s^{-j} < \frac{n_{-j}}{4} \right\}. \end{aligned}$$

Then $l_{b,s} = |A_s \sqcup B_s| = |A_s| + |B_s|$ and since each $e \in A_b$ induces at least $\frac{n_{-i}}{4} + \frac{n_{-j}}{4} \geq \frac{n}{4}$ semi-saturated edges it holds $l_{b,b} \geq \frac{n}{4}$. Again, we by the assumptions on $m$

$$m \leq \frac{n^2}{256} \tag{26}$$

we derive

$$m \geq |A_s| \left( \frac{n_{-i}}{4} + \frac{n_{-j}}{4} \right) \geq \frac{|A_s| n}{4} \Rightarrow \frac{n^2}{256} \geq \frac{|A_s| n}{4} \Leftrightarrow |A_s| \leq \frac{4n^2}{256n} \leq \frac{n}{64} \leq l_{b,b}.$$

Thus, we have $|A_b| \leq l_{b,b}$.

For $e \in B_s$ we find the probability to transition from saturated to non-saturated by first choosing with probability $\frac{1}{2}$ the distinguished endpoint $e_*$ which has less than $\frac{n_{-i}}{4}$ adjacent unsaturated nodes of different color. Since there are at most $\frac{n}{8}$ saturated vertices we can rewire $e_i$ with at least

$$n_{-i} - \frac{n_{-i}}{4} - \frac{n}{8} = \frac{3n_{-i}}{4} - \frac{n}{8} \geq \frac{3n}{8} - \frac{n}{8} = \frac{n}{4}$$

vertices to obtain a semi-saturated edge, such that

$$\begin{aligned} &\mathbb{P}[\text{transition from } e \text{ to semi-saturated} | G_t = G, e \text{ has been split}] \\ &\quad \geq \frac{1}{2}\frac{n}{4n} = \frac{1}{8}. \end{aligned} \tag{27}$$

Collecting all these estimates from equation (19) to (27) we get

$$\begin{aligned}
&\mathbb{E}[h(G_t) - h(G_{t+1})|G_t = G] \\
=& \frac{p}{m}\left(l_{b,n}\left(18p_{n,m} - p_{n,b}\right) + l_{b,b}\left(19p_{b,m} + p_{b,n} - p_{b,s}\right) + l_{b,s}p_{s,b}\right) \\
\geq& \frac{p}{m}\left(l_{b,n} + l_{b,b}p_{b,n} + l_{b,s}p_{s,b}\right) \\
\geq& \frac{p}{m}\left(l_{b,n} + |B_b|\frac{1}{8} + |B_s|\frac{1}{8}\right) \geq \frac{p}{8m}\left(\frac{1}{2}l_{b,n} + \frac{1}{2}|A_b| + |B_b| + |B_s|\right) \\
\geq& \frac{p}{16m}\left(l_{b,n} + l_{b,b} + |B_s|\right) \\
\geq& \frac{p}{16m}\left(l_{b,n} + \frac{1}{2}l_{b,b} + \frac{1}{2}|A_s| + |B_s|\right) \geq \frac{p}{32m}\left(l_{b,n} + l_{b,b} + l_{b,s}\right) \\
\geq& \frac{p}{640m}\left(20l_{b,n} + 20l_{b,b} + 20l_{b,s}\right) \geq \frac{p}{640m}\left(18l_{b,n} + 19l_{b,b} + 20l_{b,s}\right).
\end{aligned}$$

Note that the cases $l_{b,n} = 0$, $l_{b,b} = 0$, $l_{b,s} = 0$, $|A_b| = 0$, $|B_b| = 0$, $|A_s| = 0$ or $|B_s| = 0$ allow the same estimates. This is due to the following reasons. If $l_{b,b} = 0$ and $l_{b,s} = 0$, then $l_{b,n} > 0$ and one can skip all estimates concerning $l_{b,b}$ and $l_{b,s}$. If $l_{b,n} = 0$, then $|A_b| = 0$ such that $|B_b| = l_{b,b}$ If, either additionally or on its own, $l_{b,b} = 0$, then $|A_s| = 0$ such that $|B_s| = l_{b,s}$ and one can estimate as above. Also if $l_{b,s} = 0$, the estimates are possible since the problematic term $l_{b,s}p_{s,b}$ disappears. As for $|A_b| = 0$, $|B_b| = 0$, $|A_s| = 0$ or $|B_s| = 0$, in the cases where $l_{b,b} \neq 0$ and $l_{b,s} \neq 0$, the same line of argumentation works as in the two color case. These arguments also apply in the second case, which we are going to discuss next. By the previous reasoning, we find multiplicative drift in sense of theorem 2 in the first case with constant

$$\delta_{p,m} = \frac{p}{640m}. \tag{28}$$

Thus, also in this case, by the drift theorem 2, the time to segregation $T$ satisfies

$$\mathbb{E}[T] \leq \frac{1 + \ln \mathbb{E}[h(G_0)]}{\delta_{p,m}} = \mathcal{O}\left(\frac{m}{p}\log m\right).$$

such that under the given evolution process we have segregation almost surely in finite time.

In the second case, where we assume that there is a $i \in \{1, \ldots, N\}$ with $\alpha_i \leq \frac{1}{8}$ and

$$m \leq \frac{n^2}{16N^2},$$

there are also at most $\frac{n}{8}$ saturated nodes. To show this we define the following objects. Let $(i_1, \ldots, i_l)$ be a vector of indices such that

$$\left|\bigcup_{j=1}^{l} V_{i_j}\right| \leq \frac{n}{8}$$

and for all $(i'_1, \ldots, i'_{l'})$ satisfying

$$\left| \bigcup_{j=1}^{l'} V_{i'_j} \right| \leq \frac{n}{8}$$

it holds

$$\left| \bigcup_{j=1}^{l'} V_{i'_j} \right| \leq \left| \bigcup_{j=1}^{l} V_{i_j} \right|. \tag{29}$$

$\bigcup_{j=1}^{l} V_{i_j}$ is therefore a largest union of differently colored vertex sets, containing at most $\frac{n}{8}$ vertices. Node that by assumption, $l \geq 1$. By assumption is $n_i$ even such that the smallest number of edges $\tilde{m}$ needed to saturate all the vertices in $\bigcup_{j=1}^{l} V_{i_j}$ is given again by spanning a $\frac{n_{i_j}}{2}$-regular graph in each $V_{i_j}$. Therefore, we find that

$$\tilde{m} = \sum_{j=1}^{l} \frac{\frac{\alpha_{i_j} n}{2} \alpha_{i_j} n}{2} = \frac{n^2}{4} \sum_{j=1}^{l} (\alpha_{i_j})^2 \geq \frac{n^2}{16N^2} \geq m.$$

Thus, the assumption on $m$ allows at most $\frac{n}{8}$ saturated vertices.

The remainder of the proof works analogously as in the first case where only the estimates change slightly. We obtain with the same calculations as before

$$p_{n,m} \geq \frac{1}{4N},$$

as well as

$$p_{b,m} \geq \frac{1}{8N}.$$

For the size of $A_s$ and $A_b$ we get the estimate that

$$m \geq |A_s| \left( \frac{n_{-i}}{4} + \frac{n_{-j}}{4} \right) \geq \frac{|A_s| n}{4} \Rightarrow \frac{n^2}{16N^2} \geq \frac{|A_s| n}{4} \Leftrightarrow |A_s| \leq \frac{4n^2}{16N^2 n} \leq \frac{n}{36} \leq l_{b,b}$$

and

$$m \geq |A_b| \frac{n_{-i}}{4} \geq \frac{|A_b| n}{8} \Rightarrow \frac{n^2}{16N^2} \geq \frac{|A_b| n}{8} \Leftrightarrow |A_b| \leq \frac{8n^2}{16N^2 n} = \frac{n}{2N^2} \leq \frac{n}{18} \leq l_{b,n}$$

such that we get the same results $|A_s| \leq l_{b,b}$ and $|A_b| \leq l_{b,n}$. All other estimates remain identical since they where solely based on the fact that we have at most $\frac{n}{8}$ saturated vertices which is satisfied in both cases.
By employing the potential

$$h_N(G_t) = \frac{9N}{2} l_{b,n} + \left( \frac{9N}{2} + 1 \right) l_{b,b} + \left( \frac{9N}{2} + 2 \right) l_{b,s}$$

we get the expected change

$$\begin{aligned} &\mathbb{E}[h_N(G_t) - h_N(G_{t+1}) | G_t = G] \\ &= \frac{p}{m} \left( l_{b,n} \left( \frac{9N}{2} p_{n,m} - p_{n,b} \right) + l_{b,b} \left( \left( \frac{9N}{2} + 1 \right) p_{b,m} + p_{b,n} - p_{b,s} \right) + l_{b,s} p_{s,b} \right). \end{aligned}$$

Using all the estimates of the transition probabilities and the techniques from the previous case we arrive at

$$\begin{aligned}
&\mathbb{E}\left[h_N(G_t) - h_N(G_{t+1}) | G_t = G\right] \\
&= \frac{p}{m}\left(l_{b,n}\left(\frac{9N}{2}p_{n,m} - p_{n,b}\right) + l_{b,b}\left(\left(\frac{9N}{2}+1\right)p_{b,m} + p_{b,n} - p_{b,s}\right) + l_{b,s}p_{s,b}\right) \\
&\geq \frac{p}{32m}\left(l_{b,n} + l_{b,b} + l_{b,s}\right) \\
&\geq \frac{2p}{32(9N+4)m}\left(\frac{9N+4}{2}l_{b,n} + \frac{9N+4}{2}l_{b,b} + \frac{9N+4}{2}l_{b,s}\right) \\
&\geq \frac{p}{16(9N+4)m}\left(\frac{9N}{2}l_{b,n} + \frac{9N+2}{2}l_{b,b} + \frac{9N+4}{2}l_{b,s}\right).
\end{aligned}$$

Note that the cases $l_{b,n} = 0$, $l_{b,b} = 0$, $l_{b,s} = 0$, $|A_b| = 0$, $|B_b| = 0$, $|A_s| = 0$ or $|B_s| = 0$ have already been discussed in the first case. Therefore, we have proven multiplicative drift in sense of theorem 2 with constant

$$\delta_{N,p,m} = \frac{p}{16(9N+4)m} \tag{30}$$

in the second case. Therefore, once again, by the drift theorem 2, the time to segregation $T$ satisfies

$$\mathbb{E}[T] \leq \frac{1 + \ln \mathbb{E}[h(G_0)]}{\delta_{N,p,m}} = \mathcal{O}\left(\frac{mN}{p}\log m\right).$$

Thus, under the given evolution process we have segregation almost surely in finite time.

□

## 5 Conclusion

In this work we showed that segregation happens under the prescribed process and derived the expected time to segregation to be of order $\mathcal{O}\left(\frac{m}{p}\log m\right)$ in the two color case and of order $\mathcal{O}\left(\frac{mN}{p}\log m\right)$ in the $N$ color case. In the following we want to briefly discuss the implications and possible extensions of this work.

Let us consider, again, each node as an agent with corresponding opinion according to her color. Interpret a bichromatic edge as a conflict-prone relationship between two agents of different opinion. The process can then be understood as conflict avoidance behaviour between the two agents. Either they break up their relationship or decide to agree to disagree. In the second case the relationship stays intact but remains conflict-prone such that it may still lead to conflicts and break up later on in the process. This is based on the fact that the two agents only break up with positive probability $p$ if their relationship is conflict-prone

and chosen during one time step. Our model, thus, gives insight to the behaviour of conflict-avoiding agents in interrelated heterogeneous finite populations.
Within this work we have not considered the case where there is a large crowd having the same opinion. This is related to the $N$-color problem discussed in section where the largest set of some color $i$ takes up more than half the population. Multiplicative drift in this case remains an open question.

Further extensions of this work include continuous colors to model agents with slightly as well as considerably deviating opinions. Splitting and rewiring of edges based on the euclidean distance of the colors, accounting for attachment to similarly aligned agents in the model, is one possibility to extend our model directly. Simulations suggest that segregation also happens in such a setting. The methods applied in this work based on the discrete colors do not seem to be extendable and the expected time to segregation therefore has to be approached differently, which also remains an open question.
Many more questions and models on segregation of networks under stochastic dynamics can be formulated and their investigation may give further insides to real world phenomena.

## References


1. Thomas C. Schelling. Models of Segregation. *American Economic Review*, 59(2):488–93, 1969.
2. Thomas C. Schelling. Dynamic Models of Segregation. *The Journal of Mathematical Sociology*, 1(2):143–186, jul 1971.
3. M Girvan and M E J Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821–6, jun 2002.
4. Adam Douglas Henry, Paweł Prałat, and Cun-Quan Zhang. Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 108(21):8605–10, may 2011.
5. Nicholas C. Wormald. The differential equation method for random graph processes and greedy algorithms , Nov. 1995.
6. Andreas Göbel, Timo Kötzing, and Martin S. Krejca. Intuitive Analyses via Drift Theory. may 2018.
7. Johannes Lengler. Drift Analysis. dec 2017.