Positive Algorithmic Bias Cannot Stop Fragmentation in Homophilic Social Networks

Chris Blex[1,2] and Taha Yasseri[1,2,*]

[1]Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3JS, UK

[2]The Alan Turing Institute, 96 Euston Rd, London NW1 2DB, UK

January 10, 2020

Abstract

Fragmentation, echo chambers, and their amelioration in social networks have been a growing concern in the academic and non-academic world. This paper shows how, under the assumption of homophily, echo chambers and fragmentation are system-immanent phenomena of highly flexible social networks, even under ideal conditions for heterogeneity. We achieve this by finding an analytical, network-based solution to the Schelling model and by proving that weak ties do not hinder the process. Furthermore, we derive the minimal algorithmic bias needed for the mitigation of such fragmentation. However, we show that the effect of such algorithmic bias on stopping fragmentation is limited.

Keywords: social networks, echo chambers, algorithmic bias, Schelling model, mathematical modelling, homophily

1 Introduction

Echo chambers or filter bubbles have been a concern since the early years of the commercial internet, despite lacking a clear-cut definition. Sunstein (2017) warned that too much personalisation may lead to 'online segregation', where individuals would surround themselves with people of ostensibly similar characteristics or ideologies. Van Alstyne and Brynjolfsson (2005) claim that the Internet makes it easier to find like-minded people and thus facilitates the creation of fringe communities that have a common ideology. These tendencies lead to the so-called *Fragmentation Thesis* (s. Bright, 2018) capturing the emergence of increasingly politically-driven divisions in online discussion networks. Fragmentation of (online) social networks has often been attributed to the sociological concept of homophily (Lazarsfeld and Merton, 1954; Mcpherson et al., 2001). Homophily states that individuals, which are similar on some sociological, economic, genetic or other plane are more likely to interact or form social ties. Homophily has a natural ally in *Gratification Theory*, a concept from psychology (Rosengren, 1974). It describes the selective behaviour of individuals seeking to gratify certain pre-formed preferences in a high-choice environment, such as social media. Thus, social media platforms require individuals to be highly selective in their consumption of political content or engagement in political discussion. Given that these preferences are naturally homophilic, individuals are likely to be drawn to content and discussion partners aligned with their own beliefs, which in turn is encouraged by the affordances of a social media platform (Cho, 2003; Vaccari et al., 2016). Thus social media might exacerbate an underlying sociological tendency of homogenous clusters within social networks.

Whilst clustering in online social networks is a well-established empirical fact (e.g. Adamic and Glance, 2005; Newman, 2006; Conover et al., 2011; Barberá, 2015a), a similarly interesting question is whether they are somewhat inevitable (Sasahara et al., 2019) . The segregation of offline networks takes longer given geographical and temporal constraints and is thus prone to stochastic or social shocks upsetting the fragmentation process. However online social networks are much freer in their dynamics. From these considerations this paper derives its first research question:

RQ1: *Is fragmentation in (online) social networks inevitable, given pres-*

ence of homophily and the high-pace dynamic of the structure of the network?

Echo chambers are clearly a worrying phenomenon, since they are bad for social cohesion and democratic, deliberative decision-making processes (Sunstein, 2017; Bright, 2018). Algorithmically biased recommender systems have often been blamed for fostering these segregative tendencies. Thus, if part of the problem is algorithmic so should be its remedy. Therefore this paper derives its second research question:

RQ2: *Can algorithmic bias counteract homophilic fragmentation of social networks? If so, what is the minimal algorithmic bias needed for fragmentation not to occur?*

This paper tries to answer these research questions by building and analytically solving a network representation of the Schelling model. The latter originally models spatial and residential segregation based on homophilic preferences of individuals (Schelling, 1971). Our model uses the benchmark case of an 'ideal liberal' network, i.e. a network in which individuals have no inherent biases in forming their connections (e.g. preferential attachment, social, political or racial biases inherent to the network structure, algorithmic biases) or limitations with whom and with how many they will forge a connection. This paper refers to this as the 'ideal liberal' case, since it corresponds to the ideal state of a liberal society and indeed the self-image of many social media platforms. Subsequently, the paper shows how even in networks with ideal conditions for heterogeneity a limited amount of homophily is sufficient to cause complete fragmentation in the limit. The paper supports this claim by proving that the result also obtains when including weak ties. Weak ties are secondary connections between individuals, i.e. "a friend of a friend" and have often been cited as having powerful hidden influences on social networks with regards to access to resources, information sharing, or network heterogeneity (Granovetter, 1973).

Lastly, the paper tries to derive the minimal algorithmic bias needed to stop fragmentation in the network. Each individual node is confronted with the "average opinion" of the network, in this case a mean-field approximation of network heterogeneity. This will cause some nodes to change their type, i.e. from green to red, thereby changing all of its homogeneous ties to heterogeneous ones. The result shows that complete heterogeneity in the network is intractable for network administrators. This poses some serious policy and ethical concerns for the accountability of administrators of social media sites. Firstly, it posits the question that if social media administrators are able to bias the tie creation process in favour of increased network heterogeneity, then should they be allowed to, obliged, or even litigated to do so. Secondly, if social media sites may be incapable of fostering sufficient heterogeneity to prevent unsustainable levels of polarisation, the question begs whether they should be held accountable for political instability or social unrest. These are pressing ethical questions, since values such as individual freedom and privacy rights may conflict with wider political and social concerns of political uncertainty or discord.

This paper is organised in the following sections: Section 2 presents a review of related work, especially focussing on the Schelling model of segregation and the argument of weak ties posing a potential amelioration to the problem of network segregation. Section 3 presents the *Homophily Theorem*, which is a network operationalisation of the Schelling model. Section 4 presents the *Weakness of Weak Ties Theorem*, which acts as an extension to this model. It shows how even when considering weak, i.e. secondary ties, the fragmentation process is uninhibited. Section 5 derives the minimal amount of algorithmic bias required for a network not to fragment. Section 6 briefly summarises the social science interpretations and implications of these findings.

2 Literature Review

Echo chambers or filter bubbles seem to lack a coherent definition or operationalisation. One strand of studies defines echo chambers as highly-selective news diets (Gentzkow and Shapiro, 2011; Boxell et al., 2017; Dubois and Blank, 2018). Other studies have analysed echo chambers and selective exposure based on media migration (Hollander, 2015) or measuring media diets based on URL-click data against a random baseline (Nikolov et al., 2015a). Most studies on echo chambers rest on community detection and network clustering algorithms. Many of them have shown how networks seem to cluster along party lines or ideological beliefs on social media platforms (Adamic and Glance, 2005; Newman, 2006; Conover et al., 2011).
Many studies on echo chambers have tried to argue that echo chambers are overstated by pointing to the existence of weak ties (Bakshy et al., 2012;

Barberá, 2015b; Barberá et al., 2015; Barberá, 2015a; Hollander, 2015). Weak ties are supposed to counteract clustering movements by exposing users to opinions and information outside of their peer group. However, the probability of this being a decisive counteracting factor seems to be overstated if merely relying on the existence of weak ties. Iyengar and Hahn (2009) show that exposure to different views can actually entrench existing opinions even more, and Yardi and Boyd (2010) claim that whilst exposure to opposing views might be higher, meaningful engagement with them is low. Therefore, exposure to differing opinions may not be a sufficient metric to determine whether a society is polarised. Moreover, Marin (2012) shows that information holders in social networks base their decisions to share or withhold information on desire to help, reputational concerns, reluctance to appear intrusive, or fear of awkwardness resulting from negative consequences. This shows that when it comes to effective information sharing and information reception, weak ties alone are insufficient. The more important notion seems to be *bandwidth*. The latter captures the level of trust, reputation, or usefulness that a neighbouring node is associated with. Crudely speaking, a distant relative may be outside one's own echo chamber, but the effect of them posting controversial news stories on one's timeline is unlikely to affect one's opinion or the strength of the echo chamber effect. Crazy uncles rarely cause political epiphanies.

The underlying social force of echo chambers is homophily (Lazarsfeld and Merton, 1954; Mcpherson et al., 2001). Homophily is a well-documented sociological phenomenon (s. McPherson et al. for a comprehensive review) and has been frequently observed on social media (Barberá, 2015b; Barberá et al., 2015; Barberá, 2015a; Nikolov et al., 2015b; Tucker et al., 2018).
Homophily is also the underlying principle of a famous model of spatial segregation, e.g. in housing. The renowned Schelling model (Schelling, 1971) shows how even limited homophilous preferences lead to high levels of segregation: Let there be a real line with dots of two different colours randomly assorted. Let each dot have a preference for its neighbours to be of a similar colour. Should one dot move from a heterogenous neighbourhood to a more homogenous neighbourhood this trivial increases homogeneity in both neighbourhoods. But also other dots, which may have less homophilic tendencies are now proportionately surrounded by more dots of a different colour, thus reinforcing their desire to move as well. It becomes clear, how the process unravels and segregation becomes self-sustained, even if starting with very

mild homophilic preferences and perfectly heterogenous neighbourhoods.

Whilst, the original model was designed on a real line with nodes moving from one position to another in a one-dimensional space and the preferences driving segregation are hotly debated (Clark and Fossett, 2008), there has been a plethora of studies trying to generalise this model to more complex domains since. Vinkovic and Kirman (2006) adapt the model to cluster formation in physical systems, whilst Stauffer and Solomon (2007) compare residential separation to phase separation in physics. Some studies have extended the model to two- or multidimensional spaces in simulations and analytic proofs (Dall'Asta et al., 2008; Immorlica et al., 2015; Barmpalias et al., 2016, 2018). Studies more rooted in the model's home discipline of economics have included factors such as housing markets (Zhang, 2004) or used it in a game-theoretic study of collective action (Iwanaga and Namatame, 2013). Pancs et al. (2007) show how even if all actors have a strict preference for integration, the segregation result still holds. The majority of studies introduce a tolerance parameter as a perturbation to the process and find that the segregation process can be stopped for specific values of such parameters (Gauvin et al., 2009; Gracia-Lázaro et al., 2011; Hazan and Randon-Furling, 2013; Immorlica et al., 2015). Immorlica et al. (2015) find an analytic result for an unperturbed one-dimensional Schelling model, starting from a random configuration, but finding a high probability of segregation. Some models have translated the model into network environments (Banos, 2012). One study finds a network-based analytical solution to the Schelling model (Henry et al., 2011). Nodes rewire their edges stochastically based on different levels of aversion. The termination of an edge is based on attribute distance between two actors. The model finds a conversion to a distribution and derives a measure of attribute distance, measuring the degree of segregation. The intuitive result of the model is that segregation will always emerge outside of additional endogenous or exogenous drivers of network structure.

Translating the Schelling model from a real line to a network conjures some similarities and overlaps with the Axelrod model (Axelrod, 1997). The latter is a model of social influence and similarly to Schelling was based on moving around black and white dots on a checker board. Under the assumption that the more neighbours interact the more similar they become he shows how local convergence can lead to global polarisation. With probability equal to their cultural similarity a randomly chosen site on the lattice will adopt one of the cultural features of a randomly chosen neighbours. Axelrod shows how a) the number of stable regions increases with number of possible traits and decreases with range of interaction, and b) the number of stable regions decreases with more cultural features and with large territories.

The model has sparked a number of extension and combinations with the Schelling model. Guerra et al. (2010) translate the model from a lattice to a scale-free network with dynamic links and show that feature consensus is reached faster than global consensus. Lanchier and Scarlatos (2013) include homophily into Axelrod's model of social influence. That is, an interaction between two nodes persists until either feature consensus is reached or one of the nodes terminates the tie and moves somewhere else. Thus, the model combines Schelling's idea of segregative homophily with Axelrod's social influence. The paper proves in a two-state model with arbitrary number of features that when the number of features exceeds the number of states per features, the Axelrod model leads to clusters. Gracia-Lázaro et al. (2011) present another version of the Axelrod-Schelling model showing that a process of natural selection of advantageous traits seem to emerge. Rodríguez and Moreno (2010) include the mass media as an actor into the model as a node with more heavily weighted edges and a larger neighbourhood and thus more influence. They show in a numerical simulation that the monocultural state is attained with stronger dependency on mass media strength and that network size seems to drive system into a polarised society where all possible cultural configurations are present.

A recent paper by Sasahara et al. (2019) builds another hybrid Axelrod-Schelling model. They show how a network fragments and polarises naturally from a non-polarised starting point and validate their model with simulations on Twitter data. Whilst using Schelling homophily for rewiring, the driving mechanism for them is bounded social influence. Similarly, Chitra and Musco (2019) build a hybrid model of social influence and homophily using the Friedkin-Johnson model (Friedkin and Johnsen, 1999). They include a network administrator into their model, who seeks to minimise disagreement between users thus fostering filter bubbles. In fact, changing total edge weights by 40 percent increases polarisation 40 fold. They furthermore show that for a network generated by a Stochastic Block Model (SBM) that there may be low polarisation and fragile consensus, which however is very easily perturbed leading to mass polarisation. Yet, they also find that if the network administrator seeks to mitigate filter bubbles it only increases disagreement by five percent. A related paper comes from Sîrbu et al. (2019). The paper adapts the opinion dynamics model of bounded confidence by account for algorithmic bias. Discussion partners are paired at random and converge in their opinion, if their original opinion difference is sufficiently small. In the converse case, they keep their opinion as it was before. Sîrbu et al. (2019) modify this model by making the probability of two individuals interacting proportional to their opinion difference. That is, the smaller the opinion differential the more likely the interaction between two individuals. They show in a simulation that this leads to higher opinion fragmentation, increased polarisation of opinions, as well as a stark decrease in the rate of opinion convergence.

3 The Homophily Theorem

3.1 Model Intuition

A space is populated by individuals (nodes) of two types, e.g. red and green. The nodes are unconnected at the start of the model. In the 0th timestep they each form one connection to one other node with equal chance for a similar or dissimilar connection. The formation of these edges is unbiased, i.e. it is equally probable that a green node connects to another green node or a red node. From this timestep onwards, the edge formation is preferential. That is, at each timestep nodes form connections with other nodes, with a slight preference to connect to a node of similar colour. The process is as follows: e.g. if a green node has a lot of connections to red nodes there is a high probability that the next connection it forms is to another green node. The more connections to dissimilar nodes, the higher the probability of forming further connections with similar nodes. This is an application of a Schelling model and follows a simple intuition: If a node prefers to be amongst similar nodes, but is currently surrounded by dissimilar nodes, it will have a stronger preference for the next connection to be to a node of a similar kind. Note that for the model it is irrelevant whether these are strong or weak homophilic preferences. Additionally, nodes lose connections at each timestep with a certain probability. The longer a connection has persisted, the less likely a node is to lose this connection. I.e. the older the friendship, the more likely it is to persist. Letting this mechanism run its course, the network becomes entirely clustered into a green and red cluster in the limit. That is, the probability of forming a connection with a dissimilar node converges to zero.

3.2 Mathematical Formulation

Let $p_t \in [0,1]$ be the probability of one node making a connection to a node of the same type.

Let $q_t \in [0,1]$ be the probability of a node making a connection to node of a different type.

Trivially $p_t = 1 - q_t$.

Let there be a network with an infinite number of nodes n, but no edges before time $t = 0$. At $t = 0$ let each node form an edge with another node with no preferential bias with probability $p_0 \sim Bernoulli(0.5, 0.5)$. Thus, there are $\frac{n(n-1)}{2}$ pairs of connected nodes.

From $t = 1$ let the model behave like thus at each time step:

Let $E_t \in \mathbb{Z}^+$ be the total number of edges with $E_t = E_{s,t} + E_{d,t}$, where $E_{s,t}$ is the number of edges between similar nodes and $E_{d,t}$ the number of edges between different nodes. Let $\mathbb{E}[E_{d,i,t}]$ be the expected number of edges from node i to nodes of a different type to i.

$\mathbb{E}[E_{d,i,t}] = q_{i,t}E_t + k_d E_{d,i,t-1}$, where k_d is the proportion of nodes of a different type from the previous period, that are still attached to node i, i.e. the proportion of different connections kept. This process can be conveniently described by a hazard function. Therefore, let k_d be a Weibull-Process

$k_d = 1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1}$, where $\lambda \geq 1$ is a scale parameter, and $\gamma_d < 1$

This implies:

$\mathbb{E}[E_{d,i,t}] = q_{i,t}E_t + \left(1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1}\right) E_{d,i,t-1}$

Let $E_{s,i,t}$ behave symmetrically , with $\gamma_d < \gamma_s$:

$\mathbb{E}[E_{s,i,t}] = p_{i,t}E_t + k_s E_{s,i,t-1} = p_{i,t}E_t + (1 - \lambda^{-\gamma_s}\gamma_s t^{\gamma_s - 1})\ E_{s,i,t-1}$

Iterating this equation backwards yields

$\mathbb{E}[E_{d,i,t}] = \sum_{j=0}^{T} q_{i,t-j}E_{t-j} \prod_{t=1}^{\infty} (1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1})$
with $E_0 = \frac{n(n-1)}{2}$ and thus $E_{d,0} = \frac{n(n-1)}{4}$

For all $t > 0$ let $p_t = f(\frac{E_{d,t-1}}{E_t})$, where $f' \geq 0$ and $f'' < 0$. That is, the probability increases concavely in the number of edges formed to dissimilar nodes.

Theorem 1 *For any concave monotonically increasing function* $f : [0,1] \to [0,1]$, *p_t converges to 1 and thus $\mathbb{E}[E_{d,t}]$ converges to zero.*

3.3 Proof

Without loss of generality, consider the following system of the equations stipulated above in continuous time:

$$p_t = 1 - q_t = f(\frac{E_{d,t-1}}{E_t}) \tag{1}$$

$$\mathbb{E}[E_{d,i,t}] = \prod_{t=0}^{T}(1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1}) \int_0^T q_{i,t}E_t dt \tag{2}$$

where $f' > 0$ and $f'' < 0$, $\lambda \geq 1$, and $\gamma_d < 1$.

In the steady-state $\frac{\partial p_t}{\partial t} = 0$ at $f'(\frac{E_{d,t-1}}{E_t}) = 0$ and $\frac{\partial}{\partial t}\mathbb{E}[E_{d,i,t}] = 0$ and thus
$\frac{\partial}{\partial t}\mathbb{E}[E_{d,i,t}] = \frac{\partial}{\partial t}\prod_{t=0}^{T}(1 - \lambda^{-\gamma_d}\gamma_d t)^{\gamma_d - 1}) \int_0^T q_{t,i}E_t dt +$
$\prod_{t=0}^{T}(1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1})\frac{\partial}{\partial t}\int_0^T q_{i,t}E_t dt = 0$

N.B: In steady-state $q_{i,t} = q^*$ for all t and $\frac{\partial q^*}{\partial t} = 0$. Therefore, the equation reduces to

$$\frac{\partial}{\partial t}\prod_{t=0}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_i^* E_t dt =$$

$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_i^* E_t dt = 0$$

N.B: $q^* = 1 - p^*$

$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T (1-p_i^*) E_t dt = 0$$

Therefore there are two steady states, one $p_i^* = 1$ and trivially one at $t = 0$. N.B. the latter is not applicable in this model, since preferential attachment starts in $t = 1$

The steady-state is reached at

$$t^* = (\frac{\lambda^{-\gamma_d}}{\gamma_d})^{-\frac{1}{1-\gamma_d}}$$

This value for the steady-state time implies $\mathbb{E}[E_{d,i,t}] = 0$, which implies $p_t \to 1$. The latter has some intuitive implications: The time for complete fragmentation is increasing in γ_d and λ. The parameter γ_d essentially measures the rate at which old friendships decline over time. The older the friendship, the less likely it is to be severed. Thus, the higher γ_d, i.e. the lower the rate at which friendship ties are severed, the higher t^* and therefore the longer it takes to reach complete fragmentation of the network. N.B the steady-state is reached in finite time for $\gamma_d > 0$ and finite λ.

The proof above shows how social networks with limited homophily will gravitate towards complete fragmentation, even under ideal conditions for heterogeneity given a presence of limited homophily. The time it takes for the network to fragment completely increases in the severance rate of old friendship and the shape parameter of the Weibull process. Given no exact functional form other than an operationalisation of Schelling homophily was imposed on p_t, the model can incorporate other operationalisations of homophily, e.g. attribute vectors. It furthermore proves the Schelling result of segregation in a network setting.

4 The Weakness of Weak Ties Theorem

This theorem will act as an extension to the *Homophily Theorem*. It will add a form of resistance to the fragmentation based on weak ties of one group to another. Weak ties have been extensively discussed in the literature, since Granovetter's seminal paper on how weak ties are often decisive in resource and information extraction (Granovetter and Soongt, 2016). In social media analyses weak ties have often been cited as counteracting mechanisms to the existence of echo chambers (Barberá et al., 2015; Barberá, 2015a,b; Hollander, 2015; Bakshy et al., 2015)). The *Weakness of Weak Ties Theorem* will attempt to show that the result of the *Homophily Theorem* still holds even with the existence of weak ties.

4.1 Mathematical Formulation

Let node i be attached to a similar node k with probability p_i. Node k will be connected to a dissimilar node with probability $q_k = 1 - p_k$. Thus the expected number of dissimilar ties, primary and secondary (i.e. weak), for node i are given by the following expression

$$\mathbb{E}[E_{d,i,t}] = q_{i,t}E_t + \left(1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1}\right) E_{d,i,t-1} + p_i(1 - p_k)E_{d,k,t}$$

After substituting the expression for $E_{d,k,t}$, iterating the equation backwards and considering it in continuous time:

$$\mathbb{E}[E_{d,i,t}] = \int_0^T q_{i,t}E_t \prod_{t=0}^{T} (1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1}) +$$
$$(1 - q_{i,t})q_{k,t} \int_0^T q_{k,t}E_t \prod_{t=0}^{\infty} (1 - \lambda^{-\gamma_d}\gamma_d t^{\gamma_d - 1})$$

Theorem 2 *For any concave monotonically increasing function* $f : [0,1] \to [0,1]$, p_t *converges to* 1 *and thus* $\mathbb{E}[E_{d,t}]$ *converges to zero, even in the presence of secondary ties.*

4.2 Proof

In the steady state $q_{i,t} = q_i^*$ and $q_{k,t} = q_k^*$, however these are not imposed to be equal *a priori*. N.B. $\frac{\partial q_i^*}{\partial t} = \frac{\partial q_k^*}{\partial t} = 0$. From the proof above it is known that

$$\frac{\partial}{\partial t}\prod_{t=0}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_i^* E_t dt =$$
$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=0}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_i^* E_t dt = 0$$

and therefore

$$\frac{\partial}{\partial t}\prod_{t=0}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_k^* E_t dt =$$
$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T q_k^* E_t dt = 0.$$

Additionally, note that $\frac{\partial q_k^*}{\partial t} = 0$ and therefore the term $\frac{\partial q_k^*(1-q_i^*)}{\partial t}$ vanishes. Thus, the system reduces to

$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T (1-p_i^*)E_t dt +$$
$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{T}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T (1-p_k^*)E_t dt = 0$$

which can be expressed more clearly as

$$(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{\infty}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})\int_0^T [(1-p_i^*)+(1-p_k^*)]E_t dt = 0$$

From above it is known that the first, trivial solution is given by $t = 0$. For the second solution consider the second term:

$$\int_0^T [(1-p_i^*)+(1-p_k^*)]E_t dt = 0$$

Note, that E_t is always positive. Therefore, the solution can only be found via

$$(1-p_i^*)(1-p_k^*) = 0$$

Since $p_i^* \in [0,1]$ and $p_k^* \in [0,1]$, the only possible solution is given by

$p_i^* = p_k^* = 1$ implying $q_i^* = q_k^* = 0$, which is the exact same result as the *Homophily Theorem*.

Thus, the *Homophily Theorem* even persists in the presence of secondary, or weak ties. Interestingly, the steady-state is even reached at the same speed. This indicates a robustness of the *Homophily Theorem* and acts as a warning that weak ties may not be enough to hedge against echo chambers.

5 Algorithmic Bias

5.1 Mathematical Model

This paper will attempt to stop the fragmentation by finding the minimal algorithmic bias that produces a non-zero limit for the expected number of different ties $\mathbb{E}[E_{d,i,t}]$. For this, nodes can now change which faction they belong to, e.g. they can change their colour (change opinion). For simplicity, let this process only occur for when nodes have been exposed to algorithmic bias. The intuition behind this might be described as individuals being shown more content contrary to their views than they normally would be exposed to given their network and thus change their opinion. I.e. a red node with a lot of red connections will see more 'green' content than they normally would and thus changes colour with a certain probability. This is operationalised as follows: At each timestep, a certain proportion of nodes will be switching their opinion/colour based on being shown the average opinion/average colour of the network. The latter will be given by $\frac{p_t - q_t}{p_t + q_t}$. The algorithmic bias function ϕ determines what proportion of nodes will switch their colour. Thus, it also determines how many edges change from being between similar nodes to dissimilar nodes (or the reverse for a negative weight). For algebraic simplicity let T be infinite. Let the exact function of the optimal and minimal amount of algorithmic bias be unknown, but its law of motion be given by the weighted average opinion of the network $w\frac{p_t - q_t}{p_t + q_t}$. Thus, the expected number of edges between dissimilar nodes is given by:

$$\mathbb{E}[E_{d,t}] = \int_0^\infty \sum_{i=0}^{\infty} q_{i,t} E_{,it} dt + k_d E_{d,t-1} + \phi_t [E_{t-1} - E_{d,t-1}] \tag{3}$$

Iterating this equation backwards and substituting in for k_d this yields:

$$\mathbb{E}[E_{d,t}] = [\int_0^\infty \sum_{i=0}^{\infty} q_{i,t} E_{i,t} dt + \int_0^\infty \phi_t E_{t-1} dt] \prod_{t=0}^{\infty} (1 - \lambda^{-\gamma_d} \gamma_d t^{\gamma_d - 1}) - \phi_t) \tag{4}$$

ϕ is unknown, but its law of motion is given by:

$$\dot{\phi} = w\frac{p_t - q_t}{p_t + q_t} \tag{5}$$

Assuming that the network administrator wants to interfere as little as possible with the network mechanism, the problem of finding the optimal algorithmic bias becomes a constrained minimisation problem:

min ϕ_t such that $\mathbb{E}[E_{d,t}] > 0$

This yields the following Lagrangian:

$$L = \phi_t + \alpha[[\int_0^\infty \sum_{i=0}^\infty q_{i,t}E_{i,t}dt + \int_0^\infty \phi_t E_{t-1}dt]\prod_{t=0}^\infty(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1}) - \phi_t] \tag{6}$$

Since ϕ_t is unknown, the way to solve this Lagrangian is by setting up a Euler-Lagrange equation of the form $\frac{\partial L}{\partial \int \phi_t} - \frac{\partial}{\partial w_t}\frac{\partial L}{\partial \phi_t} = 0$. The first term is given by:

$\frac{\partial L}{\partial \int \phi_t} = \alpha E_{t-1}[\prod\limits_{t=0}^\infty(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1}) - \phi_t]$

The second term is given by:

$\frac{\partial}{\partial w_t}\frac{\partial L}{\partial \phi_t} = \frac{\partial}{\partial w}(1 - \alpha\prod\limits_{t=1}^\infty \phi_t[\int_0^\infty \sum\limits_{i=0}^\infty q_{i,t}E_{i,t}dt + \int_0^\infty \phi_t E_{t-1}dt)$
$= -\alpha E_{t-1}\prod\limits_{t=1}^\infty \phi_t \frac{\partial}{\partial w}\int_0^\infty \phi_t dt$

This implies:

$\prod\limits_{t=0}^\infty(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1}) = \phi_t - \frac{\partial}{\partial w}\int_0^\infty \phi_t dt$

Taking a time derivative and substituting in the expression for $\dot{\phi}$ yields:

$(1-\gamma_d)t^{\gamma_d-2}\prod\limits_{t=1}^\infty(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1}) = w\frac{1}{1-2q_t} - \int_0^\infty \frac{1}{1-2q_t}dt$

Which implies an optimal weight of:

$$w^* = (1-2q_t)[(1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{\infty}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1}) - \tfrac{1}{2}log(1-2q_t)]$$

N.B.:

$\lim_{q_t \to 0.5^-} w^* = +\infty$
$\lim_{q_t \to 0.5^+} w^*$ is undefined.

This means that the optimal weight and thus bias necessary to produce a fully heterogenous network (N.B. the original network the Schelling process was started on!) diverges to infinity.

Next, it is worth checking whether the system reaches a steady-state. The derivative of $\mathbb{E}[E_{d,t}]$ is given by:

$$\tfrac{\partial}{\partial t}\mathbb{E}[E_{d,i,t}] = (1-\gamma_d)t^{\gamma_d-2}\prod_{t=1}^{\infty}(1-\lambda^{-\gamma_d}\gamma_d t^{\gamma_d-1})[\textstyle\int_0^\infty \sum_{i=0}^{\infty} q^* E_{i,t}dt + \int_0^\infty \phi^* E_{t-1}dt]$$
$$= 0$$

This system has three mathematically possible steady-states. The first one is given trivially by $t = 0$, which is not permissible given the assumptions of the model (s. section 3). The second steady-state is given by the exact same time as in the *Homophily Theorem*. At this steady-state, $\mathbb{E}[E_{d,t}]$ is given by:

$$\mathbb{E}[E_{d,t}] = -\prod_{t=0}^{\infty}\phi_t[\textstyle\int_0^\infty \sum_{i=0}^{\infty} q_{i,t}E_{i,t}dt + \int_0^\infty \phi_t E_{t-1}dt]$$

This expression can never be strictly bigger than zero without violating the constraints of the network administrator. Thus, this steady-state is also not permissible given the assumptions made. The third steady-state is given by:

$$t = -\tfrac{1}{\Delta E_t}\tfrac{\phi^*}{q^*}$$

, which also implies that either $w < 0$, which is impermissible since $q_t \in [0, 0.5]$, given the network administrator constraint $p_t < q_t$ implying $q_t > \frac{1}{2}$, and thus also violating the constraint. Therefore, there is no permissible steady-state to this system.

Whilst there is no weight that produces perfect heterogeneity (i.e. $q_t = 0.5$), w^* can still be used to maximise network heterogeneity as much as computationally possible.Therefore, the optimal weight w^* maximising network heterogeneity depends negatively on the probability of making different connections and positively on the rate of the Weibull process. This is intuitive, since the higher the probability of a different connection the less need there is for the network administrator to intervene. Similarly, the Weibull process is culpable for connections being lost and thus needs to be offset by the algorithmic bias. The fact that the system has no steady-state is also intuitive, since any non-zero algorithmic bias will keep the system from converging. I.e. as soon as q_t risks to converge to zero the algorithmic bias will ensure that there will be another different connection. Thus, neither full fragmentation nor full heterogeneity can be reached in a stable way. The most interesting result is the limit behaviour of the algorithmic bias weight w. Given the constraint it makes sense that it is bounded by $q_t = \frac{1}{2}$. If the weight goes beyond a certain critical value, it risks turning too may different connections back into dissimilar connections. The algorithmic bias does not discriminate in its targeting between a node that is surrounded by many similar or dissimilar nodes. The lower bound is also intuitive as a weight close to zero makes the system collapse back into that of the *Homophily Theorem*. Interestingly, in order to maintain a network perfectly balanced in terms of heterogenous and homogenous ties an infinite amount of algorithmic bias is needed. It is therefore impossible for the network administrator to achieve a perfectly heterogenous system. At best, they can prevent complete or near-complete fragmentation by preventing the probability of making dissimilar connections q_t to converge to zero or trying to maximise it as much as computationally possible, given the divergence behaviour of w^*.

6 Discussion and Conclusion

The *Homophily Theorem* is a form of a stylised Impossibility Theorem. That means that even in ideal, stylised conditions heterogeneity in social networks with any degree of homophily is impossible. The type of homophily or operationalisation is irrelevant since the function p_t is as general a formulation of homophily as it could possibly be. Whilst the assumption and operationalisation seem abstract and unrealistic at first, it is worth noting that the model is thus capable of a number of different interpretations. Edges can be interpreted as friendship ties, or (more realistically) as connections in a discussion or sharing network. Thus, the severance of ties can simply mean two users stopping to talk to each other rather than the more drastic social action of unfriending. Furthermore, a severance of ties can also represent that the tie simply has become irrelevant. Whilst it may exist *pro forma* in form of a digital connection, the lack of interaction renders it meaningless. Lastly, the Weibull process can alternatively be thought of as the strength of a tie diminishing rather than the probability of it vanishing. The *Weakness of Weak Ties Theorem* makes the results of the *Homophily Theorem* more robust. It shows that even weak ties do not only not hedge against fragmentation, but do not even slow down the process. It is however important to mention that the model omits certain social factors which may drive network unification, such as a common goal, the rule of law, or cultural similarities. One such example can be found in Torok et al. (2013), where the collaborative project of a Wikipedia article increases consensus in a network over time.

Notwithstanding, it can be shown that homophily massively increases in scale and scope in social networks with high flexibility of node attachment and detachment, such as online social networks. It is this flexibility generated by a site's affordances that facilitates users to gratify their desires, such as homophily quickly and almost costlessly. Therefore, fragmentation becomes a natural point of gravitation for these networks. Whilst complete fragmentation is obviously unrealistic, it is important to see that it is a natural tendency for these types of networks. It is worth noting that the main assumption of this model is obviously that fragmentation is driven by homophily and that it ignores social influence. The question which of the two prevails more in e.g. opinion networks on social media has not been empirically answered yet. Should opinions be sufficiently robust on social media and confirmation biases sufficiently strong, it seems more realistic that homophily is the driving force of tie generation and attrition in opinion networks. Furthermore, Iyengar and Hahn (2009) show that exposure to different views can actually entrench existing opinions even more, i.e. social influence may also worsen rather than ameliorating the problem.

The model of minimal algorithmic bias shows that a network administrator may have some tools to alleviate this. Opposite to the main idea of Chitra and Musco (2019) in this model the network administrator has no incentive to minimise disagreement and seeks to stop fragmentation. The model shows that in order to do so, the administrator has to especially counteract the severance of ties. This could for instance mean that old connections, which have become irrelevant are now sought to be reinvigorated, e.g. by recommender systems pushing them back to the top of the feed. However, Sasahara et al. (2019) point out that recommending things that will be ignored is not a good strategy. They propose that preventing triadic closure would be a even more neutral interference. Yet, they also agree that complete severing of ties should be discouraged (e.g. by the inclusion of affordances such as snoozing). The most concerning result is that the network administrator is incapable of maintaining a perfectly balanced network. The algorithmic bias needed to achieve this quickly diverges to infinity. At best, the network administrator can avoid complete fragmentation, but never foster more heterogenous networks. This poses big ethical and regulatory questions on whether and how network administrators can or should interfere. It furthermore may indicate that the social forces of fragmentation in such highly flexible networks can only tentatively be mitigated by network-based or algorithmic mechanisms.

This paper suggests how echo chambers are seemingly inevitable in social networks with high flexibility of node attachment and detachment. Even in ideal conditions for heterogeneity (such as no network inherent biases for tie creation or algorithmic bias) a limited amount of homophily is sufficient to cause the network to fragment in the limit. Furthermore, this paper has shown how finding an optimal and minimal amount of algorithmic bias may be included to ameliorate the problem. However, the result shows that the network administrator is very constrained in their capabilities. Approaching a network where the proportion of homogenous and heterogenous ties are in perfect balance causes the amount of required algorithmic bias to diverge. Thus, there are major regulatory and ethical questions concerning the role of network administrators. Firstly, if network-based mechanisms such as biased recommender system can barely mitigate the problem, what is the responsibility of network administrators and social media sites (especially now that Pandora's box has been opened)? Secondly, how much should network administrators intervene in this process and how much does such intervention warrant regulation?

On a technical level the paper produces a straightforward translation of the Schelling model into a network-based model. Furthermore, whilst the Schelling model has been shown as robust in simulations it hitherto has had only one analytical solution (however not in network form). This is a considerable methodological contribution, since as Brandt et al. (2012) point out that the Schelling model is "surprisingly difficult to analytically prove or even rigorously define the segregation phenomenon observed qualitatively in simulations. Another contribution is the result that weak ties do not ameliorate the fragmentation process. Even stronger, weak ties do not even slow down fragmentation. This is a warning call, since there is a prevalent view in the literature that weak ties hedge against the building of echo chambers (Bakshy et al., 2012; Barberá, 2015b; Barberá et al., 2015; Barberá, 2015a; Hollander, 2015).

Future work could be conducted in testing the fundamental assumption of the model, namely homophily in opinion networks. That is, are dynamics in opinion networks on social media mainly driven by homophily, due to psychological phenomena like confirmation bias, or are they sufficiently malleable by neighbouring nodes. The model could be extended to incorporate such social influences (N.B. including a social influence model into the algorithmic bias equation does not qualitatively change the result). Another extension could be proving that the model holds in arbitrary dimensions of or different operationalisations of homophily. Finally, it could be shown how the model changes if the attrition process varies on node-level.

Acknowledgements

The authors would like to thank Oliver M. Crook for his advice on using Euler-Lagrange equations for finding the minimal algorithmic bias. The authors would also like to thank Renaud Lambiotte for his advice. Chris Blex was funded by the Alan Turing Institute Studentship. Taha Yasseri was partially supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1. The Funder had no role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

References

Adamic, L. A. and Glance, N. (2005). The Political Blogosphere and the 2004 US Election: Divided They Blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, pages 36 – 43.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203 – 226.

Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, 348(6239):1130–1132.

Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. (2012). The Role of Social Networks in Information Diffusion. In *Proceedings of the 21st International Conference of the World Wide Web*, pages 519–528.

Banos, A. (2012). Network Effects in Schelling's Model of Segregation: New Evidence from Agent-based Simulation. *Environment and Planning B: Planning and Design*, 39(2):393–405.

Barberá, P. (2015a). Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation using Twitter Data. *Political Analysis*, 23(1):76–91.

Barberá, P. (2015b). How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the U.S. *APSA Conference*, page 46.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10):1531–1542.

Barmpalias, G., Elwes, R., and Lewis-Pye, A. (2016). Unperturbed Schelling Segregation in Two or Three Dimensions. *Journal of Statistical Physics*, 164(6):1460–1487.

Barmpalias, G., Elwes, R., and Lewis-Pye, A. (2018). Digital morphogenesis via Schelling segregation. *Nonlinearity*, 31(4):1593–1638.

Boxell, L., Gentzkow, M., and Shapiro, J. M. (2017). Greater Internet Use is not Associated with Faster Growth in Political Polarization among US Demographic Groups. *Proceedings of the National Academy of Sciences*, 115(3):201706588.

Brandt, C., Immorlica, N., Kamath, G., and Kleinberg, R. (2012). An Analysis of One-Dimensional Schelling Segregation. pages 1–27.

Bright, J. (2018). Explaining the Emergence of Political Fragmentation on Social Media: The Role of Ideology and Extremism. *Journal of Computer-Mediated Communication*, 23(1):17–33.

Chitra, U. and Musco, C. (2019). Understanding Filter Bubbles and Polarization in Social Networks. pages 1–18.

Cho, J. (2003). Beyond a Access : The Digital Divide and Internet Uses and Gratifications. *IT & Society*, 1(4):46–72.

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

Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., and Flammini, A. (2011). Political polarization on Twitter. *Icwsm*, 133(26):89–96.

Dall'Asta, L., Castellano, C., and Marsili, M. (2008). Statistical Physics of the Schelling Model of Segregation. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(7).

Dubois, E. and Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information Communication and Society*, 21(5):729–745.

Friedkin, N. E. and Johnsen, E. C. (1999). Social influence networks and opinion change.

Gauvin, L., Vannimenus, J., and Nadal, J. P. (2009). Phase diagram of a Schelling segregation model. *European Physical Journal B*, 70(2):293–304.

Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *Quarterly Journal of Economics*, 126(4):1799–1839.

Gracia-Lázaro, C., Floría, L. M., and Moreno, Y. (2011). Selective advantage of tolerant cultural traits in the Axelrod-Schelling model. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 83(5):1–8.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380.

Granovetter, M. and Soongt, R. (2016). Threshold Models of Diversity : Chinese Restaurants , Residential Segregation , and the Spiral of Silence Author (s): Mark Granovetter and Roland Soong Published by : American Sociological Association Stable URL : http://www.jstor.org/stable/271045 Your. 18(1988):69–104.

Guerra, B., Poncela, J., Gómez-Gardeñes, J., Latora, V., and Moreno, Y. (2010). Dynamical organization towards consensus in the Axelrod model on complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 81(5).

Hazan, A. and Randon-Furling, J. (2013). A Schelling model with switching agents: Decreasing segregation via random allocation and social mobility. *European Physical Journal B*, 86(10):1–9.

Henry, A. D., Pralat, P., and Zhang, C.-Q. (2011). Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences*, 108(21):8605–8610.

Hollander, B. B. A. (2015). Out or Tuning From 1998 To 2006. pages 23–40.

Immorlica, N., Kleinberg, R., Lucier, B., and Zadimoghaddam, M. (2015). Exponential Segregation in a Two-Dimensional Schelling Model with Tolerant Individuals. pages 1–20.

Iwanaga, S. and Namatame, A. (2013). Collective behavior in cascade and schelling model. *Procedia Computer Science*, 24:217–226.

Iyengar, S. and Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, 59(1):19–39.

Lanchier, N. and Scarlatos, S. (2013). Fixation in the one-dimensional Axelrod model. *Annals of Applied Probability*, 23(6):2538–2559.

Lazarsfeld, P. and Merton, R. (1954). Friendship As A Social Process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, 18(1):18 – 66.

Marin, A. (2012). Don't mention it: Why people don't share job information, when they do, and why it matters. *Social Networks*, 34(2):181–192.

Mcpherson, M., Smith-lovin, L., and Cook, J. M. (2001). B IRDS OF A F EATHER : Homophily in Social Networks. *Annual Review of Sociology*, 27:415 – 444.

Newman, M. E. (2006). Modularity and Community Structure in Networks. In *Proceedings of the National Academy of Sciences 103.23*, pages 8577 – 8582.

Nikolov, D., Oliveira, D. F. M., Flammini, A., and Menczer, F. (2015a). Measuring Online Social Bubbles. pages 1–22.

Nikolov, D., Oliveira, D. F. M., Flammini, A., and Menczer, F. (2015b). Measuring Online Social Bubbles.

Pancs, R., Vriend, N. J., Mary, Q., and Equilibria, M. N. (2007). Schelling ' s Spatial Proximity Model of Segregation. *Journal of Public Economics*, 91(1-2):1–24.

Rodríguez, A. H. and Moreno, Y. (2010). Effects of mass media action on the Axelrod model with social influence. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 82(1):1–21.

Rosengren, K. E. (1974). Uses and gratifications: A paradigm outlined. In *The uses of mass communications: Current perspectives on gratifications research, 3*, pages 269–286.

Sasahara, K., Chen, W., Peng, H., Ciampaglia, G. L., Flammini, A., and Menczer, F. (2019). On the Inevitability of Online Echo Chambers. pages 1–23.

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

Sîrbu, A., Pedreschi, D., Giannotti, F., and Kertész, J. (2019). Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model. *PLoS ONE*, 14(3):1–20.

Stauffer, D. and Solomon, S. (2007). Ising, Schelling and self-organising segregation. *European Physical Journal B*, 57(4):473–479.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Torok, J., Iniguez, G., Yasseri, T., Miguel, M., Kaski, K., and Kertesz, J. (2013). Opinions, Conflicts, and Consensus: Modelling Social Dynamics in a Collaborative Environment. *Physical Review Letters*, 110.

Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., and Nyhan, B. (2018). Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *Ssrn*, (March):1–95.

Vaccari, C., Valeriani, A., Barberá, P., Jost, J. T., Nagler, J., and Tucker, J. A. (2016). Of Echo Chambers and Contrarian Clubs: Exposure to Political Disagreement Among German and Italian Users of Twitter. *Social Media and Society*, 2(3).

Van Alstyne, M. and Brynjolfsson, E. (2005). Global Village or Cyber-Balkans? Modeling and Measuring the Integration of Electronic Communities. *Management Science*, 51(6):851–868.

Vinkovic, D. and Kirman, A. (2006). A physical analogue of the Schelling model. *Proceedings of the National Academy of Sciences*, 103(51):19261–19265.

Yardi, S. and Boyd, D. (2010). Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science, Technology & Society*, 30(5):316–327.

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