Proceedings Paper:
Barmpalias, G, Elwes, RH and Lewis-Pye, A (2014) Digital Morphogenesis via Schelling Segregation. In: IEEE Annual Symposium on Foundations of Computer Science. 2014 Annual Symposium on Foundations of Computer Science, 18-21 Oct 2014, Philadelphia. IEEE Computer Society Press , 156 - 165.

https://doi.org/10.1109/FOCS.2014.25

Digital morphogenesis via Schelling segregation

George Barmpalias*†, Richard Elwes‡ and Andy Lewis-Pye§
*State Key Lab of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing,
Email: barmpalias@gmail.com
†School of Mathematics, Statistics and Operations Research, Victoria University, Wellington, New Zealand.
‡Department of Mathematics, University of Leeds, UK.
Email: r.elwes@gmail.com
§Department of Mathematics, London School of Economics, UK.
Email: andy@aemlewis.co.uk

***Abstract*—Schelling's model of segregation looks to explain the way in which particles or agents of two types may come to arrange themselves spatially into configurations consisting of large homogeneous clusters, i.e. connected regions consisting of only one type. As one of the earliest agent based models studied by economists and perhaps the most famous model of self-organising behaviour, it also has direct links to areas at the interface between computer science and statistical mechanics, such as the Ising model and the study of contagion and cascading phenomena in networks.**

While the model has been extensively studied it has largely resisted rigorous analysis, prior results from the literature generally pertaining to variants of the model which are tweaked so as to be amenable to standard techniques from statistical mechanics or stochastic evolutionary game theory. In [2], Brandt, Immorlica, Kamath and Kleinberg provided the first rigorous analysis of the unperturbed model, for a specific set of input parameters. Here we provide a rigorous analysis of the model's behaviour much more generally and establish some surprising forms of threshold behaviour, notably the existence of situations where an *increased* level of intolerance for neighbouring agents of opposite type leads almost certainly to *decreased* segregation.

***Keywords*-Schelling segregation, Ising, spin glass, networks, morphogenesis, algorithmic game theory.**

I. Introduction

While Alan Turing is best known within the mathematical logic and computer science communities for his work formalising the algorithmically calculable functions, it is interesting to note that his most cited work [9] is actually that relating to morphogenesis. Turing wanted to understand certain biological processes: the gastrulation phase of embryonic development, the process whereby dappling effects arise on animal coats, and phyllotaxy, i.e. the arrangement of leaves on plant stems. One can consider the more general question, however, as to how morphogenesis occurs – how structure can arise from an initially random, or near random configuration. Along these lines, one of the major contributions of the economist and game theorist Thomas Schelling was an elegant model of segregation, first described in 1969 [7], which turns out to provide a very simple model of such a morphogenic process. This model looks to describe how individuals of different types come to organise themselves spatially into segregated regions, each of largely one type. Today it has become perhaps the best known model of self-organising behaviour, and was one of the reasons cited by the Nobel prize committee upon awarding Schelling his prize in 2005. Although the explicit aim was initially to model the kind of racial segregation observed in large American cities, Schelling himself pointed out that the analysis is sufficiently abstract that any situation in which objects of two types arrange themselves geographically according to a certain preference not to be of a minority type in their neighbourhood, could constitute an interpretation. As described in [10], for example, Schelling's model can be seen as a finite difference version of differential equations describing interparticle forces and applied in modelling cluster formation. Many authors (see for example [3], [4], [8]) have pointed out direct links to spin-1 models used to analyse phase transitions – by introducing noise into the dynamics of the underlying Markov process one can arrive at the Boltzmann distribution for the set of possible configurations, with the 'energy' typically corresponding to some measure of the mixing of types. From there one can immediately deduce that the modified (now ergodic) process spends a large proportion of the time in completely segregated states, with this proportion tending to 1 as the analogue of the temperature is taken to 0. So had Schelling been aware of these connections to variants of the Ising model, he could have based his work on a long history of physics research.

Our own avenue into these questions, however, came via the work of computer scientists [2] and the study of cascading phenomena on networks, a good introduction to which can be found in [6]. The dynamics of the Schelling process (as will be clear once it has been formally defined below) are almost identical to many of those used to model the flow of information or behaviour on large social or physical networks such as the internet, the principal differences here being the initial conditions considered and the use of a much simpler underlying graph structure. An immediate concern, given the results of this paper, is as to whether techniques developed here can be extended and applied to understand emerging clustering phenomena on the various (normally more complex) random graph structures studied by those in the networks community. Along these lines, Henry, Prałat and Zhang have described a simple but elegant model of network clustering [5], inspired by the Schelling model. Their model doesn't display the kind of involved threshold behaviour, however, that one might expect to be exhibited by a direct translation of the Schelling process to random underlying graph structures.

We concentrate here on the one-dimensional version of the model, as in [2]. The model works as follows. One begins with a large number n of nodes (individuals) arranged in a circle. Each node is initially assigned a *type*, and has probability $\frac{1}{2}$ of being of type α and probability $\frac{1}{2}$ of being of type β (the types of distinct individuals being independently distributed). We fix a parameter w, which specifies the 'neighbourhood' of each node in the following way: at each point in time the neighbourhood of the node u, denoted $\mathcal{N}(u)$, is the set containing u and the w-many closest neighbours on each side – so the neighbourhood consists of $2w + 1$ nodes in total. The second parameter $\tau \in [0, 1]$ specifies the proportion of a node's neighbourhood which must be of their type for them to be happy. So, at any given moment in time, we define u to be *happy* if at least $\tau(2w + 1)$ of the nodes in $\mathcal{N}(u)$ are of the same type as u. One then considers a discrete time process, in which, at each stage, one pair of unhappy individuals of opposite types are selected uniformly at random and are given the opportunity to swap locations. We work according to the assumption that the swap will take place as long as each member of the pair has at least as many neighbours of the same type at their new location as at their former one (note that for $\tau \leq \frac{1}{2}$ this will automatically be the case). The process ends when (and if) one reaches a stage at which no further swaps are possible.

Our contribution

Much of the difficulty in providing a rigorous analysis stems from the large variety of absorbing states for the underlying Markov process. Many authors have therefore worked with variants of the model in which perturbations are introduced into the dynamics so as to avoid this problem. In [2], Brandt, Immorlica, Kamath and Kleinberg used an analysis of locally defined stable configurations, combined with results of Wormald [11], to provide the first rigorous analysis of the unperturbed one-dimensional Schelling model, for the case $\tau = \frac{1}{2}$. In this paper we shall consider what happens more generally for $\tau \in [0, 1]$ (for the unperturbed model), and we shall observe, in particular, that some remarkable threshold behaviour occurs. While some aspects of the approach from [2] remain, in particular the focus on locally defined stable configurations which can be used to understand the global picture, the details of the methods of their proof (the use of 'firewall incubators', and so on) are entirely specific to the case $\tau = \frac{1}{2}$, and so largely speaking we shall require different techniques here. The picture which emerges is one in which one observes different behaviour in five regions. For κ which is the unique solution in $[0, 1]$ to

$$(1/2 - \kappa)^{1-2\kappa} = (1 - \kappa)^{2-2\kappa}\,, \tag{1}$$

($\kappa \approx 0.353092313$) these regions are: (i) $\tau < \kappa$, (ii) $\tau = \kappa$, (iii) $\kappa < \tau < \frac{1}{2}$, (iv) $\tau = \frac{1}{2}$, (v) $\tau > \frac{1}{2}$. In fact we shall not consider the case $\tau = \kappa$, but the behaviour for all other values of τ is given by the theorems below. Perhaps the most surprising fact is that, in some cases, increasing τ almost certainly leads to decreased segregation. The assumption is always that we work with $n \gg w$, i.e. all results hold for all n which are sufficiently large compared to w. A *run* of length d is a set of d-many consecutive nodes all of the same type. *Complete segregation* refers to any configuration in which there exists a single run to which all α nodes belong.

The first theorem deals with low values of τ, and formally establishes the (perhaps rather intuitive) idea that very low levels of intolerance lead to low levels of segregation:

Theorem 1.1: Suppose $\tau < \kappa$ and $\epsilon > 0$. For all sufficiently large w, if a node u is chosen uniformly at random, then the probability that any node in $\mathcal{N}(u)$ is ever involved in a swap is $< \epsilon$. Thus there exists a constant d such that, for sufficiently large w, the probability u belongs to a run of length $> d$ in the final configuration is $< \epsilon$.

As one increases τ beyond the threshold κ, however, the dynamics of the process qualitatively change:

Theorem 1.2: Suppose $\tau \in \left(\kappa, \frac{1}{2}\right)$ and $\epsilon > 0$. There exists a constant d such that (for all w and all n such that $w \ll n$) the probability that u chosen uniformly at random will belong to a run of length $\geq e^{w/d}$ in the final configuration, is greater than $1 - \epsilon$.

So, to summarise Theorem 1.2 less formally, increasing τ beyond κ suddenly causes high levels of segregation, in the form of run-lengths which are exponential in w. Furthermore, by analysing the proof of Theorem 1.2, we shall be able to prove (a formalised version of the statement) that increasing τ in this interval actually decreases run-lengths, i.e. increasing τ increases the required value of d. The next case is that dealt with in [2], where one sees polynomially bounded run-lengths:

Theorem 1.3 ([2]): Suppose $\tau = \frac{1}{2}$. There exists a constant $c < 1$ such that for all $\lambda > 0$, the probability that u chosen uniformly at random will belong to a run of length greater than λw^2 in the final configuration, is bounded above by c^λ.

The final case is when $\tau > \frac{1}{2}$. Here a combinatorial argument can be used to argue that complete segregation will eventually occur. Note that, if $\frac{1}{2} < \tau \leq \frac{w+1}{2w+1}$, then the process is identical to that for $\tau = \frac{1}{2}$, since in both cases a node requires $w + 1$ many nodes of its own type in its neighbourhood in order to be happy.

Theorem 1.4: Suppose that $\tau > \frac{1}{2}$, and that w is sufficiently large that $\tau > \frac{w+1}{2w+1}$. Then, with probability tending to 1 as $n \to \infty$, the initial configuration is such that complete segregation is inevitable.[1]

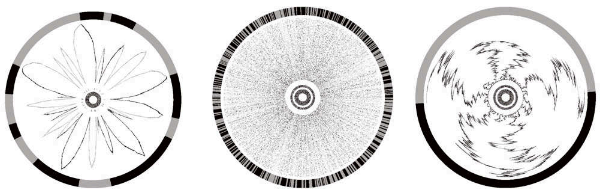

Figure 1: Processes for $\tau = 0.38, \tau = 0.5, \tau = 0.7$.

The outcomes of some simulations are illustrated in Figure 1. In the processes depicted here the number of nodes $n = 100000$, $w = 60$ and in the diagrams individuals of type α are coloured light grey and individuals of type β are coloured black. The inner ring displays the initial mixed configuration (in fact the configuration is sufficiently mixed that changes of type are not really visible, so that the inner ring appears dark grey). The outer ring displays the final configuration. Just immediately exterior to the innermost ring is another, which displays individuals which are unhappy in the initial configuration. The process by which the final configuration is reached is indicated in the space between the inner rings and the outer ring in the following way: when an individual changes type this is indicated with a mark, at a distance from the inner rings which is proportional to the time at which the change of type takes place. In fact, for the case $\tau > \frac{1}{2}$ one has to be a little careful in talking about the 'final' configuration – there will, almost certainly, always be unhappy individuals of both types able to swap, but once a completely segregated configuration is reached all future configurations must remain completely segregated.

II. Notation and terminology

In describing the model earlier, we spoke in terms of nodes or individuals swapping locations at various stages of the dynamic process. In fact, it is notationally easier to consider a process whereby one simply has a set of n nodes, with two unhappy nodes of opposite type selected at each stage (if such exist), which may then both change type (when this occurs we shall still refer to the nodes as 'swapping', but now they are swapping type rather than location). Thus nodes are identified with indices for their locations amongst the set $\{0, 1, ..., n-1\}$, and unless stated otherwise, addition and subtraction on these indices are performed modulo n. In the context of discussing a node u_1, for example, we might refer to the immediate neighbour on the right as node u_1+1.[2] As noted before, for any node u, we let $\mathcal{N}(u)$ denote the neighbourhood of u, which is the interval $[u-w, u+w]$. For any set of nodes I, suppose that x is the number of α nodes in I, while $y = |I| - x$. Then $\Theta(I) := x - y$ and is called the *bias* of I. By the bias of a node we mean the bias of its neighbourhood. Recall that by a *run* of length $m+1$ we mean an interval $[u, u+m]$ in which all nodes are of the same type. We shall be particularly interested in local configurations which are stable, in the sense that certain nodes in them can never be caused to change type. Note that if an interval of length $w+1$ contains at least $\tau(2w+1)$ many α nodes, then each of those α nodes is happy so long as the others do not change type, meaning that, in fact, no α nodes in that interval will ever change type. We say that such an interval of length $w+1$ is α-stable (and similarly for β). An interval of length $w+1$ is *stable* if it is either α-stable or β-stable. We shall also make use of a particular kind of stable interval which was used in [2]: a *firewall* is a run of length at least $w+1$. We write 'for $0 \ll w \ll n$', to mean 'for all sufficiently large w and all n sufficiently large compared to w'. We define the *harmony index* corresponding to any given configuration to be the sum over all nodes of the number of their own type within their neighbourhood. For $\tau \leq \frac{1}{2}$, this harmony index is easily seen to strictly increase whenever an unhappy node changes type, which combined with the existence of an upper bound $n(2w+1)$, implies that the process must terminate after finitely many stages. For $\tau > \frac{1}{2}$, we shall argue that with probability tending to 1 as $n \to \infty$, the initial configuration is such that complete segregation eventually occurs with probability 1. Once complete segregation has occurred it is easy to see that all future states must be completely segregated.

III. The case $\tau < \kappa$

The case $\tau < \kappa$ is the easiest case, and we only sketch the argument here. The proof in its entirety can be found in the full version of this paper [1]. The basic idea is that we wish to find the value of τ at which stable intervals become more likely than unhappy nodes in the initial configuration. For such τ, taking w large, we shall have that stable intervals are *much* more likely than unhappy nodes in the initial configuration. If u is selected uniformly at random then we shall very likely find stable intervals of both types on either side of u before any unhappy element, meaning that u never changes type. The following lemma is the first step towards formalising this idea:

Lemma 3.1: Consider the initial configuration. Let κ be as specified in 1 and let P be any polynomial. If $\tau < \kappa$ and u is a node chosen uniformly at random, then for $0 \ll w \ll n$,

[1] Note that, while Theorem 1.4 requires $\tau > \frac{w+1}{2w+1}$, Theorem 1.2 does not. Briefly, this is because the choice of d can be made to so as to deal with finitely many small w anyway, meaning that Theorem 1.2 is essentially a statement about what happens for large w.

[2] Since we work modulo n it is worth clarifying some details of the interval notation: for $0 \leq b < a < n$, we let $[a, b]$ denote the set of nodes ('interval') $[a, n-1] \cup [0, b]$ (while $[b, a]$ is, of course, understood in the standard way).

the probability u belongs to a stable interval is more than $P(w)$ times the probability that u is unhappy. If $\tau > \kappa$ then the reverse is true, i.e. the probability u is unhappy is more than $P(w)$ times the probability that u belongs to a stable interval.

We need something more than this though. For $\gamma \in \{\alpha, \beta\}$, let γ^* be the opposite type, i.e. $\gamma^* = \alpha$ if $\gamma = \beta$, and $\gamma^* = \beta$ otherwise.

Lemma 3.2: Suppose $\tau < \kappa$ and consider the initial configuration. For any $\epsilon > 0$, for $0 \ll w \ll n$, if u_0 is selected uniformly at random, then with probability $> 1 - \epsilon$, there exist $u_{-2} < u_{-1} < u_0 < u_1 < u_2$ such that:

- There are no unhappy nodes in the interval $[u_{-2}, u_2]$.
- For $i \in [-2, 2] - \{0\}$, u_i belongs to a γ_i-stable interval, I_i say. Also $\gamma_{-2} = \gamma^*_{-1}$ and $\gamma_2 = \gamma^*_1$.

To complete the argument, we now show that the existence of stable intervals of opposite types on both sides before any unhappy nodes, suffices to prevent type changes, i.e. that Lemma 3.2 gives the desired result. So suppose otherwise, and let v be the first node in the interval $[u_{-2}, u_2]$ to become unhappy. Without loss of generality suppose v is an α node. In order for v to become unhappy, another α node $v' \in \mathcal{N}(v)$ must change to type β. Since $v' \notin [u_{-2}, u_2]$, we either have $v' < u_{-2} \leq v$, or else $v \leq u_2 < v'$. Suppose that the first case holds, the other is similar. Then, together with the fact that any α-stable interval to which u_{-2} belongs is of length $w+1$, $v' \in \mathcal{N}(v)$ implies that v belongs to any stable interval to which u_{-2} belongs, and so cannot become unhappy. This gives the required contradiction.

IV. The case $\kappa < \tau < \frac{1}{2}$

In what follows we shall work with some fixed τ in the interval $\left(\kappa, \frac{1}{2}\right)$, some fixed $\epsilon > 0$, and we shall assume that n is large compared to w. Again, we shall only sketch the proof and most of the necessary lemmas will simply be stated without proof. The full proof can be found in [1]. We want to show that there exists a constant d such that for all sufficiently large w the probability that a randomly chosen node will belong to a run of length $\geq e^{w/d}$ (in the final configuration) is greater than $1 - \epsilon$. Of course, proving the result for all sufficiently large w suffices to give the result for all w since one can simply adjust the choice of d to deal with finitely many small values, but we shall make frequent use of the fact that we need only work for all sufficiently large w in what follows and so rephrasing the theorem in this way is instructive.

A difficulty that arises in this context is that when there are different numbers of unhappy α and β nodes, it fails to be the case that every unhappy node is equally likely to be chosen as part of a swapping pair. If there are more unhappy α nodes than unhappy β nodes at a given stage, for example, then unhappy β nodes belong to more unhappy pairs of opposite type than do their unhappy α counterparts, and so are more likely to be chosen as part of an unhappy pair. Applying technology developed by Wormald [11], however, one can show that the discrete Schelling process can be sufficiently accurately modelled by a continuous one which is governed by a system of differential equations. From there one can demonstrate that the number of unhappy nodes of each type actually remains very evenly balanced, at least until a suitably late stage of the process. In what follows here, we shall subdue mention of such issues and simply assume that each unhappy node has an equal chance of being chosen to swap at each stage. We refer the reader who wishes to see a careful treatment of these complications to the full version of the paper [1].

Our entire analysis takes place relative to a node u_0, chosen uniformly at random. Roughly, the aim is to establish that in the final configuration u_0 very probably belongs to a firewall of considerable length. The argument consists of two main parts: first we consider what can be expected from the vicinity of u_0 in the initial configuration, and then we consider how events are likely to develop in subsequent stages.

Before we begin with the technicalities, let us consider very informally what can be expected from the vicinity of u_0 in the initial configuration. Since $\kappa < \tau < \frac{1}{2}$, for large w we shall have that unhappy nodes are much more likely than stable intervals, but that unhappy nodes themselves are few and far between. Starting at u_0 and moving to the left, (since n is large) we can expect to find a first unhappy node, l_1 say, and it will very likely be the case that $[l_1, u_0]$ is an interval of considerable length, containing no stable intervals. To the left of l_1 and inside $\mathcal{N}(l_1)$, there may be some other unhappy nodes. If we move now to $l_1 - (2w + 1)$, however, and repeat the process (with $l_1 - (2w + 1)$ taking the place of u_0), then so long as w is large enough, we can expect the same to happen again, i.e. we find a first unhappy node l_2 and $[l_2, l_1 - (2w + 1)]$ is an interval of considerable length, containing no stable intervals. In fact, for any fixed k which does not depend on w, if we repeat this process k many times, then so long as w is large enough (and how large we have to take w will depend on k) we can be pretty sure that the same thing will happen at every one of those k-many steps. Now let us establish this informal picture more carefully.

The initial configuration

Recall that for any set of nodes I, $\Theta(I)$ is the *bias* of I and that by the bias of a node we mean the bias of its neighbourhood. In general, if $x_1, ..., x_k$ are independent random variables with $\mathbf{P}(x_i = 1) = \mathbf{P}(x_i = -1) = \frac{1}{2}$ when $1 \leq i \leq k$, then letting $X = \sum_{i=1}^{k} x_i$ Hoeffding's inequality gives, for arbitrary $\lambda > 0$:

$$\mathbf{P}(|X| > \lambda\sqrt{k}) < 2e^{-\lambda^2/2}.$$

Now we use this to bound the probability that a node u has bias in the initial configuration which will cause it to be unhappy, should u be of the minority type in its neighbourhood. So, we wish to bound the probability that the number of α nodes in $\mathcal{N}(u)$ is $> (1-\tau)(2w+1)$ or the number of β nodes in $\mathcal{N}(u)$ is $> (1-\tau)(2w+1)$. This corresponds to a bias $\Theta(\mathcal{N}(u))$ of $> (1-2\tau)(2w+1)$ or $< -(1-2\tau)(2w+1)$.

Definition 4.1: When $|\Theta(\mathcal{N}(u))| > (1-2\tau)(2w+1)$ we say that u has **high bias**, denoted $\mathtt{Hb}(u)$. If this holds for the initial configuration, we say that $\mathtt{Hb}^*(u)$ holds.

Definition 4.2: For the remainder of this section we define $d = 2/(1-2\tau)^2$.

The following lemma then follows from a direct application of Hoeffding's inequality.

Lemma 4.3 (Likely happiness): Let u be a node chosen uniformly at random. For any $\epsilon' > 0$ and for all sufficiently large w, the probability that $\mathtt{Hb}^*(u)$ holds is $< \epsilon' e^{-w/d}$.

Defining the nodes l_i and r_i. For now, we fix some $k_0 > 0$. We shall choose a specific value of k_0 which is appropriate later – for now, however, we make the promise that our choice of k_0 will not depend on w.

For $1 \le i \le k_0$ we define a node l_i to the left of u_0 and also a node r_i to the right. We let l_1 be the first node v to the left of u_0 such that $\mathtt{Hb}^*(v)$ holds, so long as this node is in the interval $[u_0 - \frac{1}{4}n, u_0]$ (otherwise l_1 is undefined). Then, given l_i for $i < k_0$ we let l_{i+1} be the first node v to the left of $l_i - (2w+1)$ such that $\mathtt{Hb}^*(v)$ holds, so long as no nodes in the interval $[l_{i+1}, l_i]$ are outside the interval $[u_0 - \frac{1}{4}n, u_0]$ (otherwise l_{i+1} is undefined). We let r_1 be the first node v to the right of u_0 such that $\mathtt{Hb}^*(v)$ holds, so long as this node is in the interval $[u_0, u_0 + \frac{1}{4}n]$. Given r_i for $i < k_0$ we let r_{i+1} be the first node v to the right of $r_i + (2w+1)$ such that $\mathtt{Hb}^*(v)$ holds, so long as no nodes in the interval $[r_i, r_{i+1}]$ are outside the interval $[u_0, u_0 + \frac{1}{4}n]$.

The reason for considering the intervals $[u_0 - \frac{1}{4}n, u_0]$ and $[u_0, u_0 + \frac{1}{4}n]$ in the above, is that we wish to be able to move left from u_0 to l_{k_0} without meeting any of the nodes r_i.

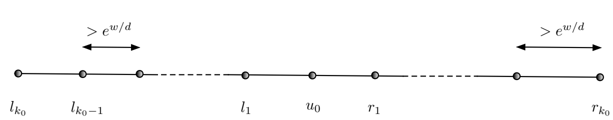

Figure 2: Picking out nodes of high bias near u_0.

*Definition 4.4 (**Good spacing**):* Let d be as in Definition 4.2. It is notationally convenient to let $l_0 = u_0 = r_0$. We let **Good spacing** be the event that for $1 \le i \le k_0$:

(i) l_i and r_i are both defined, and;
(ii) $|l_i - l_{i-1}| > e^{w/d}$ and $|r_i - r_{i-1}| > e^{w/d}$.

Note that for any fixed w, as $n \to \infty$ the probability that any l_i or r_i is undefined (for $i \le k_0$) goes to 0. By Lemma 4.3, and since the probability that any node in an interval I has high bias in the initial configuration is at most $\Sigma_{u \in I} \mathbf{P}(\mathtt{Hb}^*(u))$, for any $\epsilon' > 0$ and for any fixed $k_0 \ge 1$ we can ensure that $\mathbf{P}(\textbf{Good spacing}) > 1 - \epsilon'$ by taking w sufficiently large (and by taking n sufficiently large compared to w). Thus, for $0 \ll w \ll n$, the picture we are presented with is almost certainly as in Figure 2.

Building the informal picture

Recall that, by a *run* of length $m+1$ we mean an interval $[u, u+m]$ in which all nodes are of the same type, and that a *firewall* is a run of length at least $w+1$. The basic observation on which we now wish to build is as follows: if the interval $[u-w, u]$ is a firewall of type α, then when $u+1$ is of type β, it cannot be happy unless the interval $[u+1, u+w+1]$ is β-stable. So, since we are dealing with $\tau \le \frac{1}{2}$, firewalls will spread until they hit stable intervals of the opposite type (so long as there exist unhappy nodes of both types).

With this in mind, let us now consider informally what can be expected to happen in the neighbourhood of l_i. Suppose that l_i is initially of type β and is unhappy in the initial configuration. Then with probability close to 1 for sufficiently large w, there will not be any unhappy nodes of type α in the neighbourhood of l_i in the initial configuration. If l_i changes type, then this will make the bias in its neighbourhood still more positive, which may cause further nodes of type β to become unhappy. If these change to type α then this will further increase the bias, potentially causing more nodes to become unhappy, and so on. The following definitions formalise some of the ways in which this process might play out, and in particular the possibility that this process might play out without interference from what happens in other neighbourhoods $\mathcal{N}(l_j)$ or $\mathcal{N}(r_j)$.

Definition 4.5: For $0 < i < k_0$ we say that l_i **completes at stage** s if both:

1) No node in $\mathcal{N}(l_i)$ is unhappy at stage s, and this is not true for any $s' < s$.
2) There exist x_0 and x_1 with $l_{i+1} + 2w < x_0 < l_i - 2w < l_i + 2w < x_1 < l_{i-1} - 2w$, such that by the end of stage s, no node in $[x_0 - w, x_0]$ or $[x_1, x_1 + w]$ has changed type.

We say that l_i **completes** if it completes at some stage. We also define completion for r_i analogously.

Definition 4.6: We say that l_i (or r_i) **originates a firewall** if it completes at some stage s and (i) it belongs to firewall at stage s, and (ii) all type changes in $\mathcal{N}(l_i)$ (or $\mathcal{N}(r_i)$) at stages $\le s$ are of the same kind (i.e. all α to β, or all β to α).

The informal idea, is that we now wish to show that each l_i and each r_i has some reasonable chance of originating a firewall (and that this reasonable chance is bounded below by some value which doesn't depend on w). Then we can choose k_0 so that the probability none of the l_i originate a firewall or none of the r_i originate a firewall is $\ll \epsilon$, i.e. with probability close to 1 firewalls will originate either side of u_0 within the interval $[l_{k_0}, r_{k_0}]$. Then, letting i_1 be the least i such that l_i originates a firewall, and letting i_2 be the least i such that r_i originates a firewall, we wish to show that with probability close to 1 the firewalls originated at l_{i_1} and r_{i_2} will spread until u_0 is contained in one of them. Since these two firewalls have originated at nodes which are at distance at least $e^{w/d}$ apart, u_0 ultimately belongs to a firewall of at least this length. So to sum up:

> The approximate reason Theorem 1.2 holds is that u_0 can be expected to join a firewall which – *precisely because* unhappy nodes are rare in the initial configuration – originated at a long distance from u_0.

In order to make this basic picture work, however, we need to be careful about the formation of stable intervals in $[l_{k_0}, r_{k_0}]$. As noted above, firewalls will spread until they hit stable intervals of the opposite type. Now suppose that, with i_1 and i_2 as above, $i_1 = i_2 = 2$ and, for now, suppose that α-firewalls are originated at both l_2 and r_2. In order to show that these two firewalls will spread until they meet each other, it will be helpful first of all, to be able to assume that in the initial configuration there are no stable subintervals of $[l_{k_0}, r_{k_0}]$. This will follow quite easily for large w, from our previous analysis of the ratio between the probability of unhappy nodes and stable intervals. A further danger that we have to be able to avoid, however, is that, while l_1 and r_1 do not originate firewalls, they *do* get as far as creating β-stable intervals.

Definition 4.7: Given i with $0 < i < k_0$, let $u = l_i$ or $u = r_i$. Let $u_1 = u - (2w + 1)$ and $u_2 = u + (2w + 1)$. We say that u **subsides** if it completes at some stage s, and:

- There are no nodes in $[u_1, u_2]$ belonging to stable intervals at stage s;
- No nodes in $\mathcal{N}(u_1)$ or $\mathcal{N}(u_2)$ have been unhappy at any stage $\leq s$.

So we need to be able to show, in fact, that with probability close to 1 each l_i and r_i either originates a firewall or subsides. To do this clearly involves a careful analysis what is likely to happen in each of the neighbourhoods $\mathcal{N}(l_i)$ and $\mathcal{N}(r_i)$. First of all, the large distances between these nodes mean that, for fixed k_0 and sufficiently large w, we can expect all of the l_i and r_i (for $0 < i < k_0$) to complete, so that one can understand the early stages of the process for each of these neighbourhoods by considering each in isolation. We then wish to show:

> *The required dichotomy:* in the neighbourhood of each l_i and r_i, either a small number of type changes will occur before completion, or else a large number of type changes will occur before completion and a firewall will be created.

Now, if we strengthen our original requirement that there are no stable subintervals of $[l_{k_0}, r_{k_0}]$ in the initial configuration, to a requirement that there are no subintervals which are 'close' to being stable in the initial configuration (where 'close' is to be made precise in such a way as to ensure that when a small number of type changes occur in the neighbourhood of l_i before completion, these are not enough to create any stable intervals), then we shall have that with probability close to 1 each l_i and each r_i either originates a firewall or else completes without creating stable intervals.

Once all this is in place, there is then one further hurdle. In the above, we assumed that the firewalls originating at l_2 and r_2 are both α-firewalls. If they are firewalls of opposite type, however, we still have some work to do in order to prove that u_0 will almost certainly end up belonging to one of these two firewalls.

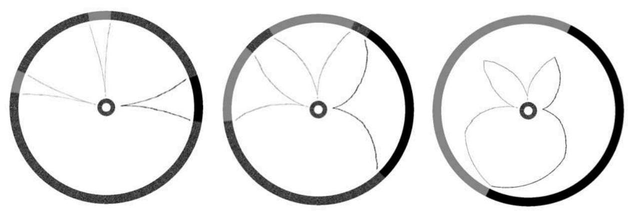

Figure 3: An example of the developing process with $n = 10^5, w = 75, \tau = 0.37$, at stages $5 \cdot 10^3, 15 \cdot 10^3$ and 24986.

Formalising the intuitive picture

Thus far we have defined an event, **Good spacing**, which depends upon the value k_0. We are yet to specify k_0, but have promised that this choice will not depend on w. For any $\epsilon' > 0$ and for any fixed $k_0 \geq 1$ we can ensure that $\mathbf{P}$(**Good spacing**) $> 1 - \epsilon'$ by taking w sufficiently large (and by taking n sufficiently large compared to w).

In order to specify how the type changing process can be expected to develop in the vicinity of u_0, we shall now proceed to define a finite number of other events of this kind. This finite set of events (having **Good spacing** as a member) we call shall call Π. Our aim is to show that, for any $\epsilon' > 0$, the probability that all the events in Π occur is greater than $1 - \epsilon'$ for all sufficiently large w (and for n sufficiently large compared to w). Suppose that this is established for some $\Pi' \subset \Pi$. To establish the result for $\Pi' \cup \{Q\}$ with $Q \in \Pi - \Pi'$, it then suffices to prove for each $\epsilon' > 0$ and all sufficiently large w, that the probability of Q given P is greater than $1 - \epsilon'$, where P is any conjunction (possibly empty) of the events in Π'. Of course, we choose that P which is most convenient to work with.

In discussing the 'required dichotomy' above, a requirement was suggested, that there should be no subintervals of $[l_{k_0}, r_{k_0}]$ which are 'close' to being stable in the initial configuration. In fact an appropriate formalisation of this idea is easy to describe, and we now do so.

Definition 4.8: For any $\tau' \in (0,1)$, we say that an interval of length $w+1$ is **τ'-stable**, if it contains $\tau'(2w+1)$ many α-nodes or $\tau'(2w+1)$ many β nodes.

*Definition 4.9 (**Stable clear**):* Once and for all, fix some τ_0 with $\kappa < \tau_0 < \tau$. Let **Stable clear** $\in \Pi$ be the event that l_i and r_i are defined for all $1 \le i \le k_0$, and there do not exist any τ_0-stable subintervals of $[l_{k_0}, r_{k_0}]$ in the initial configuration.

Lemma 4.10: For fixed k_0 and $\epsilon' > 0$, $\mathbf{P}$(**Stable clear**) $> 1-\epsilon'$ for all w sufficiently large (and all n sufficiently large compared to w).

Now, in order to establish the required dichotomy, we need to build up a clear picture of what the neighbourhoods $\mathcal{N}(l_i)$ and $\mathcal{N}(r_i)$ can be expected to look like. The distributions for these intervals in the initial configuration are a little difficult to attack directly, however, due to the nature of their definition. In choosing l_1 we move left until we find the *first* node which has high bias – this gives an asymmetry to the given information concerning the neighbourhood. Roughly, we might expect something like a hypergeometric distribution, but how good is this as an approximation? What we shall do, in fact, is first of all to understand what can be expected from the neighbourhood of a node which is chosen uniformly at random from among those with borderline bias:

Definition 4.11: Let us say that a node u has **borderline bias**, denoted $\texttt{Bb}(u)$, if:

$$|\Theta(\mathcal{N}(u))| = \text{Min}\{\theta \in 2\mathbb{N}+1 : \theta > (1-2\tau)(2w+1)\},$$

i.e., u has high bias but decreasing the modulus of the bias by the minimum possible amount of 2 would cause it not to have high bias. We say that $\texttt{Bb}^*(u)$ holds if u has borderline bias in the initial configuration. Note that each of the nodes l_i and r_i $(0 < i \le k_0)$ has borderline bias.

In what follows it is often convenient to work with some fixed $k \ge 1$ and to divide an interval $I = [a,b]$ into k parts of equal length. This occasions the minor inconvenience that the length of the interval might not be a multiple of k, motivating the following definition:

Definition 4.12: Let $I = [a,b]$ and suppose $k \ge 1$. We define the subintervals: $I(1:k) := [a, a+\lfloor (b-a)/k \rfloor] := [I(1:k)_1, I(1:k)_2]$ and $I(j:k) := [a+\lfloor (j-1)(b-a)/k \rfloor + 1, a+\lfloor j(b-a)/k \rfloor] := [I(j:k)_1, I(j:k)_2]$ for $2 \le j \le k$.

In Definition 4.12 the intervals are counted from left to right, but it is also useful to work from right to left:

Definition 4.13: Let $I = [a,b]$ and suppose $k \ge 1$. For $1 \le j \le k$ we define $I(j:k)^- = I(k-j+1:k)$, $I(j:k)^-_1 = I(k-j+1:k)_1$ and $I(j:k)^-_2 = I(k-j+1:k)_2$.

Lemma 4.14 (Smoothness Lemma): Suppose u is such that the proportion of α nodes in $I := \mathcal{N}(u)$ is θ, and that u is selected uniformly at random from nodes with this property. Then for any fixed $k \ge 1$ and $\epsilon' > 0$, for all sufficiently large w the following holds with probability $> 1-\epsilon'$: for every j with $1 \le j \le k$, the proportion of the nodes in $I(j:k)$ which are α nodes, lies in the interval $[\theta-\epsilon', \theta+\epsilon']$.

Lemma 4.14 basically tells us that if we choose a node u with borderline bias uniformly at random, then for large w we can expect the bias to move towards 0 fairly smoothly as we move to $u+(2w+1)$ or $u-(2w+1)$. In order to see roughly why this is true, suppose that $|\Theta(\mathcal{N}(u))| = \rho$ and let θ be the proportion of the nodes in $\mathcal{N}(u)$ which are of type α. Let $I = [u, u+(2w+1)]$ and, for some k, consider the sequence of evenly spaced nodes $v_j = I(j,k)_2$. Now in forming the neighbourhood of v_j, we lose an interval of length (almost exactly) $\lfloor j(2w+1)/k \rfloor$ from the neighbourhood of u, which by Lemma 4.14 we can expect to have a proportion of α nodes very close to θ. We also gain an interval of the same length from outside $\mathcal{N}(u)$, which we can expect to have a proportion of α nodes very close to $\frac{1}{2}$. This means a bias for v_j close to $\rho\frac{k-j}{k}$. The following definition allows us to express this more formally:

Definition 4.15: Suppose that $\texttt{Hb}^*(u)$ holds. Let $I_1 = [u-(2w+1), u]$ and $I_2 = [u, u+(2w+1)]$. Let $|\Theta(\mathcal{N}(u))| = \rho$ and let θ be the proportion of the nodes in $\mathcal{N}(u)$ which are of type α. Suppose that $k \ge 1$ is even and $\epsilon' > 0$. For $1 \le j \le k$ let $v_j = I_2(j:k)_2$ and let $v_{-j} = I_1(j:k)^-_1$. We say that $\texttt{Smooth}_{k,\epsilon'}(u)$ holds if both:

- For every j, $1 \le |j| \le k$, $|\Theta(\mathcal{N}(v_j)) - \rho\frac{k-j}{k}|/w < \epsilon'$.
- For every j with $1 \le j \le k/2$ the proportion of the nodes in $I_1(j:k)^-$ which are of type α lies in the interval $[\theta-\epsilon', \theta+\epsilon']$, and similarly for $I_2(j,k)$.

We say that $\texttt{Smooth}^*_{k,\epsilon'}(u)$ holds if $\texttt{Smooth}_{k,\epsilon'}(u)$ holds in the initial configuration.

Corollary 4.16 (Smoothness Corollary): Suppose u is selected uniformly at random among those nodes such that $\texttt{Bb}^*(u)$ holds. For all $k \ge 1$ and $\epsilon' > 0$, and for all sufficiently large w, $\texttt{Smooth}^*_{k,\epsilon'}(u)$ holds with probability $> 1-\epsilon'$.

While Smoothness Corollary 4.16 tells us what can be expected from the neighbourhood of a node chosen uniformly at random from among those with borderline bias, this does not immediately allow us to infer anything about what can be expected from the neighbourhood of each l_i and r_i. What we need is that, if we choose a node u uniformly at random and then move left (or right) until we find a first node v with high bias, then with probability close to 1 $\texttt{Smooth}^*_{k,\epsilon'}(v)$ holds. With some work we can establish the following:

Lemma 4.17 (Smoothness for l_i and r_i): For any node u, let x_u be the first node to the left of u which has high bias in the initial configuration. For any $\epsilon' > 0$ and $k \ge 1$, if $0 \ll w \ll n$ and u is chosen uniformly at random, then x_u is defined and $\texttt{Smooth}^*_{k,\epsilon'}(x_u)$ holds with probability $> 1-\epsilon'$. An analogous result holds when 'left' is replaced by 'right'.

In order to see that Lemma 4.17 suffices to establish probable smoothness for all of the l_i and r_i, note first that k_0 is fixed while we take w large. At step i of the iteration which defines the sequence $l_1, l_2, ..$, the fact that l_i has borderline bias tells us nothing about the neighbourhood of $l_i-(2w+1)$ or the nodes to the left of this neighbourhood (but at a distance small compared to n).

We are now ready to define the third event in Π:

*Definition 4.18 (**Smooth**):* Let τ_0 be as in Definition 4.9. Once and for all, choose k_1 such that $\frac{1}{k_1} \ll \tau-\tau_0$, and choose k_2 and ϵ_0 such that $k_1 \ll k_2 \ll \frac{1}{\epsilon_0}$ and k_2 is a multiple of k_1. We define **Smooth** to be the event that all the l_i and r_i are defined for $1 \leq i \leq k_0$, and that when $u=l_i$ or $u=r_i$, $\texttt{Smooth}^*_{k_2,\epsilon_0}(u)$ holds.

By Lemma 4.17, when $k_0 \ll w$ the probability that **Smooth** does not occur is $\ll \epsilon$.

The process to completion

Having established a clearer picture of what can be expected from the initial configuration, we now look to understand what will happen in the early stages, in the neighbourhood of each l_i or r_i. First of all, we can establish that these nodes can be expected to complete:

Lemma 4.19 (l_i and r_i complete): For any $\epsilon'>0$, if $0 \ll w \ll n$ then for all $i \in [1,k_0)$, l_i and r_i will (be defined and will) complete with probability $>1-\epsilon'$.

*Definition 4.20 (**Completion**):* We define **Completion** ($\in \Pi$) to be the event that all the l_i and r_i (for $1 \leq i < k_0$) are defined and complete.

We are now ready to prove the required dichotomy.

Lemma 4.21 (The required dichotomy): Suppose that w is large and that **Good spacing**, **Stable clear**, **Smooth** and **Completion** all hold. Then for $i<k_0$, l_i and r_i will each either subside or originate a firewall.

Proof: We prove the result for l_i, and the proof for r_i is essentially identical.

Note first that the choice of k_1 in Definition 4.18 means, in particular, that $10w/k_1$ type changes in any given neighbourhood cannot create stable intervals, given that **Stable clear** holds (the numbers here are fairly arbitrary). Note also, that satisfaction of **Smooth** suffices to ensure that there are not unhappy nodes of both types in the neighbourhood of l_i in the initial configuration. Now suppose that l_i completes at stage s and has positive bias in the initial configuration (the case for negative bias is essentially identical).

It is useful at this point to establish names for a number of relevant intervals. We let $u_1=l_i-(2w+1)$ and $u_2=l_i+(2w+1)$. Then we define:

- $J=\mathcal{N}(u_1)\cup\mathcal{N}(l_i)\cup\mathcal{N}(u_2)$.
- $I=[u_1,u_2]$.
- $I_1=[u_1,l_i]$, $I_2=[l_i,u_2]$.
- $K^1_j=I_1(j:k_1)^-\cup I_2(j:k_1)$.
- $K^2_j=I_1(j:k_2)^-\cup I_2(j:k_2)$.

In Definition 4.18 we assumed that k_2 is a multiple of k_1, so we may let m be such that $k_2=mk_1$. It is convenient to assume that k_2 is even. Now we divide into two cases.

l_i **subsides**. First of all, suppose that at stage s, there is a β-node u in the interval K^1_1. Then u must be happy at stage s. The fact that **Smooth** is satisfied, together with the fact that l_i completes at stage s, means that prior to stage s, the only type changes in the interval J are from type β to type α, so that u must be happy at *every* stage $\leq s$. Now, since $k_1 \ll k_2 \ll \frac{1}{\epsilon_0}$ and $\texttt{Smooth}^*_{k_2,\epsilon_0}(l_i)$ holds, any nodes in $I-(K^1_1\cup K^1_2)$ have lower bias than u, and hence are happy, in the initial configuration. It then follows by induction on the stages $\leq s$ that no node in $J-(K^1_1\cup K^1_2)$ changes type prior to stage s. In order to see this suppose that it holds prior to stage $s' \leq s$. Then at stage s', if $v \in I-(K^1_1\cup K^1_2)$, it still has lower bias than u and so cannot change type from β to α, since u is happy so v must be. If $v \in J-I$, then v changing type would contradict the fact that l_i completes.

We therefore get at most $|(K^1_1\cup K^1_2)|<10w/k_1$ many type changes in the interval J prior to completion. As observed above, this means that no stable intervals are created and l_i subsides, as required.

l_i **originates a firewall**. So suppose instead that, at stage s, all nodes in the interval K^1_1 are of type α. Given m as above, another way of putting this, is that all nodes in $\bigcup_{j\leq m}K^2_j$ are of type α at stage s. We now show by induction on $r \geq m$ that, when $r \leq k_2/2$, any nodes in K^2_r must be of type α at stage s – i.e. that all nodes in $\mathcal{N}(l_i)$ are α nodes at stage s. So suppose that $m \leq r < k_2/2$ and that the hypothesis holds for all $r' \leq r$. Consider $u \in K^2_{r+1}$. Let ρ be the bias of l_i in the initial configuration. First let us form a lower bound for the bias of u in the initial configuration. The fact that **Smooth** holds means that the leftmost and rightmost nodes in K^2_r have bias at least $\rho-\frac{r}{k_2}\rho-\epsilon_0 w$. Then, since the bias can change by at most 2 if we move left or right one node, we conclude that u has bias

$$\rho_1 \geq \rho-\left(\frac{r}{k_2}\rho+\epsilon_0 w+\frac{2(2w+1)}{k_2}\right)$$

in the initial configuration. Now we have to take into account all of the β nodes in $\bigcup_{j\leq r}K^2_r$ which have changed type. In fact, so that we can be sure that each change of type affects the bias of u, we shall consider just those which lie between l_i and u. Let θ be the proportion of the nodes in $\mathcal{N}(l_i)$ which are α nodes in the initial configuration (recalling that l_i has borderline bias at that point) – so that $\rho=(2\theta-1)(2w+1)$. Then the number of β nodes in $\bigcup_{j\leq r}K^2_j$ in the initial configuration, which lie between l_i and u, is at least:

$$(1-\theta-\epsilon_0)(2w+1)r/k_2.$$

Each change of type for one of these nodes means an increase of 2 in the bias of u, so that at stage s, u has bias ρ_2 which is at least

$$\rho-\left(\frac{r}{k_2}\rho+\epsilon_0 w+\frac{2(2w+1)}{k_2}\right)+\frac{2(1-\theta-\epsilon_0)(2w+1)r}{k_2}.$$

So we have $\rho_2 \geq \rho+(2w+1)\cdot A$, where

$$A:=\frac{(1-\theta)2r}{k_2}-\frac{\epsilon_0 w}{2w+1}-\frac{2}{k_2}-\frac{r}{k_2}(2\theta-1)-\frac{\epsilon_0 2r}{k_2}.$$

We are left to compare the terms

$$\frac{(1-\theta)2r}{k_2}, \quad \frac{\epsilon_0 w}{2w+1}, \quad \frac{2}{k_2}, \quad \frac{r}{k_2}(2\theta-1) \text{ and } \frac{\epsilon_0 2r}{k_2}.$$

Since $1/k_2 \gg \epsilon_0$, the second term is much smaller than the third. Since $\theta \in (0.5, 0.65)$, $r/k_2 \geq 1/k_1$ and $k_1 \ll k_2$, the first term is much larger than the third (to see that $\theta \in (0.5, 0.65)$ recall that l_i has borderline bias in the initial configuration and $\tau > \kappa > 0.35$). Since ϵ_0 is small, the first term is also much larger than the last. The result then follows for large w, since $2-2\theta$ is always more than double $2\theta - 1$ for $\theta \in (0.5, 0.65)$, meaning that the first term is more than double the fourth term, and thus $\rho_2 \geq \rho$, meaning that if u is a β node, it will be unhappy. ■

The following lemma is proved via approximation to a biased random walk:

Lemma 4.22 (Reasonable chance of firewall): Suppose that **Good spacing** holds. There exists $\delta > 0$ which does not depend on w, such that if $1 \leq i < k_0$, then l_i originates a firewall with probability $> \delta$ (and similarly for r_i).

Then the following definition gives our final event in Π and also specifies k_0.

*Definition 4.23 (Defining **Firewall** and choosing k_0):* We let **Firewall** be the event that one of the l_i $i < k_0$ is defined and originates a firewall, and that the same holds for some r_j, $j < k_0$. According to Lemma 4.22, given $\epsilon > 0$ we can choose k_0 once and for all, which is large enough such that the probability **Firewall** does not occur is $\ll \epsilon$ for $0 \ll w \ll n$.

Finally, Lemma 4.24 completes the proof of Theorem 1.2.

Lemma 4.24 (u_0 ultimately joins a firewall): Suppose that all events in Π hold. Let i be the least such that l_i is defined and originates a firewall, and let j be least such that r_j is defined and originates a firewall. For any $\epsilon' > 0$, for all sufficiently large w, with probability $> 1-\epsilon'$, u_0 will eventually be contained in one of the two firewalls originated at l_i and r_j.

Increasing τ in the interval $[\kappa, \frac{1}{2}]$ decreases run-lengths

Let us fix $\epsilon > 0$ and τ_1 and τ_2 such that $\kappa < \tau_1 < \tau_2 < \frac{1}{2}$. For some large w and $n \gg w$ suppose that we run two version of the process, one for τ_1 and the other for τ_2, and then let ℓ_1 and ℓ_2 be the corresponding run-lengths to which u_0 (chosen uniformly at random) belongs in the final configuration. Our aim in this subsection is to observe that, so long as w is sufficiently large, the probability that $\ell_1 > \ell_2$ is greater than $1-\epsilon$.

In order to see this we may reason as follows. In the proof of Theorem 1.2, we chose k_0 in such a way we could be almost certain at least one of the l_i and at least one of the r_i would originate firewalls. We could equally well have chosen k_0 so that the following fails to occur with probability $\ll \epsilon$: two of the l_i originate firewalls of opposite type, and similarly two of the r_i originate firewalls

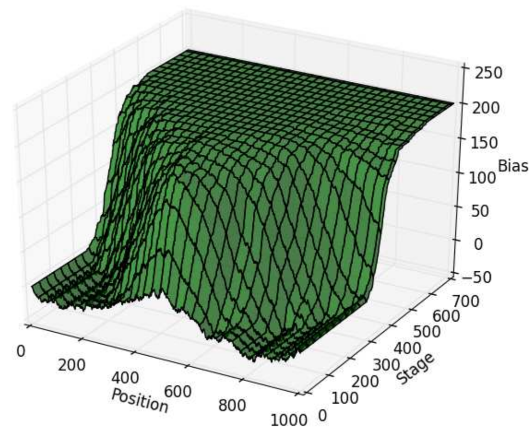

Figure 4: The creation of a firewall (from simulation).

of opposite type. Our previous analysis then suffices to show that the following fails to occur with probability $\ll \epsilon$ for sufficiently large w; u_0 ultimately belongs to a firewall of length $> \min\{|u_0 - l_1|, |u_0 - r_1|\}$ and of length $< [l_{k_0}, r_{k_0}]$. So it suffices to show that with probability $> 1-\epsilon$, the length of the interval $[l_{k_0}, r_{k_0}]$ for the process with τ_2 is less than the value $\min\{|u_0 - l_1|, |u_0 - r_1|\}$ for the process with τ_1. This then follows for sufficiently large w, by applying Lemma 3.3 of [1] to the events P_u, that u is τ_2-unhappy, and Q_u that u is τ_1-unhappy.

V. The case $\tau > 0.5$

Note that if $\frac{1}{2} < \tau \leq \frac{w+1}{2w+1}$ then the process is identical to that for $\tau = \frac{1}{2}$. We therefore assume in what follows that w is sufficiently large to ensure $\tau > \frac{w+1}{2w+1}$, which is equivalent to the condition that adjacent nodes of opposite types cannot both be happy. Note that once a completely segregated configuration is reached, all future states are completely segregated. Let x be the number of α nodes in the initial configuration, and let $x_p = x/n$. Our task is to show that for sufficiently large n it is possible to reach a completely segregated configuration from any other, so long as x_p is close to $\frac{1}{2}$. To prove this, however, one must be able to ensure the existence of unhappy individuals of both types at each step along the way. The following lemma is perhaps surprisingly tricky to prove:

Lemma 5.1: With probability tending to 1 as $n \to \infty$, x satisfies the property that there are unhappy α nodes (in fact, unhappy α nodes outside any interval of length $4w+1$) in *any* configuration on a ring of size n with x many α nodes. A similar result holds for β.

To complete the argument, we then build a list of configurations from which it is possible to reach complete segregation. Consider first any configuration which is not completely segregated, but which has a run of length at least $2w$. Without loss of generality, suppose that this is a run of α nodes occupying the interval $[a, b]$, where this interval is chosen to be of maximum possible length. If the nodes a and b are both happy then the length of the interval ensures that all nodes in the run are happy – this follows by induction on the distance from the edge of the interval. In this case let u be an unhappy α node and let $c \in \{a, b\}$ be distance at least $w+1$ from u. Then u and the β neighbour of c may legally be swapped, increasing the length of the run by at least 1. So suppose instead that at least one of the individuals a and b is not happy, and without loss of generality suppose that a has bias less than or equal to b. Then a and $b+1$ may legally be swapped. Performing this swap causes position $b+1$ to have at least the same bias as b did before the swap, and causes $a+1$ to have at most the same bias as a did before the swap. Thus, the swap has the effect of shifting the run one position to the right and may be repeated until the length of the run is increased by at least 1, i.e. for successive $i \geq 0$ we can swap the nodes $a+i$ and $b+i+1$, so long as the latter is of type β. The first stage at which the latter is of type α the length of the run has been increased. Putting these observations together, we conclude that from any configuration which has a run of length at least $2w$ it is possible to reach full segregation.

Next consider a configuration in which the longest run $[a, b]$ is of length at least w, but strictly less than $2w$. We shall suppose that $[a, b]$ contains α nodes, the case for β nodes is similar. Let c be the first α node strictly to the left of a. If c is unhappy, then we may legally swap c and $a-1$, strictly increasing the length of the longest run. If c is happy then the distance between c and a is at most w and we may successively swap unhappy α nodes from outside the interval $[c-w, a+2w]$ with the nodes $c+i$ for $1 \leq i < a-c$ (starting with $i=1$ and proceeding in order), in order to strictly increase the length of the longest run. This follows because as each node $c+i$ performs the swap, it will become happy.

It remains to show that we can always move to a configuration with a run of length at least w. The proof appears in [1].

Acknowledgments

Andy Lewis-Pye (previously Andrew Lewis) was supported by a Royal Society University Research Fellowship. Barmpalias was supported by the Research Fund for International Young Scientists from the National Natural Science Foundation of China, grant number 613501-10535, and an International Young Scientist Fellowship from the Chinese Academy of Sciences; partial support was also received from the project Network Algorithms and Digital Information from the Institute of Software, Chinese Academy of Sciences and a Marsden grant of New Zealand.

References

[1] G. Barmpalias, R. Elwes, A. Lewis-Pye, Digital morphogenesis via Schelling segregation (full version), http://arxiv.org/abs/1302.4014.

[2] C. Brandt, N. Immorlica, G. Kamath, R. Kleinberg, An Analysis of One-Dimensional Schelling Segregation, *Proc. 44th Annual ACM Symposium on Theory of Computing (STOC 2012).*

[3] L. Dall'Asta, C. Castellano, M. Marsili, Statistical physics of the Schelling model of segregation, *J. Stat. Mech*, 7, 2008.

[4] L. Gauvin, J. Vannemenus, J.-P. Nadal, Phase diagram of a Schelling segregation model, *European Physical Journal B*, 70, 293–304, 2009.

[5] A.D. Henry, P. Prałat, and C. Zhang, Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences*, 108(21):8605-8610, May 2011.

[6] J. Kleinberg, Cascading Behavior in Networks: Algorithmic and Economic Issues, *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani, eds., Cambridge University Press, 2007.

[7] T. Schelling, Models of segregation, *The American Economic Review*, 488-493, 1969.

[8] D. Stauffer and S. Solomon, Ising, Schelling and self-organising segregation, *European Physical Journal B*, 57, 473–479, 2007.

[9] A.M. Turing, The Chemical Basis of Morphogenesis , *Phil. Trans. R. Soc. London B*, 237, 37-72, 1952.

[10] D. Vinković and A. Kirman, A Physical Analogue of the Schelling Model, *Proceedings of the National Academy of Sciences*, no. 51, volume 103, 19261-19265, 2006.

[11] N.C. Wormald, Differential equations for random processes and random graphs, *Annals of Applied Probability*, 5:1217-1235, 1995.

[12] H.P. Young, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton, NJ: Princeton University Press, 1998.

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

[14] J. Zhang, Residential segregation in an all-integrationist world, *Journal of Economic Behavior & Organization*, 54(4):533-550, 2004.

[15] J. Zhang, Tipping and residential segregation: A unified Schelling model, *Journal of Regional Science*, 51:167-193, Feb. 2011.