# Axelrod's Model in Two Dimensions

Junchi Li
Department of Mathematics, Duke University
Supervisor: Rick Durrett

May 2014

## 摘 要

In 1997 R. Axelrod introduced a model in which individuals have one of $Q$ possible opinions about each of $F$ issues and neighbors interact at a rate proportional to the fraction of opinions they share. Thanks to work by Lanchier and collaborators there are now a number of results for the one dimensional model. Here, we consider Axelrod's model on a square subset of the two-dimensional lattice starting from a randomly chosen initial state and simplify things by supposing that $Q$ and $F$ large. If $Q/F$ is large then most neighbors have all opinions different and do not interact, so by a result of Lanchier the system soon reaches a highly disordered absorbing state. In contrast if $Q/F$ is small, then there is a giant component of individuals who share at least one opinion. In this case we show that consensus develops on this cluster.

# 1 Voter Model and Axelrod's Model in One Dimension

## 1.1 Voter model on integer lattices

In the basic voter model [19, 6], individuals have one of two opinions about an issue. The prototypical example is what political party they belong to. The voter model on the $d$-dimensional integer lattice can be defined as follows. For $i \in \mathbb{Z}^d$, $\xi_t(i) \in \{0, 1\}$ represents the binary opinion of voter $i$ at time $t$. For each directed edge $(i, j)$ only at $t \in \{T_n^{(i,j)}, n \geq 1\}$ of a rate 1/2 Poisson arrivals $i$ decides to take the opinion of $j$, i.e. $\xi_t(i) = \xi_{t-}(j)$. Obviously, the dynamics is a pure jump process which takes place on the state space $\{0, 1\}^{\mathbb{Z}^d}$. The voter model on integer lattices is a simple model characterizing voting behaviors and there have been a number of results, of which the most classical and important one is

**Theorem 1.** *(i)* ***Clustering*** *occurs when $d \leq 2$: for any $\xi_0$ and distinct pair $i, j \in \mathbb{Z}^d$ we have*

$$P(\xi_t(i) \neq \xi_t(j)) \to 0 \quad \textit{as } t \to \infty$$

*(ii)* ***Coexistence*** *occurs when $d \geq 3$: let $\xi_t^\theta$ denote the process starting from an initial state in which the events $\{\xi_0^\theta(i) = 1\}$ are independent and have probability $\theta$. As $t \to \infty$ the dynamics $\xi_t^\theta$ converges in distribution to $\xi_\infty^\theta$, a stationary distribution where $\{\xi_\infty^\theta(i) = 1\}$ has probability $\theta$ for each $i \in \mathbb{Z}^d$.*

*Proof of Theorem 1.* As illustrated in the case of one dimensions in Figure 1.1, the key to the proof is to use a "dual process" $\{\zeta_s^{t,i} : i \in \mathbb{Z}^d\} \subseteq \mathbb{Z}^d$ which traces the opinion of each voter backwards in time to determine the source of the opinions at time $t$. For each $i \in \mathbb{Z}^d$, $\zeta_s^{t,i}$ simply stays put at $i$ until the first time $s$ that $t - s = T_n^{(i,j)}$ for some $n, j$, where we set $\zeta_s^{t,i} = j$. By Markov property the process can be defined recursively until the next $T_n^{(i,j)}$. Such process forms the dynamics of *coalescing random walks*, i.e. each voter independently moves as a simple random walk until they hit each other. Afterwards they move together. We have the duality property that

$$\xi_t(i) = \xi_{t-s}(\zeta_s^{t,i}), \quad 0 \leq s \leq t. \qquad (*)$$

Coming back to the proof, we note for $d = 1, 2$ using duality property $(*)$

$$P(\xi_t(i) \neq \xi_t(j)) \leq P(\xi_0(\zeta_t^{t,i}) \neq \xi_0(\zeta_t^{t,j})) \leq P(\zeta_t^{t,i} \neq \zeta_t^{t,j})$$

which tends to 0 since the process is recurrent, completing part (a). For part (b) we refer the readers to Section 2 of [6] so they can persuade themselves that it is enough to prove the convergence of $P(\xi_t^\theta(i) = 0 \text{ for all } i \in B)$ for each $B \in \mathbb{Z}^d$. Observe by dual relations

$$P(\xi_t^\theta(i) = 0 \text{ for all } i \in B) = E(1-\theta)^{|\zeta_t^{B,t}|}$$

where $\zeta_s^{B,t}$ is the bundle version of $\zeta_s^{i,t}, i \in B$. Note that the distribution of $\zeta_t^{B,t}$ is the same as the marginal distribution of coalescing random walk $\zeta_t^B$, which is nonincreasing of $t$. It hence follows from the bounded convergence theorem that $P(\xi_t^\theta \cap B = \varnothing)$ has a limit and the proof is complete. $\square$

## 1.2 Axelrod's model

In reality, there are many issues that people have opinions about, such as income tax policy, environmental issues, gambling, gun control, gay marriage, etc., and one can have a number of different opinions about each issue. Axelrod [2] formulated a model in which each individual holds one of $Q$ opinions about each of $F$ issues, so the state of the process at time $t$ is described by giving for each voter $x$ and issue $f$, $\xi_t(x, f) \in \{1, 2, \dots Q\}$. Each oriented pair of connected voters $(x, y)$ interacts at times of a Poisson process at rate $1/2$. At these times an issue $f = 1, 2, \dots, F$ is picked at random. If $\xi_t(x, f) = \xi_t(y, f)$ then $x$ picks an issue from the ones they disagree on, and changes its opinion to agree with $y$ on that particular issue. If they already agree on all issues no change occurs.

Note that while the motivation may suggest that the $Q$ values are positions that range from strongly in favor to strongly against, individuals act only if their opinions agree exactly on the issue chosen. A number of researchers have studied the variation in which $F = 1$, the state of an individual is a point along a political continuum $[0, 1]$ and only neighboring individuals whose opinions differ by less than $\varepsilon \in (0, 1]$ will interact. Deffuant et al [5] considered a version in which an interaction opinions $a$ and $b$ results in $a+\mu(b-1)$ and $b+\mu(a-b)$. If $\varepsilon \geq 1/2$ then the system converges to consensus, but if $\varepsilon < 1/2$ one ends up with roughly $1/2\varepsilon$ opinions. For

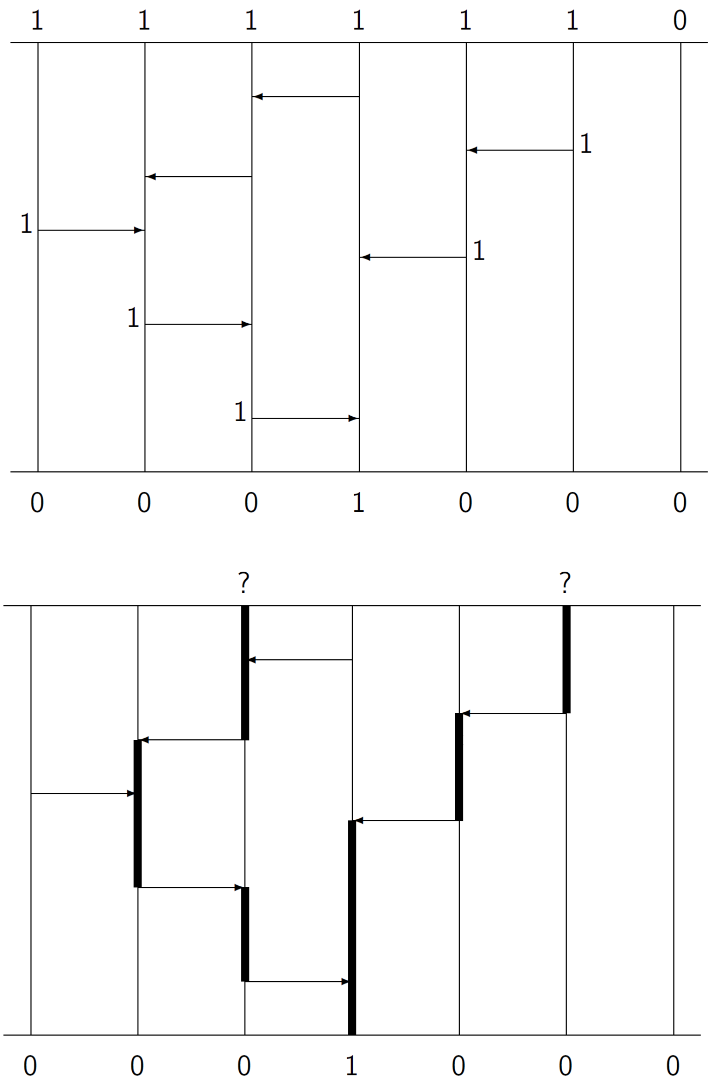


图 1: Illustration of voter model duality more on this model see section II.F in Castellano et al [3]. Lanchier [15] has recently studied a version in which when two individuals with opinions that differ by $\varepsilon$ interact, one imitates the other.

The tendency for people to interact more frequently with people who are similar to them is called *homophily.* Axelrod did a number of computer experiments and found, for example that if there are 5 features and 10 opinions (which he called traits) then on a 10 by 10 grid one would end up with an average of 3.2 "cultural domains," clusters of people with the same opinion on all issues. Castellano, Marsili, and Vespignani [4] used simulation to study the model in one dimension. They concluded that if $F > 2$ then on a long line segment Axelrod's model (a) converges to a monocultural equilibrium when $Q < F$ and (b) fixates at a highly fragmented configuration when $Q > F$. The case $F = 2$ is different. The ending state is disordered for any value of $Q$. Vilone, Vespignani, and Castellano [22] confirmed these predictions using mean-field theory and more simulation. There has been an extensive study of Axelrod's model in the physics literature, see chapter IV of Castellano et al [3].

When $F = 2$ and $Q = 2$ Axelrod's model reduces to the constrained voter model of Vázquez, Krapivsky, and Redner [20]. In that model there are Leftists (1), Centrists (0), and Rightists ($-1$) but 1's and $-1$s ignore each other, i.e., a 1 becomes 0 at a rate equal to the fraction of neighbors that are 0, but we never see changes $1 \to -1$. To map Axelrod's model into this system: $11 \to L$, $10, 01 \to C$, and $00 \to R$. If we further map the constrained voter model so that $L, R \to 0$ and $C \to 1$ then the system becomes an ordinary voter model. From this we see that in $d \le 2$ one will either get fixation into an all $C$ state or a frozen mixture of $L$ and $R$. For more on this system, see Vázquez and Redner [21] or Section III.B of Castellano et al [3].

#### 1.2.1 Rigorous results

While Axelrod's model has been studied extensively by simulation and non-rigorous methods such as the pair approximation, the first rigorous results only appeared recently. Lanchier [16] proved

**Theorem 2.** *The one-dimensional model on $\mathbb{Z}$ with $Q = F = 2$ clusters, that is, for each $f$ and $x \sim y$, $P(\xi_t(x, f) \neq \xi_t(y, f)) \to 0$.*

Note that this result is different from the convergence to a disordered state seen in simulation on a finite set. The proof helps explain the difference. Lanchier looks at boundaries between the four types. $00, 11$ boundaries are static, while the other possibilities are active and perform simple random walks. When an active boundary hits a static one it may become active and in this way the density of static boundaries is reduced to 0. However, on a finite set one will eventually run out of active boundaries. This results has recently been generalized by Lanchier and Schweinsberg [18] to show that on $\mathbb{Z}$ if $Q = 2$ then the system clusters for any $F < \infty$.

More interesting and more difficult is Lanchier's result in the other direction. Consider the system on $G = \{0, 1, \dots N\}$ and let $K_\infty$ be the number of cultural domains when the process fixates.

**Theorem 3.** *If $F < Q$ then*

$$N^{-1} E K_\infty \geq \frac{Q}{Q-F}\left(1 - \frac{1}{Q}\right)^F - \frac{F}{Q-F}\left(1 - \frac{1}{Q}\right)$$

The right-hand side may be $\leq 0$ but when it is $> 0$, the result implies that the system fixates in a fragmented state, where the number of cultural domains is proportional to the system size. If $F \to \infty$ and $F/Q \to c$ then the right hand side converges to

$$\frac{1}{1-c} e^{-c} - \frac{c}{c-1} = 0 \quad \text{when} \quad e^{-c} = c$$

The last equation has a unique solution at $c_0 \approx 0.567$. Lanchier and Scarlatos [17] have recently shown that Axelrod's model on $\mathbb{Z}$ fixates when $F \leq c_0 Q$. In addition they showed that the model on $\mathbb{Z}$ with $F = 2$ and $Q = 3$ fixates.

# 2 Axelrod's Model on the Two-dimensional Lattice

## 2.1 Main Theorem

In this chapter, our goal is to prove a result for Axelrod's model on a large finite two-dimensional lattice $\{0, 1, \ldots, N\}^2$ with periodic boundary conditions where the initial condition is a random assignment of opinions. If $Q \to \infty$ and $F/Q \to c$ then on each edge $e = \{x, y\}$ the number of agreements $A_e(0) = |\{f : \xi_0(x, f) = \xi_0(y, f)\}|$ has a limiting Poisson distribution with mean $c$, and the number of agreements on different edges are independent, so the limiting probability of *active edges* that have at least one agreement at time 0 is $1 - e^{-c}$. The critical value for bond percolation on $\mathbb{Z}^2$ is $1/2$, so when $c > \ln(2)$, $1 - e^{-c} > 1/2$ and we are in the supercritical case. On the graph $\{0, 1, \ldots, N\}^2$ this means that there will be a giant component that contains a positive fraction of the $N^2$ voters and the second largest component will be $O(\log N)$. Let $\mathcal{E}_+ = \{e : A_e(0) > 0\}$ be the edges that are active at time 0. Our goal is to show that

**Theorem 4.** *If $F/Q \geq \ln(2) + \eta$ and $Q > Q_0(N, \eta)$ then with high probability Axelrod's model reaches consensus on a giant component of $\{0, 1, \ldots, N\}^2$.*

It is possible that edges $e$ that have no agreements initially will become active at some point time. We call such events *surprises*. For example if $\xi_0(x - (1, 0), f) = \xi(x + (1, 0), f)$ and $\xi_0(x, f)$ is different then $x$ imitating one of its neighbors will suddenly produce two agreements. However, in the regime we are considering with a fixed graph and large $Q$ and $F$, these surprise events are rare.

To be precise, the number of issues $f$ for which $\xi_t(x, f) = \xi_t(y, f)$ for some $x$ and $y$ is $O(N^4)$. Since we are considering the limiting behavior as $F, Q \to \infty$ with $N$ fixed, this number is $O(1)$ as far as limits in $Q$ are concerned. As the reader will see from later arguments, surprises will not occur while the agreements along edges in $\mathcal{E}_+$ remains $\leq F^{1-\delta}$. Thus we will prove our result by first ignoring surprises and then going back to bound their influence.

**Phase I.** If $e = \{x, y\}$ we let $I_e(t) = \{f : \xi_0(x, f) = \xi_0(y, f)\}$, and

$A_e(s) = |I_e(sF \log F)|$. The first step is to bound $A_e(s)$ by a Yule process $A_e^+(s)$ run at rate $\log F$ and use coupling to show that at all times $s \le 0.45$, $A_e(s) = A_e^+(s)$. Given edges $e = \{x, y\}$ and $e' = \{y, z\}$, we do not have the worry about an agreement created by $y$ imitating $x$ disturbing one between $y$ and $z$ until $A_e$ and $A_{e'}$ are $O(F^{0.5})$. Since the Yule process $A_e^+(s) \approx W_e F^s$ this is a manifestation of the birthday problem.

**Phase IIA.** When the number of agreements on edges exceeds $F^{0.5}$ then there will be some loss of agreements on an edge due to creation of agreements on its neighboring edges, but up until time $1-\delta$, $A_e(s)$ will be close to the branching process. To be precise if we let $W_e = F^{-0.45} A_e^+(0.45)$ then with high probability if $e \in \mathcal{E}_+$

$$|F^{-s} A_e(s) - W_e| \le F^{-0.2} \quad \text{for } 0.45 \le s \le 1-\delta$$

**Phase IIB.** Let $\kappa$ be a small constant. The next step is to show that up until time $T_1$ which is equal to $\inf\{s : \max_e A_e^+(s) \ge \kappa F\}$ we have if $e \in \mathcal{E}_+$ then

$$F^{-s} A_e(s) \ge W_e - F^{-0.2} - \frac{3\widehat{W}\kappa}{1-\kappa} \quad \text{for } 1-\delta \le s \le T_1$$

where $\widehat{W} = \max_e W_e + F^{-0.2}$. If $\kappa$ is small enough then the right-hand side will be positive for all $e \in \mathcal{E}_+$.

**Phase IIIA.** Let $\mathtt{t}$ be a tree that is a subset of the initial giant component $\chi$ and let $p$ be the number of edges in the tree. (The somewhat unusual notation comes from the fact that we first wrote this argument for a path.) Let $A_{\mathtt{t}}$ be the number of issues on which all members of the tree agree. Let $\Delta_p = 1 - 1/p$ for $p \ge 2$. As the reader will see this is the time at which $p$-edge agreements first appear. The first step is show that there is a constant $K_{\mathtt{t}}$ so that if $W_{\mathtt{t}} = K_{\mathtt{t}}(1/p!) \prod_{e \in \mathtt{t}} W_e$ then with high probability

$$|F^{-ps+(p-1)} A_{\mathtt{t}}(s) - W_{\mathtt{t}}| \le F^{-\delta/3} \quad \text{for } \Delta_p + 2\delta/p \le s \le 1 - \delta/p$$

Note that $-ps + (p-1)$ when $s = \Delta_p$ so that is the first time that $p$ edge agreements occur. When $p = 2$, $\Delta_p = 1/2$.

**Phase IIIB.** We next extend the result in Phase IIIA up to time $T_1$. There are constants $k_{\mathtt{t}} > 0$ so that with high probability if $F \ge F_{\mathtt{t}}$ then

$$F^{-ps+(p-1)} A_{\mathtt{t}}(s) \ge k_{\mathtt{t}} \quad \text{for } 1 - \delta/p \le s \le T_1$$

If we take $\mathtt{t}$ to be a spanning tree then this result tells us that all individual in any connected component (later called *component* for short) $\psi \subset \mathcal{E}_+$ agree on a positive fraction of issues.

**Phase IV.** If we ignore surprises and shift to running time at rate $F$ after time $T_1$, then all edges see agreements being created at a positive rate. If we let $\mathcal{A}_e(t) = |I_e(T_1 F \log F + tF)|$ then comparing with a nonhomogeneous voter model allows us to drive the process to consensus on $\mathcal{E}_+$ by time $T_1 F \log F + c_N F \log\log F$.

**Phase V.** Finally we have to worry about surprises getting in the way of consensus. We will show that any edge $e$ created by a surprise has $\leq F^\delta$ agreements up to time $T_1$. Even in the process with surprises, the initial giant component $\chi$ will again have agreement on a positive fraction of the issues at time $T_1$. It will maintain that property until there is agreement on $\chi \cup e$ in the case of one edge surprise, or on $\chi \cup e \cup \chi'$ where $e$ connects to a component $\chi'$. We will define these terms more precisely in Chapter 3.

## 2.2 Growth phase I

The first step is to construct Axelrod's model using a graphical representation in order to couple the early stages to a branching process. Let $e = \{x, y\}$ be an edge and let $I_e(t) = \{f = 1, \dots, F : \xi_t(x, f) = \xi_t(y, f)\}$ be the set of issues that individuals at the ends of $e$ agree upon at time $t$ and $|I_e(t)|$ be its cardinality. For each $e \in \mathcal{E}$, we define a process $J_e(t)$ with $J_e(0) = I_e(0)$. To couple $I_e(t)$ with $J_e(t)$, we introduce for each oriented edge $(x, y)$

(1) A Poisson process $\{U^n_{(x,y)}, n \geq 1\}$ with rate $1/2$,

(2) $\{V^n_{(x,y)}(k) : n \geq 1, k \geq 0\}$ an array of independent draws from discrete uniform distribution on $\{1, \dots, F\}$.

Since the graph is finite, the sequence of Poisson arrivals associated with all edges can be ordered. Suppose the process has been constructed until time $t-$ where $t = U^n_{(x,y)}$ and we have $I_e(t-) \subset J_e(t-)$ for all $e \in \mathcal{E}$. Writing $x, y$ for the unoriented edge

- If $|I_{x,y}(t)| = F$ there is nothing to do. If not and $V^n_{(x,y)}(0) \leq |I_{x,y}(t)|$ then we read the independent sequence of choices from $\{1, \dots F\}$, $V^n_{(x,y)}(k)$, $k \geq 1$ until we find the first $V^n_{(x,y)}(K_I) \notin I_{x,y}(t-)$, then we set set $\xi_t(x, V^n_{(x,y)}(K_I)) = \xi_{t-}(y, V^n_{(x,y)}(K_I))$, and do not change any other opinions.

- If $|J_{x,y}(t)| = F$ there is nothing to do. If not and $V^n_{(x,y)}(0) \leq |J_{x,y}(t)|$ then we read the sequence $V^n_{(x,y)}(k)$, $k \geq 1$ until we find the first $V^n_{(x,y)}(K_J) \notin J_{x,y}(t-)$ and add $V^n_{(x,y)}(K_J)$ to $J_{x,y}(t)$.

To check that $I_e(t) \subset J_e(t)$ for all $e$ at time $U^n_{(x,y)}$, it suffices to check that $I_{x,y}(t) \subset J_{x,y}(t)$, since the other $J_e$ don't change, and assuming there are no surprises, the other $I_e$ can only decrease. To check the inequality for $e = x, y$ note that when $I_{x,y}(t-) = J_{x,y}(t-)$, $K_I = K_J$ and the same point will be added to both sets. When $I_{x,y}(t-) \neq J_{x,y}(t-)$, we may have $K_I < K_J$ but in this case the point added to $I_{x,y}(t)$ will be one that is already in $J_{x,y}(t)$.

For calculations it is convenient to extend $|J_{x,y}(t)|$ after the time $T^+_{x,y}$ it becomes absorbed at $F$. We define $B^+_e(t) = |J_{x,y}(t)|$ for $t \leq T^+_{x,y}$ and extend $B_e^+(t)$ to be a Yule process at rate $1/F$, i.e.

$$B_e^+(t) : k \to k+1 \qquad \text{at rate } k/F.$$

From the coupling the following is obvious:

**Lemma 2.1.** *For each $e \in \mathcal{E}$ and $t \geq 0$, $|I_e(t)| \leq B_e^+(t)$.*

We are interested in understanding the growth of $|I_e(t)|$ until it gets up to size $F$, so we speed up time. Let

$$A_e(s) = |I_e(sF \log F)| \qquad A_e^+(s) = B_e^+(sF \log F) \tag{2.1}$$

Lemma 2.1 implies $A_e(s) \leq A_e^+(s)$, for all $s \geq 0$. It is easy to understand the behavior of $A_e^+(t)$

**Lemma 2.2.** *If $s < 1$, $F^{-s} A_e^+(s)$ converges weakly to Gamma$(|I_e(0)|, 1)$ as $F \to \infty$.*

证明. Let $Z(t)$ be the Yule process starting from $Z(0) = 1$. It is well known that, see e.g. [1], that

$$\lim_{t \to \infty} e^{-t} Z(t) = W \ a.s.$$

where $W$ has an exponential distribution with mean 1. $A_e^+(t)$ is the sum of $|I_e(0)|$ independent Yule processes, so the desired result follows by changing variables $t = s \log F$. □

Our next step is to show that with high probability the coupling is exact up to time $s = 0.45$. Let

$$\Omega_1 = \{A_e^+(s) = A_e(s) > 0, \forall s \leq 0.45, e \in \mathcal{E}\}.$$

**Lemma 2.3.** *As $F \to \infty$, $P(\Omega_1) \to 1$.*

证明. Discrepancies in the coupling, which we call *collisions*, occur when we set $\xi_t(x, f) = \xi_{t-}(y, f)$ and there is an edge $e' = (x, z)$ so that $f \in I_{e'}(t-)$. Let $B(x, y, z, s)$ be the event that this occurs for fixed $x \sim y$ and $z \sim x$ by time $sF \log F$. Let

$$H = \left\{ \max_{e \in \mathcal{E}} A_e^+(0.45) \leq F^{0.49} \right\}.$$

Since $\mathcal{E}$ is finite, Lemma 2.2 implies $P(H) \to 1$. It is easy to see that if one takes $s = 0.45$ then

$$P(B(x, y, z, 0.45)^c | H) \geq \prod_{k=1}^{F^{0.49}} \left(1 - \frac{k}{F - F^{0.49}}\right) \geq \left(1 - \frac{F^{0.49}}{F - F^{0.49}}\right)^{F^{0.49}} \to 1$$

The equation comes from the fact that in general when the $k$-th issue is randomly selected it cannot be the same as any of the $k - 1$ preceding issues. Since there are only finitely many $x, y, z$ to consider the desired result follows. □

### 2.3 Growth phase II

Let $\kappa > 0$ be a small constant to be determined later. Our goal is to show that $A_e(s)$ is well approximated by $A_e^+(s)$ defined in (2.1) up to a time $T_1 = \inf\{s : \max_e A_e^+(s) \geq \kappa F\}$.

#### 2.3.1 Upper bound on $A_e^+(s)$

Let $\mathcal{E}$ be the set of all edges, and let $\mathcal{F}_t^+$ be the $\sigma$-field generated by $(A_e^+(s) : e \in \mathcal{E}, s \leq t)$. The first order of business is to upper bound the growth of $A_e^+(s)$. Let

$$W_e(0.45) = F^{-0.45} A_e^+(0.45)$$

and recall from the previous section that the $A_e^+(t)$ are independent Yule processes run at rate $\log F$ defined for all time.

**Lemma 2.4.** *For each $e \in \mathcal{E}$,*

$$E^+ \left( \sup_{s \geq 0.45} F^{-s} A_e^+(s) - W_e(0.45) \right)^2 \leq C W_e(0.45) F^{-0.45}$$

*where $E^+$ is shorthand for $E(\,\cdot\,|\mathcal{F}_{0.45}^+)$.*

证明. $(F^{-s} A_e^+(s) - W_e(0.45) : s \geq 0.45)$ is a martingale under $P^+$. Classical theory of branching processes [1] gives for $t \geq 0.45$

$$E^+(A_e^+(t) - e^{(t-0.45)\log F} A_e^+(0.45))^2 \leq C_1 A_e^+(0.45) F^{2(t-0.45)}$$

so multiplying by $e^{-t\log F} = F^{-t}$

$$E^+(F^{-t} A_e^+(t) - W_e(0.45))^2 \leq C_1 W_e(0.45) F^{-0.45}$$

Using the $L^2$ maximal inequality for martingales

$$E^+ \left( \sup_{0.45 \leq s \leq t} F^{-s} A_e^+(s) - W_e(0.45) \right)^2 \leq C_2 W_e(0.45) F^{-0.45}$$

Since the right-hand side is independent of $t$ the desired result follows. $\square$

Let $\overline{W} = \max_{e \in \mathcal{E}} W_e(0.45) \in \mathcal{F}_{0.45}$, let $\widehat{W} = \overline{W} + F^{-0.2}$, and we remind the reader that

$$T_1 = \inf\{s : \max_e A_e^+(s) \geq \kappa F\} \tag{2.2}$$

It immediately follows from Lemma 2.4 along with Chebyshev's inequality that with high probability under measure $P^+$

$$\sup_{e\in\mathcal{E}} A_e^+(s) \le \widehat{W} F^s, \qquad \forall s \ge 0.45.$$

Set $T_1^o = 1 + \frac{\log(\kappa/\widehat{W})}{\log F}$ to be the approximation of $T_1$ precisely defined in (2.2), then

$$\sup_{e\in\mathcal{E}} A_e^+(T_1^o) \le \widehat{W} F \frac{\kappa}{\widehat{W}} = \kappa F, \qquad \text{indicating that } T_1^o \le T_1 \tag{2.3}$$

### 2.3.2 Lower bound on $A_e(s)$

$A_e$ changes by

$$\begin{aligned} &+1 \qquad \text{at rate } A_e \log F \\ &-1 \qquad \text{at rate } \frac{1}{2}\sum_{e'\sim e} \frac{A_e - A_{e,e'}}{F - A_{e'}} A_{e'} \log F \end{aligned}$$

where $e' \sim e$ indicates $e'$ and $e$ shares a common voter, and $A_{e,e'}$ is the number of issues for which there is agreement on $e$ and on $e'$. To get a lower bound on $A_e(t)$, we define the birth-and-death chain $A_e^-$ which changes by

$$\begin{aligned} &+1 \qquad \text{at rate } A_e^- \log F \\ &-1 \qquad \text{at rate } \frac{3\widehat{W} F^s}{F - \widehat{W} F^s} A_e^- \log F \end{aligned}$$

where the 3 comes from the fact that on the two dimensional lattice there are 6 $e' \sim e$ and oriented edges have events at rate $1/2$.

**Lemma 2.5.** *We can define our processes on the same probability space so that*

$$A_e^-(s) \le A_e(s) \le A_e^+(s), \qquad \textit{for all } 0.45 \le s \le T_1^o.$$

证明. Since $A_e - A_{e,e'} \le A_e$ and $A_e, A_{e'} \le \widehat{W} F^s$ the fraction in the definition of $A_e^-$ is larger than the one in the definition of $A_e$

$$A_{e'} \frac{A_e - A_{e,e'}}{F - A_e} \le \widehat{W} F^s \frac{A_e}{F - \widehat{W} F^s}$$

The fact that $A_e^- \le A_e$ may look like trouble, but these are birth and death chains so it is enough that the rate of jumping by $-1$ is larger in $A_e^-$ when $A_e^- = A_e$. $\square$

To analyze $A_e^-$ we begin by computing the infinitesimal mean and variance of $U_e^-(s) = F^{-s} A_e^-(s)$. Note

$$U_e^-(s+h) - U^-(s) = F^{-(s+h)}(A_e^-(s+h) - A_e^-(s)) + (F^{-(s+h)} - F^{-s})A_e^-(s) \tag{2.4}$$

In terms of stochastic differentials

$$dU_e^-(s) = F^{-s} dA_e^-(s) - (\log F) F^{-s} A_e^-(s) ds. \tag{2.5}$$

From this we see that the infinitesimal mean

$$\begin{aligned} b(U_e^-) &= F^{-s}\left(1 - \frac{3\widehat{W}F^s}{F - \widehat{W}F^s}\right) A_e^- \log F - (\log F) F^{-s} A_e^-(s) \\ &= -(\log F) F^{-s}\left(\frac{3\widehat{W}F^s}{F - \widehat{W}F^s}\right) A_e^- \end{aligned} \tag{2.6}$$

To compute the infinitesimal variance from first principles, note since $A_e^-$ jumps by $\pm 1$, as $h \to 0$ the second moment of the first difference in (2.4) is

$$\sim hF^{-2s}\left(1 + \frac{3\widehat{W}F^s}{F - \widehat{W}F^s}\right) A_e^- \log F$$

while the second moment of the second difference is $h^2((\log F) F^{-s} A_e^-(s))^2$. Cauchy-Schwarz inequality implies that the cross term in $(U_e^-(s+h) - U^-(s))^2$ is $O(h^{3/2})$, so the infinitesimal variance is

$$v_s(U_e^-) = F^{-s}\left(1 + \frac{3\widehat{W}F^s}{F - \widehat{W}F^s}\right) U_e^- \log F \tag{2.7}$$

To bound $A_e^-(t)$ we will use the $L^2$ maximal inequality on

$$U_e^-(s) - U_e^-(0.45) - \int_0^t b(U_e^-(s))\, ds$$

**Lemma 2.6.** *For $e \in \mathcal{E}$ we have with high probability that for all $0.45 \le t \le T_1^o$*

$$U_e^-(s) \ge W_e(0.45) - \frac{3\widehat{W}\kappa}{1-\kappa} - F^{-0.2}$$

证明. Lemmas 2.5 and 2.4 imply that $U_e^-(s) \le F^{-s} A_e^+(s) \le \widehat{W}$. For $0.45 \le s \le T_1^o$ the infinitesimal variance term

$$v(U_e^-(s)) \le F^{-0.45}\left(1 + \frac{3\widehat{W}F^{s-1}}{1-\kappa}\right)\widehat{W}$$

which is $o(F^{-0.4})$ as $F \to \infty$. Using the $L^2$ maximal inequality it follows that as $F \to \infty$

$$P^* \left( \sup_{0.45 \le t \le 0.7} \left| U_e^-(t) - W_e(0.45) + \int_{0.45}^t (\log F) \left( \frac{3\widehat{W} F^{s-1}}{1 - \widehat{W} F^{s-1}} \right) U_e^-(s)\, ds \right| > F^{-0.2} \right) \to 0.$$

Using $U_e^-(s) \le \widehat{W}$ for $s \ge 0.45$, and $\widehat{W} F^{s-1} \le \kappa$ and a little calculus we see that if $0.45 \le t \le T_1^o$

$$0 \le \int_{0.45}^t \frac{3\widehat{W} F^{s-1} \log F}{1 - \widehat{W} F^{s-1}} U_e^-(s)\, ds \le \frac{3\widehat{W}}{1-\kappa} F^{t-1} \widehat{W} \tag{2.8}$$

To clean up the last result note that if $t \le T_1^o$, $F^{t-1}\widehat{W} \le \kappa$ and the desired result follows. $\square$

### 2.4 Growth of tree agreements

In this section we show that, assuming no surprises, the number of agreements on any component $\psi$ reaches a positive fraction of issues at time $T_1 F \log F$ where $T_1$ was defined in (2.2). One can do this by using the infinitesimal analysis techniques as in Section 2.3 to bound the growth of tree agreements and use a *spanning tree* of $\psi$ to conclude the result.

#### 2.4.1 Upper bound

Let $\mathtt{t}$ be a *tree* which is a subgraph of $\mathcal{E}_+$, and let $p$ be its number of edges. Let $A_{\mathtt{t}}$ be the number of issues that all voters on the tree agree on. Our first step is to get an upper bound for $A_{\mathtt{t}}(s)$.

**Lemma 2.7.** *Let $E_0^*$ be the expected value conditional on $A_e(0)$*

$$E_0^* A_{\mathtt{t}}(s) \leq F^{ps-(p-1)} \prod_{e \in \mathtt{t}} A_e(0)$$

For $p \geq 2$ this implies that $A_{\mathtt{t}}(s) = 0$ with high probability when $s < \Delta_p = 1 - 1/p$. We will see that the upper bound is the right order of magnitude but too big by a constant factor. In what follows, for example in the proof of Lemma 2.8, it is convenient to define $\Delta_1 = 0.45$.

证明. The upper bound processes $A_e^+(t) = |J_e(t)|$ are independent and have $J_e(t) \supset I_e(t)$. $P(f \in J_e(t) | A_e^+(t)) \leq A_e^+(t)/F$, and $E_0^* A_e^+(t) = F^s A_e(0)$ so

$$E \left| \bigcap_{e \in \mathtt{t}} J_e(t) \right| \leq F (F^{s-1})^p \prod_{e \in \mathtt{t}} A_e(0)$$

and we have proved the desired result. □

In this subsection we are ignoring surprises, so growth can only occur at the end points of the path, i.e., $A_{\mathtt{t}}$ has dynamics

$$\begin{aligned} +1 \quad &\text{at rate} \quad \frac{1}{2} \sum_{e \sim L^{\mathtt{t}},\, e \in \mathtt{t}} A_e \frac{A_{\mathtt{t} \setminus \{e\}} - A_{\mathtt{t}}}{F - A_e} \log F \\ -1 \quad &\text{at rate} \quad \frac{1}{2} \sum_{e^* \sim \mathtt{t},\, e^* \notin \mathtt{t}} A_{e^*} \frac{A_{\mathtt{t}} - A_{e^* \cup \mathtt{t}}}{F - A_{e^*}} \log F \end{aligned}$$

where $e \sim L^{\mathtt{t}}$ (resp. $e \sim \mathtt{t}$) means that $e$ is connected to a leaf of $\mathtt{t}$ (reps. the tree itself $\mathtt{t}$), and the overall factor of $\log F$ comes from the fact time $s$ corresponds to time $sF \log F$ in the original process. The factor of $F$

disappears because the probability of picking an agreement is $A_e/F$. Note these rates are associated with $\mathtt{t}\backslash\{e\}$ for $e \in L^{\mathtt{t}}$, which is a tree again, and $e^* \cup \mathtt{t}$, which might be a tree or a unicyclic graph, where in the latter case $A_{e^*\cup\mathtt{t}}$ is defined analogously.

**Lemma 2.8.** *Let $W_{\mathtt{t}} = (1/p!)\prod_{e\in\mathtt{t}} W_e(0.45)$ and $\delta > 0$. With high probability*

$$\left|F^{-ps+(p-1)}A_{\mathtt{t}}(s) - W_{\mathtt{t}}\right| \le F^{-\delta/3}$$

*for all $\Delta_p + 2\delta/p \le s \le 1 - \delta/p$.*

证明. We proceed by induction on $p$. When $p = 1$ this follows from Lemmas 2.4 and 2.6. To do the induction step, note that since we stop at time $1-\delta/p$

$$\frac{1}{2}\sum_{e\sim L^{\mathtt{t}}, e\in\mathtt{t}} A_e \frac{A_{\mathtt{t}\backslash\{e\}} - A_{\mathtt{t}}}{F - A_e}\log F \approx [pW_{\mathtt{t}} + o(F^{-\delta/3p})]\cdot F^{ps-(p-1)}$$

and the upper bound in Lemma 2.7 implies that the $-A_{\mathtt{t}}$ in the rate for positive jumps and the rate for negative jumps can be ignored.

To analyze the behavior of $A_{\mathtt{t}}(s)$, we let

$$U_{\mathtt{t}}(s) = \frac{A_{\mathtt{t}}(s)}{F^{ps-(p-1)}} \qquad \text{for } s \ge \Delta_p + \delta/p$$

Computing as in (2.6) and (2.7) we conclude that when $s \ge \Delta_p$ the infinitesimal mean and variance are, omitting the error term,

$$\begin{aligned} b(U_{\mathtt{t}}) &= -p(\log F)F^{-ps+(p-1)}A_{\mathtt{t}}(s) + pF^{-ps+(p-1)}b(A_{\mathtt{t}}) \\ &\approx p\log F(0.5|L^{\mathtt{t}}|\cdot W_{\mathtt{t}} - U_{\mathtt{t}}(s)), \qquad (2.9) \\ v(U_{\mathtt{t}}) &\le p\log F\cdot F^{-ps+(p-1)}W_{\mathtt{t}} = o(F^{-\delta/3}). \qquad (2.10) \end{aligned}$$

Let $T(p,\delta) = 1 - 1/p + \delta/p$ and $P^*_{p,\delta} = P(\cdot|\mathcal{F}_{T(p,\delta)})$. By definition $-T(p,\delta)p + (p-1) = \delta/p$, so the $L^2$ maximal inequality implies that as $F \to \infty$

$$P^*_{p,\delta}\left(\sup_{t\in[T(p,\delta),T_1]}\left|U_{\mathtt{t}}(s) - U_{\mathtt{t}}(T(p,\delta)) - \int_{T(p,\delta)}^t b(U_{\mathtt{t}}(s))\,ds\right| \ge 0.5F^{-\delta/3}\right) \to 0 \qquad (2.11)$$

**Lemma 2.9.** *Suppose $y$ is continuous with*

$$\left|y(t) - y(0) + D\int_0^t (y(s) - C)ds\right| \le \varepsilon \qquad (2.12)$$

*for any* $0 \le t \le T$ *where* $C$ *and* $D > 0$ *are fixed constants. Then we have*

$$\sup_{t\in[0,T]} |y(t) - y^*(t)| \le 2\varepsilon$$

*where* $y^*(t) = y(0)e^{-Dt} + C(1 - e^{-Dt})$ *is the unique solution to the integral equation*

$$y^*(t) = y(0) - D\int_0^t (y^*(s) - C)ds.$$

Intuitively, the integral equation is close to $y'(t) = -D(y(t) - C)$, so the solution should be close to the solution of the differential equation.

证明. Let $Y(t) = \int_0^t (y(s) - C)ds$. The integral inequality (2.12) implies

$$Y'(t) + (C - y(0)) + DY(t) \le \varepsilon$$

Since $Y'(t) + DY(t) \le \varepsilon + y(0) - C$ we have

$$\frac{d}{dt}\left(e^{Dt}Y(t)\right) \le (\varepsilon + y(0) - C)e^{Dt}$$

Using $Y(0) = 0$ and integrating

$$e^{Dt}Y(t) \le (y(0) - C + \varepsilon)\int_0^t e^{Ds}ds = (y(0) - C + \varepsilon)\frac{e^{Dt} - 1}{D}$$

and we have

$$Y(t) \le (\varepsilon + y(0) - C)\frac{1 - e^{-Dt}}{D}. \tag{2.13}$$

Similarly if we start with $Y'(t) + (C - y(0)) + DY(t) \ge -\varepsilon$ we end up with

$$Y(t) \ge (-\varepsilon + y(0) - C)\frac{1 - e^{-Dt}}{D}. \tag{2.14}$$

To get the upper bound on $y(t)$, we use (2.12)) and the upper bound on in (2.14)

$$\begin{aligned} y(t) &\le \varepsilon + y(0) - DY(t) \\ &\le \varepsilon + y(0) - (-\varepsilon + y(0) - C)(1 - e^{-Dt}) \\ &= \varepsilon + y(0)e^{-Dt} + (C + \varepsilon)(1 - e^{-Dt}) + \varepsilon \le 2\varepsilon + y^*(t). \end{aligned}$$

The proof of the lower bound is similar. □

**Remark 2.1.** *The reader should note that in the first two lines of the last computation we use both* $Y'(t) + (C - y(0)) + DY(t) \le \varepsilon$ *and* $\ge -\varepsilon$ *(to get* (2.14)*), so there is no one sided version of this result.*

Back to the proof of Lemma 2.8. Using Lemma 2.9 with $D = p\log F$, $C = W_{\mathsf{t}}$ and $\varepsilon = F^{-\delta/3}$ with the formula for $y^*(t)$, we have

$$\left|U_{\mathsf{t}}(T(p,\delta)+t) - U_{\mathsf{t}}(T(p,\delta))e^{-pt\log F} + W_{\mathsf{t}}(1-e^{-pt\log F})\right| \le F^{-\delta/3}$$

By Lemma 2.7, $E_0^* U_{\mathsf{t}}(T(p,\delta)) \le C_{\mathsf{t}}$. So if $t \ge 2\delta/p$, $U_{\mathsf{t}}(T(p,\delta))e^{-pt\log F} = o(F^{-\delta/3})$, and $W_{\mathsf{t}} e^{-pt\log F} = o(F^{-\delta/3})$, which with (2.11) proves the desired result. □

### 2.4.2 Results up to $T_1$

For our results to be useful they have to hold up to $T_1 = \inf\{T : \max_e A_e^+(t) \ge \kappa F\}$.

**Lemma 2.10.** *Let $W_{\mathsf{t}}^+ = W_{\mathsf{t}} + F^{-\delta/3}$. If $\kappa > 0$ is small and if $F$ is large then with high probability*

$$F^{-pt+(p-1)} A_{\mathsf{t}}(t) \le 2W_{\mathsf{t}}^+ \quad \text{for } 1-\delta/p \le t \le T_1$$

证明. To begin, we note that the Lemma 2.8 implies

$$A_{\mathsf{t}}(1-\delta/p) \le W_{\mathsf{t}}^+ F^{1-\delta}$$

Noting that the infinitesimal variance of $A_{\mathsf{t}}$ is always $\le F\log F$ and using the $L^2$ maximal inequality, we have that as $F \to \infty$

$$P_{1-\delta/p}^* \left( \sup_{1-\delta/p \le t \le T_1} A_{\mathsf{t}}(t) - A_{\mathsf{t}}(1-\delta/p) - pW_{\mathsf{t}}^+ F \int_{1-\delta/p}^t \frac{F^{ps-p}\log F}{1-\widehat{W}F^{s-1}}\, ds \ge F^{0.6} \right) \to 0$$

For $1-\delta/p \le t \le T_1$, using $1-\widehat{W}F^{s-1} \ge 1-\kappa$ gives

$$\int_{1-\delta/p}^t \frac{F^{p(s-1)}\log F}{1-\widehat{W}F^{s-1}}\, ds \le \frac{1}{1-\kappa}\cdot\frac{1}{p}F^{pt-p}$$

Multiplying the last bound by $pW_{\mathsf{t}}^+ F$ and combining our estimates, we see that

$$A_{\mathsf{t}}(t) \le F^{0.6} + W_{\mathsf{t}}^+ F^{1-\delta} + \frac{W_{\mathsf{t}}^+}{p(1-\kappa)} F^{pt-(p-1)} \quad \text{for } 1-\delta/p \le t \le T_1$$

and the desired result follows when $\kappa < 1 - \frac{1}{2p}$. □

Now that we have upper bounds on the $A_{\mathsf{t}}(t)$ we can get lower bounds.

**Lemma 2.11.** *If $\kappa > 0$ is small enough then there are constants $c_{\mathrm{t}}(\kappa, p) > 0$ so that if $F \geq F_{\mathrm{t}}$*

$$F^{-pt+(p-1)} A_{\mathrm{t}}(t) \geq c_{\mathrm{t}} \quad \text{for } 1 - \delta/p \leq t \leq T_1.$$

证明. Again we use induction, Lemma 2.6 gives the result for $p = 1$. Noting that the infinitesimal variance of $A_{\mathrm{t}}$ is always $\leq F \log F$ and using the $L^2$ maximal inequality, we have that as $F \to \infty$

$$P^*_{1-2\delta/p}\left(\inf_{1-2\delta/p \leq t \leq T_1} A_{\mathrm{t}}(t) - A_{\mathrm{t}}(1 - 2\delta/p) - \int_{1-2\delta/p}^{t} b(A_{\mathrm{t}}(s))\, ds \leq -F^{0.6}\right) \to 0$$

From the rates we see that for $s \leq T_1$, $b(A_{\mathrm{t}}(s))/\log F$

$$\geq \frac{1}{2} \sum_{\mathrm{t} \ni e \sim L^{\mathrm{t}}} A_e \frac{A_{\mathrm{t}\setminus\{e\}} - A_{\mathrm{t}}}{F} - \frac{1}{2} \sum_{\mathrm{t} \not\ni e^* \sim \mathrm{t}} A_{e^*} \frac{A_{\mathrm{t}} - A_{e^* \cup \mathrm{t}}}{F(1-\kappa)}$$

We will drop the last term in the second sum which is nonpositive. Using Lemma 2.10 and the induction hypothesis, the above is

$$\geq \frac{1}{2} \sum_{\mathrm{t} \ni e \sim L^{\mathrm{t}}} F^s c_{e_1} \frac{F^{(p-1)s-(p-2)} c_{\mathrm{t}\setminus\{e\}}}{F} - \frac{1}{2} \sum_{\mathrm{t} \not\ni e^* \sim t, \text{or } \mathrm{t} \ni e \sim L^{\mathrm{t}}} F^s (2W^+_{e^*}) \frac{F^{ps-(p-1)}(2W^+_{\mathrm{t}})}{F(1-\kappa)}.$$

Rewriting $sq - (q-1) = q(s-1) + 1$ we see that the last expression has the form

$$F^{p(s-1)+1} a_{\mathrm{t}} - F^{(p+1)(s-1)+1} b_{\mathrm{t}}$$

for constants $a_{\mathrm{t}}, b_{\mathrm{t}}$ recursively defined as

$$a_{\mathrm{t}} = \frac{1}{2} \sum_{\mathrm{t} \ni e \sim L^{\mathrm{t}}} c_{\{e\}} c_{\mathrm{t}\setminus\{e\}}$$
$$b_{\mathrm{t}} = \frac{1}{2} \sum_{\mathrm{t} \not\ni e^* \sim t, \text{or } \mathrm{t} \ni e \sim L^{\mathrm{t}}} \frac{(2W^+_{e^*})(2W^+_{\mathrm{t}})}{1-\kappa}$$

So we have

$$\int_{1-2\delta/p}^{t} b(A_{\mathrm{t}}(s))\, ds \geq \frac{a_{\mathrm{t}}}{p}\left(F^{p(t-1)+1} - F^{1-2\delta}\right) - \frac{b_{\mathrm{t}}}{p+1} F^{(p+1)(t-1)+1}$$

Using $F^{t-1} \leq \kappa/\widehat{W}$ and $A_{\mathrm{t}}(1 - 2\delta/p) \geq 0$ the desired result follows by setting

$$c_{\mathrm{t}} < \frac{a_{\mathrm{t}}}{p} - \frac{b_{\mathrm{t}}}{p+1} \cdot \frac{\kappa}{\widehat{W}}$$

which is positive if $\kappa > 0$ is small enough. □

**Lemma 2.12.** *For any initially active component $\psi$ consensus, with high probability $A_\psi(T_1) \geq c_* F$ where $c_* = c_*(\kappa, \psi)$ is a positive constant. In words, consensus has been reached over a positive fraction of issues on $\psi$.*

证明. Let $\mathtt{t}$ be a spanning tree of the component $\psi$, let $t = T_1$ and $\kappa$ sufficiently small as in Lemma 2.11. Using (2.2) and (2.3) one can conclude that with high probability

$$A_\psi(T_1) \geq A_{\mathtt{t}}(T_1) \geq c_{\mathtt{t}} F^{pT_1-(p-1)} \geq c_{\mathtt{t}}(\kappa/\widehat{W})^p F > 0. \tag{2.15}$$

Letting $c_* = c_{\mathtt{t}}(\kappa/\widehat{W})^p$ completes the proof. □

## 2.5 Consensus on the giant component

Our goal in this section is to show that, under the assumption that there are no surprises, consensus on the giant component $\chi$ can be reached on all $F$ issues by time $T_1 F \log F + c_N F \log\log F$. To do such in general settings, we switch to study Axelrod's model on a given component $\psi$ running at rate $F$ (short as *Axelrod's dynamics*) starting from an initial configuration where the number of issues $\psi$ agree on is $\geq \delta F$. Define $\sigma_f$ as the first time Axelrod's dynamics reaches consensus over issue $f$ on $\psi$. The main Lemma in this section is

**Lemma 2.13.** *For Axelrod's dynamics on $\psi$ that has reached consensus over $\geq \delta F$ issues at time 0 where $\delta > 0$ is a constant, there is a constant $C$ so that the first time $\psi$ reaches consensus on all issues*

$$\max_f \sigma_f \leq (CN^6/\delta) \log\log F \tag{2.16}$$

*with high probability as $F \to \infty$*

**Remark 2.2.** *Under the assumption of no surprises, Lemma 2.12 implies that Axelrod's model on the giant component $\chi$ possess the same opinion on $c_*(\kappa, \chi)F$ issues at time $T_1 F \log F$. Since our model is Markovian we can apply Lemma 2.13 with $\delta = c_*(\kappa, \chi)$ which indicates that consensus on the giant component $\chi$ can be reached over all $F$ issues by time $T_1 F \log F + c_N F \log\log F$, where $c_N = CN^6/c_*(\kappa, \chi) > 0$ does not depend on $F$, completing the argument.*

We start the proof of Lemma 2.13 by showing

**Lemma 2.14.** *For each $f = 1, \ldots, F$ let*

$$t_0 = t_0(\delta, N) = \frac{N^6}{\delta}. \tag{2.17}$$

*There is an enlarged probability space where we can construct independent $(\tilde{\sigma}_f : f = 1, \ldots, F)$ such that $\sigma_f \leq \tilde{\sigma}_f$ for all $f = 1, \ldots, F$, and for $n = 0, 1, 2, \ldots$*

$$P\left(\tilde{\sigma}_f > n t_0\right) \leq \left(\frac{1}{2}\right)^n. \tag{2.18}$$

**Remark 2.3.** *We prove Lemma 2.14 by assume first that there are only two opinions held by voters of $\psi$ over each issue and prove*

$$P\left(\tilde{\sigma}_f > n\frac{N^4}{\delta}\right) \leq \left(\frac{1}{2}\right)^n. \tag{2.19}$$

*then generalize our argument to the case of multiple opinions to conclude* (2.19).

证明. *The two-opinion case.* We assume without loss of generality that the opinion set on each issue is $\{0,1\}$. To prove (2.19), note that in Axelrod's dynamics we pick an issue on which there is a disagreement at random, however by the idea of Poisson thinning we achieve the same dynamics by instead picking at random from all issues at rate

$$F \cdot \frac{A_e(t-)}{F} \cdot \frac{1}{(F - A_e(t-)) \vee 1} = \frac{A_e(t-)}{(F - A_e(t-)) \vee 1} \tag{2.20}$$

since when an issue is picked on which there is agreement nothing happens.

We proceed and complete the proof by constructing a coupling between Axelrod's dynamics and a stack of independent random walks. Imitation events for issue $f = 1, \ldots, F$ on edge $e$ occur at rate $\frac{A_e(t-)}{F - A_e(t-)}$. When imitation happens the opinion is either copied or eliminated with probability $1/2$ for each, independent of the dynamics before time $t$. Set

(1) $(\Lambda^0_{f,k}(t) : t \geq 0, f = 1, \ldots, F, k = 1, \ldots, |\mathcal{E}|)$ be a collection of unit Poisson processes independent across $f = 1, \ldots, F$ and $k \in [|\mathcal{E}|]$, and

(2) $(X^n_f : n \geq 1)$ be a collection of independent fair coins that take value $\{-1, +1\}$ with probabilities $1/2$ each, independent across $f = 1, \ldots, F$.

With these tools at hand one can recursively build up Axelrod's dynamics as what follows. Let $\tau_0 = 0$. For each $n = 0, 1, \ldots$ suppose the process has been defined for $t \in [0, \tau_n]$, we define the next phase of process as

- Let $D^n_f$ denotes the set of edges connected by two voters holding different opinions for issue $f$. $D^n_f$ is nonempty whenever $\tau_n < \sigma_f$. Let $\iota^n_f : [|\mathcal{E}|] \to \mathcal{E}$ be a bijective map satisfying $\iota^n_f(1) \in D^n_f$. Such map is adapted to the process filtration at $\tau_n$.
- Define $\tau^*$ as the minimal jump time of $(\Lambda_{f,e}(t) : t \geq 0)$ among all $f = 1, \ldots, F$ and $e \in \mathcal{E}$, where for $t \in (\tau_n, \tau^*]$

$$\Lambda_{f,\iota^n_f(k)}(t) = \Lambda_{f,\iota^n_f(k)}(\tau_n) + \Lambda^0_{f,k}\left(\mathbf{B}_k + \frac{A_{\iota^n_f(k)}(\tau_n)}{(F - A_{\iota^n_f(k)}(\tau_n)) \vee 1}(t - \tau_n)\right) - \Lambda^0_{f,k}(\mathbf{B}_k).$$

where $\mathbf{B}_k$ stands for the expression

$$\sum_{n'=0}^{n-1} \frac{A_{\iota_f^{n'}(k)}(\tau_{n'})}{(F - A_{\iota_f^{n'}(k)}(\tau_{n'})) \vee 1} (\tau_{n'+1} - \tau_{n'}).$$

which is $\geq \tau_n \cdot \delta$. Let $e^*$ denote the edge whose associated $\Lambda$-process achieves the minimal jump time.

- If two voters connecting $e^*$ share the same opinion on issue $f$, no dynamics occurred at $\tau^*$ so we set $\tau_n = \tau^*$ and return to Bullet 1. Otherwise set $\tau_{n+1} = \tau^*$.
- At $t = \tau_{n+1}$, we flip a coin $X_f^{n+1}$ to determine the direction of imitation on $e^*$ for issue $f$ of Axelrod's dynamics. If $X_f^{n+1}$ is head then the opinion pattern on $e^*$ flips from $10 \to 11$, i.e. the voter which holds opinion 0 takes the opinion of its neighbor acroos $e^*$; otherwise the pattern flips from $10 \to 00$.
- Update $(A_e(\tau_{n+1}) : e \in \mathcal{E})$.

We are making our coupling very explicit in order to be able to define an associated process $M_f(t)$ = the number of individuals holding opinion 1 for issue $f$. Let

$$\Lambda_f(t) = \sum_{n:\tau_n \leq t} \left( \sum_{e \in D_f^n} \Delta\Lambda_{f,e}(\tau_n) \right),$$

which counts all the jumps occurred on the edge set $D_f^n$. From the previous definition of process

$$\Lambda_f(t) \geq \Lambda_{f,1}^0 \left( \mathbf{B}_1 + \frac{A_{\iota_f^n(1)}(\tau_n)}{(F - A_{\iota_f^n(1)}(\tau_n)) \vee 1} (t - \tau_n) \right) \geq \Lambda_{f,1}^0 (\delta t). \qquad (2.21)$$

Set

$$M_f(t) = M_f(0) + \sum_{n=1}^{\Lambda_f(t)} X_f^n.$$

Since there are only two opinions our consensus time $\sigma_f$ is equal to the time $M_f(t)$ first hits the boundary of $[0, |\psi|]$. To define coupled random walks let $\tilde{\Lambda}_f(t) = \Lambda_{f,1}^0 (\delta t)$ and $\tilde{\sigma}_f$ be the first time $t$ that

$$\tilde{M}_f(t) = M_f(0) + \sum_{n=1}^{\tilde{\Lambda}_f(t)} X_f^n$$

hits the boundary of $[0, |\psi|]$. We have $\tilde{\sigma}_f$ being independent across $f = 1, \dots, F$, and because of (2.21), $\sigma_f \le \tilde{\sigma}_f$ for all $f = 1, \dots, F$. It is obvious that $\tilde{M}_f(t)$ is a time-change of simple random walk on $[0, |\psi|]$ with $|\psi| \le N^2$. Classical gambler's ruin result implies that independent of initial value of $\xi_0$

$$E_{\xi_0} \tilde{\sigma}_f \le \delta^{-1} \frac{|\psi|^2}{4} \le \frac{N^4}{4\delta} \tag{2.22}$$

From the Markov property of Axelrod's dynamics, Markov's inequality and (2.22)

$$\begin{aligned} P\left(\tilde{\sigma}_f > n \cdot \delta^{-1} N^4 \middle| \tilde{\sigma}_f > (n-1) \cdot \delta^{-1} N^4\right) &\le \sup_{\xi_0} P_{\xi_0}\left(\tilde{\sigma}_f > \delta^{-1} N^4\right) \\ &\le \left(\delta^{-1} N^4\right)^{-1} \left(\sup_{\xi_0} E_{\xi_0} \tilde{\sigma}_f\right) \le \frac{1}{2} \end{aligned} \tag{2.23}$$

where supremum is taken among all initial $\xi_0$. (2.19) follows immediately by taking products for (2.23) for $n = 1, 2, \dots$ □

*Proof of Lemma 2.14 in the general case.* For the multiple opinions case let $I$ be the ordered set of opinions $q_1, \dots, q_{|I|}$ held by all individuals on issue $f$. To generalize our previous arguments we

- Map the opinions $\{q_1\}$ and $I \backslash \{q_1\}$ to the new opinions $\bar{1}$ and $\bar{0}$;
- Let $M_f(t)$ = the number of voters taking the new $\bar{1}$. Then $M_f(t)$ absorbs at the boundary of $[0, |\psi|]$ at $\sigma_f^{(1)}$, which is upper bounded by $\tilde{\sigma}_f^{(1)}$ as in (2.19);
- If $M_f(t)$ is absorbed at $N^2$ consensus is reached and we are done. Otherwise we reduce the opinion set to $I \backslash \{q_1\}$ and start a new round by mapping the opinion sets $\{q_2\}$ and $I \backslash \{q_1, q_2\}$ as the new opinions $\bar{1}$ and $\bar{0}$ and return the first bullet.

Keep the procedures until $M_f(t)$ hits $N^2$ before 0 for some round $K$. Since $|I| \le N^2$, $K \le N^2 - 1$ and the sum $\tilde{\sigma}_f \equiv \sum_{k=1}^{N^2-1} \tilde{\sigma}_f^{(k)}$ bounds $\sigma_f$ from above with $\tilde{\sigma}_f$ being independent across $f = 1, \dots, F$, (2.22) indicates that $E\tilde{\sigma}_f \le (4\delta)^{-1} N^6$. Markov property along with Markov's inequality implies for $t_0$ defined in (2.17)

$$P\left(\tilde{\sigma}_f > n \cdot t_0 \middle| \tilde{\sigma}_f > (n-1) \cdot t_0\right) \le t_0^{-1} \left(\sup_{\xi_0} E_{\xi_0} \tilde{\sigma}_f\right) \le \frac{1}{2} \tag{2.24}$$

Lemma 2.14 follows by taking products. □

**Remark 2.4.** *To explain our proof of* (2.19) *note for an issue that* $\psi$ *reaches consensus over, arrows do not pose effects on edges of agreement. When the opinions are different we read* $10 \to 11$ *as flipping out a head* $(+1)$ *and* $10 \to 00$ *as a tail* $(-1)$*, recorded by* $X_1, X_2, \ldots$*. These fair coins are then recycled to couple the 1-counters of Axelrod's dynamics for a fixed issue with uniformly slower random walks, and a sequence of independent upper-bounding variables for the consensus time can be constructed thereafter.*

*Proof of Lemma 2.16.* Referring to (2.17) for the definition of $t_0$ we define a discrete-time Markov chain $\underline{A}_n$ to give a lower bound on $A_\psi(nt_0)$. It starts from $\underline{A}_0 = 0$ and has transition probability

$$P(\underline{A}_{n+1} = k + i | \underline{A}_n = k) = \binom{F-k}{i} \left(\frac{1}{2}\right)^{F-k} \quad \text{for } 0 \le i \le F - k.$$

From Lemma 2.14 there exists $t_0 > 0$ such that there exists a coupling $(A_\psi(nt_0), \underline{A}_n)$ such that $A_\psi(nt_0) \ge \underline{A}_n$. It is clear that

$$E(\underline{A}_{n+1} | \underline{A}_n = k) = k + \frac{1}{2}(F - k).$$

Let $\underline{b}_n = E(F - \underline{A}_n)$. Taking expectation in the last display we obtain $\underline{b}_{n+1} = \frac{1}{2}\underline{b}_n$ so

$$EA_\psi(nt_0) \ge E\underline{A}_n = F\underline{b}_n \ge F(1 - 2^{-n})$$

Taking $n = C \log\log F$, we have

$$E\big(F - A_\psi(Ct_0 \log\log F)\big) \le F \left(\frac{1}{2}\right)^{C \log\log F}.$$

One takes $C = 3/\log 2 > 0$ to obtain $EA_\psi(Ct_0 \log\log F) \ge (1 - (\log F)^{-2})F$. Since $F - A_\psi(Ct_0 \log\log F)$ is nonnegative, Markov's inequality implies

$$P\left(F - A_\psi(Ct_0 \log\log F) \ge (\log F)^{-2} F\right) \le \frac{(\log F)^{-3} F}{(\log F)^{-2} F} = (\log F)^{-1} \tag{2.25}$$

which goes to 0 as $F \to \infty$. One can then complete the proof of Lemma 2.13 using Lemma 2.14 and

**Lemma 2.15.** *We have*

$$P\left(\max_f \tilde{\sigma}_f - Ct_0 \log\log F \ge \frac{2t_0}{\log 2} (\log F)^{-1}\right) \le \frac{1}{F}. \tag{2.26}$$

证明. (2.25) implies that w.h.p. after time $Ct_0 \log\log F$, each of the individual voter models in the coupling is running at rate $\geq (\log F)^2$. By Lemma 2.14 the $\tilde{\sigma}_f$ are independent and the tail of distribution bounded geometrically. From this it follows that

$$P\left(\max_f \tilde{\sigma}_f - T_* \geq (\log F)^{-2} K t_0\right) \leq F(1/2)^K$$

Taking $K = (2/\log 2)\log F$ we obtain the desired result. □

# 3 Surprises

Having completed the argument assuming no surprises, we now return to the real dynamics where surprise events can happen. The crucial consequence of surprises is an edge with no initial agreement can be *activated by surprises* in later phases. As the readers will see, suprises will occur before time $T_1$ with probability $1-O(\kappa)$, and the agreements on edges created by surprises are $\leq (\log F)^\alpha$ until consensus is reached on the giant component $\chi$ over all undisturbed issues. We begin with the case of one surprise. In this situation our argument has two cases: the surprise adds one edge to the giant component, or connects it to an existing component.

## 3.1 The probability of no surprises before $T_1 F \log F$

We say that *coincidence* occurs initially on issue $f$ if there exist distinct voters $i, j$ which are not neighbors such that $\xi_0(i,f) = \xi_0(j,f)$. Issues where coincidence occurs initially are the only ones that can produce surprises. Without much difficuty one can prove

**Lemma 3.1.** *Let $\mathcal{N}_\mathcal{S}$ be the set of issues where coincidence occurs initially and suppose $F \leq \gamma Q$. We have $E|\mathcal{N}_\mathcal{S}| \leq \gamma N^4/2$.*

证明. Let $\mathcal{C}_f = \{\text{coincidence occurs initially on issue } f\}$. There are $\binom{N^2}{2}$ pairs of vertices which have the same opinion on issue $f$ with probability $1/Q$ so

$$P(\mathcal{C}_f) \leq \frac{1}{Q}\binom{N^2}{2} \tag{3.1}$$

Hence Lemma 3.1 follows from (3.1) and

$$E|\mathcal{N}_\mathcal{S}| = \sum_{f=1,\ldots,F} P(\mathcal{C}_f) \leq \frac{F(N^2)^2}{2Q} \leq \gamma N^4/2.$$

□

Our main Lemma in this section claims that if $\kappa > 0$ is small, with probability $1-O(\kappa)$ there are no surprises before $T_1 F \log F$.

**Lemma 3.2.** *Let $H_*(\kappa, F)$ be the event that there is no surprise event before time $T_1 F \log F$. Then $P(H_*(\kappa,F)) \geq 1 - \gamma\kappa N^6$.*

证明. We refer the readers to Section 2.2 for the definitions of $I_e(t)$ and $J_e(t)$. Let $P^+$ be the probability measure conditioned on the initial configuration. The probability that some surprise event occurs is

$$\begin{aligned} P^+(H_*(\kappa,F)^c) &\le P^+\left(\mathcal{N}_{\mathcal{S}} \cap \left(\bigcup_{e\in\mathsf{t}} I_e(T_1)\right) \neq \varnothing\right) \\ &\le P^+\left(\mathcal{N}_{\mathcal{S}} \cap \left(\bigcup_{e\in\mathsf{t}} J_e(T_1)\right) \neq \varnothing\right) \\ &\le \sum_{f_{\mathcal{S}}\in\mathcal{N}_{\mathcal{S}}} P^+\left(f_{\mathcal{S}} \in \bigcup_{e\in\mathsf{t}} J_e(T_1)\right) \le |\mathcal{N}_{\mathcal{S}}| \cdot \frac{2\kappa N^2 F}{F}. \end{aligned}$$

since $\cup_{e\in\mathsf{t}} J_e(t)$ is a uniform subset of $\{1,\ldots,F\}$ and has cardinality $\le \kappa|\mathcal{E}|F = 2\kappa N^2 F$. Lemma 3.1 implies

$$P(H_*(\kappa,F)) = 1 - EP^+(H_*(\kappa,F)^c) \ge 1 - E|\mathcal{N}_{\mathcal{S}}| \cdot 2\kappa N^2 \ge 1 - \gamma\kappa N^6.$$

□

**Remark 3.1.** *In the later sections we assume $H_*(\kappa,F)$ occurs, and the probability and expectation notations $P$ and $E$ stand for the conditional ones on $H_*(\kappa,F)$. Lemma 3.2 implies that the conditioning imposes little effect on the distributions since the probability is bounded away from 0 as $F\to\infty$.*

### 3.2 One surprise

Remark 2.2 in the beginning of Section 2.5 indicates that in the case of no surprises, Axelrod's model reach consensus on all issues on $\chi$ along with all other components at time $T_1 \log F + c_N \log\log F$. In this section we study the growth of one surprise edge connecting to an individual not in the giant component $\chi$.

We continue, as in the beginning of Section 2.5 to run Axelrod's model at rate $F$ after $T_1$. Let $e^*$ be the edge activated by a surprise and connected the main component $\psi$ to either a single voter or a component $\psi'$. Also in the first (resp. second) case, assume for $\psi$ (resp. $\psi^* = \psi, \psi'$), at time $T^*$ consensus has reached on $\psi$ (resp. $\psi^*$) over $\geq \delta F$ issues where $\delta > 0$ is a constant. Let $\tau_{e^*}$ be the first time an agreement on $e^*$ emerges. Define the new component $\mathcal{C} = \psi \cup \{e^*\} \cup \psi'$ (To unify both cases we force to write $\psi' = \varnothing$ in the case of single voter) and

$$T^o_{e^*} = \inf\{t \geq \tau_{e^*} : A_{\mathcal{C}}(t) \geq 1\}, \qquad T^w_{e^*} = \inf\{t \geq \tau_{e^*} : A_{e^*}(t) = 0\}$$

If $T^w_{e^*} < T^o_{e^*}$ then agreement on the surprise edge $e^*$ is lost, so $\psi$ is once again disconnected with $\psi'$, and it has to be reactivated by a new surprise. Otherwise we are in the case that the issues of agreements on $\mathcal{C}$ can grow. By restarting the process the former case can be combined with the latter one, and we may without loss of generality assume that latter is the case. Let $\sigma^a_f$ be the first time that *activation time*, i.e. the first time that an agreement on $e^*$ is generated on issue $f$. Let

$$\mathcal{J}_{e^*}(t) = \inf\{t : \sigma^a_f \leq t\}.$$

In words, $\mathcal{J}_{e^*}(t)$ denotes the set that records the issues over dynamics on $e^*$ occurred by time $t$. We are able to conclude that in both cases, the full component $\mathcal{C}$ reaches positive level $\delta_{e^*} F$ before the number of issues for which there is agreement on $\psi$ or $\psi'$ drops below $(\delta/2)F$. Using the coupling with a branching process we can also conclude that once an issue has been activated on $\mathcal{C}$ it takes $O(\log F)$ time for the number of issues $\mathcal{C}$ agree on to reach positive level.

### 3.2.1 Surprises in single-edge case

In this case the surprise edge $e^*$ does not connect $\psi$ to any component. We introduce for $\kappa > 0$

$$T_{e^*}(\kappa) = \inf\{t \geq T_* : |\mathcal{J}_{e^*}(t)| \geq \kappa F\}$$
$$T_2(\delta_0) = \inf\{t \geq T_* : A_{\psi \cup e^*}(t) \geq \delta_0 F\}$$

**Lemma 3.3.** *For $\kappa > 0$ and $\delta_0 > 0$ appropriately chosen, with high probability*

$$T_2(\delta_0) \leq T_{e^*}(\kappa).$$

证明. Since growth on $e^*$ is the only source of disturbance of the consensus on $\psi$, a little thought reveals that $A_\psi(t) - A_{\psi\cup e^*}(t)$, the number of issues agreed on $\psi$ but not on $e^*$, has

$$A_\psi(t-) - A_{\psi\cup e^*}(t-) \geq A_\psi(T_*) - |\mathcal{J}_{e^*}(t)|. \tag{3.2}$$

for all $t \geq T_*$. When $\kappa$ is chosen to be $\delta/2$ the above is $\geq (\delta/2)F$ for $t \in [T_*, T_{e^*}(\kappa)]$. Consider the edge activated by surprise $e^* = \{i, j\}$ where $i \in \psi$ and $j \notin \psi$. We use $\mathcal{I}_\psi(t)$ to denote the set of issues individuals of $\psi$ agree on at time $t$, so that $A_\psi(t) = |\mathcal{I}_\psi(t)|$. At time $t$ where $\mathcal{J}_{e^*}(t)$ adds issue $f$ to itself, either

(1) $i$ gives its opinion to $j$ and $f \in \mathcal{I}_\psi(t-)$, and $A_{\psi\cup e^*}(t)$ increases by 1, or

(2) $j$ gives its opinion to $i$ or $f \notin \mathcal{I}_\psi(t-)$, $A_{\psi\cup e^*}(t)$ does not increase.

By time $t \leq T_{e^*}(\kappa)$ case (i) above has probability

$$= \frac{1}{2}\frac{A_\psi(t-) - A_{\psi\cup e^*}(t-)}{F - A_{e^*}(t-)} \geq \frac{(\delta/2)F}{2F} = \frac{\delta}{4},$$

so the agreement on $\psi \cup \{e^*\}$ at $T_{e^*}$ is bounded below by Binomial$(\kappa F, \delta/4)$. To complete the proof we use the following classical large deviation result for binomial distribution

**Lemma 3.4.** *Let $X_1, \ldots, X_n$ be i.i.d. with $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$. Then*

$$P(X_1 + \cdots + X_n < n(p - \varepsilon)) \leq \exp(-\varepsilon^2 n/2).$$

*When $\varepsilon = p/2$ this is just*

$$P\left(\mathit{Binomial}(n, p) < (p/2)n\right) \leq \exp\left(-(p^2/8)n\right).$$

证明. If $\alpha > 0$ then

$$P(X_1 + \cdots + X_n \leq n(p-\varepsilon))e^{-\alpha n(p-\varepsilon)} \leq (pe^{-\alpha} + (1-p))^n$$

Taking log's, dividing by $n$, rearranging and then using $\log(1+x) \leq x$ we have

$$\begin{aligned} \frac{1}{n}\log P(X_1 + \cdots + X_n \leq n(p-\varepsilon)) &\leq \alpha(p-\varepsilon) + \log(1 + p(e^{-\alpha} - 1)) \\ &\leq \alpha(p-\varepsilon) + p(e^{-\alpha} - 1) = -\alpha\varepsilon + p(e^{-\alpha} - 1 + \alpha) \end{aligned}$$

Now $e^{-\alpha} - 1 + \alpha \leq \alpha^2/2$ for $0 < \alpha < 1$ so taking $\alpha = \varepsilon$ and using $p \leq 1$ gives

$$P(X_1 + \cdots + X_n < n(p-\varepsilon)) \leq \exp(-\varepsilon^2 n/2)$$

and completes the proof of Lemma 3.4. □

Now applying Lemma 3.4 with $n = \kappa F$, $p = \delta/4 = 2\varepsilon$ and assume $\kappa \leq 1/8$ without loss of generality we obtain

$$P\left(A_{\psi\cup e^*}(T_{e^*}(\kappa)) \geq (\kappa\delta/8)F\right) \geq P\left(\text{Binomial}\left(\kappa F, \delta/4\right) \geq (\kappa\delta/8)F\right) \geq 1 - \exp\left(-(\delta^2/128)\kappa F\right).$$

Letting $\delta_0 = \kappa\delta/8$ completes the proof of Lemma 3.3. □

To estimate $T_2$ the time reaching consensus on $\psi \cup \{e^*\}$ over $\delta_0 F$ issues, we have

**Lemma 3.5.** *With high probability $T_2(\delta_0) \leq C(\delta)\log F$ as $F \to \infty$.*

证明. From the proof of Lemma 3.3 we see that for time $t \leq T_{e^*}$, $A_{\psi\cup e^*}(t)$ can be coupled with $A^-_{\psi\cup e^*}(t)$, a Yule binary fission which starts from $A^-_{\psi\cup e^*}(T_*) = 1$ running at rate $\delta/4$, so that

$$A_{\psi\cup e^*}(t) \geq A^-_{\psi\cup e^*}(t), \qquad T_* \leq t \leq T_{e^*}. \tag{3.3}$$

Using the standard results of branching process (see e.g. [1]) the first time $A^-_{\psi\cup e^*}(t)$ reaches $\delta_0 F$, denoted by $\widehat{T}_{e^*}$, is equal to $T_* + (\delta/4)^{-1}(\log(\delta_0 F) + V^F)$ where $V^F$ converges weakly to $V^\infty$ with cumulative distribution function $P(V^\infty \leq x) = \exp(-\exp(-x))$ for $x \in \mathbb{R}$. Therefore for $C > (\delta/4)^{-1}$ the weak convergence implies

$$P\left(\widehat{T}_{e^*} - T_* \geq C\log F\right) = P\left(V^F \geq [(\delta/4)C - 1]\log F - \log\delta_0\right) \to 0 \text{ as } F \to \infty. \tag{3.4}$$

One can then finish the argument by combining Lemma 3.3, (3.4) and

$$T_{e^*}(\kappa) \le \widehat{T}_{e^*}. \tag{3.5}$$

(3.5) can be proved by contradiction as follows. Suppose $T_{e^*}(\kappa) > \widehat{T}_{e^*}$ then the coupling in (3.3) works for $t = \widehat{T}_{e^*}$ and thereby $A_{\psi\cup e^*}(\widehat{T}_{e^*}) \ge A^-_{\psi\cup e^*}(\widehat{T}_{e^*}) \ge \delta_0 F$, contradicting the definition of $T_{e^*}(\kappa)$. □

#### 3.2.2 Surprises occurs in middle of two components

Let $\sigma_f^*$ be the time that Axelrod's model reaches consensus (and hence fixation) on $\mathcal{C}$ for issue $f$, i.e. all members of the new component $\mathcal{C}$ share the same opinion on issue $f$, $|\psi'|$ denote the number of vertices in $\psi'$, $t_1^*(\kappa,\psi') = 2|\psi'|^2/\kappa$ and define the following stopping times

$$\begin{aligned}
\tau_\psi(\kappa) &= \inf\{t \ge T_* : A_\psi(t) < (\delta/2)F\} \\
\tau_{\psi'}(\kappa) &= \inf\{t \ge T_* : A_{\psi'}(t) < (\delta/2)F\} \\
T_3^*(\kappa) &= \inf\{t > 0 : |\mathcal{J}_{e^*}(t)| \ge (\kappa/2)e^{-t_1^*}F\}
\end{aligned}$$

**Lemma 3.6.** *For $\kappa > 0$ and $\delta_{e^*} > 0$ appropriately chosen such that with high probability as $F \to \infty$ $T_3^*(\kappa) < \infty$, and*

$$T_3(\delta_{e^*}) \le T_3^*(\kappa) + t_1^*(\kappa,\psi') \le \min\{\tau_\psi(\kappa), \tau_{\psi'}(\kappa)\}. \tag{3.6}$$

The idea of proof is to couple the counter $A_{\mathcal{C}}(t)$ for $t \le T_*$ with a lower-bounded one that is easier to analyze. Consider vertex $x^*$ whose opinion over an issue $f$ at its activation time $\sigma_f^a$ takes over $\psi'$. Let $G_f^{\mathcal{C}}$ denote the event that consensus is reached by time $T_3^* + t_1^*$, i.e.

$$G_f^{\mathcal{C}} = \{\sigma_f^* \le T_3^* + t_1^*,\ \xi_{\sigma_f^*}(f,x) = \xi_{\sigma_f^a}(f,x^*)\ \forall x \in \mathcal{C}\}.$$

Lastly let

$$\eta_{e^*}(\kappa,\psi') = (\delta/4)\frac{1}{2|\psi'|}. \tag{3.7}$$

**Lemma 3.7.** *For $\kappa > 0$ appropriately chosen, there is an enlarged probability space where set of independent events $(\tilde{G}_f^{\mathcal{C}} : f = 1,\dots,F)$ live, satisfying*

*(i)* $G_f^{\mathcal{C}} \supseteq \tilde{G}_f^{\mathcal{C}}$ *for* $f \in \mathcal{J}_{e^*}(T_3^*(\kappa))$*;*

*(ii)* $P\left(\tilde{G}_f^{\mathcal{C}}\right) \ge \eta_{e^*}(\kappa,\psi')$*.*

证明. Let $\{i,j\} = e^*$ with $i$ in the giant component $\psi$, and $\kappa = \delta/2$. To explain why $\eta_{e^*}$ defined in (3.7) is a lower bound, note the first factor is no greater than the probability that consensus has been reached on $\psi$ on the issue that $i$ gives her/his opinion to $j$, and hence

$$= \frac{1}{2}\frac{A_\psi(t-) - A_{\psi\cup e^*}(t-)}{F - A_{e^*}(t-)} \geq \frac{(\delta/2)F}{2F} = \delta/4,$$

once again due to (3.2). The second factor is the lower-bound probability that the opinion successfully spread to all members of $\psi'$ and therefore consensus on $\mathcal{C}$ is reached over issue $f$. We utilize the techniques provided in the proof of Lemma 2.14 and couple the dynamics with independent martingales $(M_f^{\mathcal{C}}(t), t \geq \sigma_f^a : f = 1, \ldots, F)$ which performs random walks at constant rate $\kappa \leq \kappa/(1-\kappa)$ on $\{0, 1, \ldots, |\psi'|\}$ with absorbing boundaries. With $(\tilde{\sigma}_f^* : f = 1, \ldots, F)$ defined as independent stopping time $M_f^{\mathcal{C}}(t)$ reaches $|\psi'|$, standard coupling argument gives $\sigma_f^* \leq \tilde{\sigma}_f^*$ for all $f = 1, \ldots, F$. Let

$$\tilde{G}_f^{\mathcal{C}} = \{\tilde{\sigma}_f^* - \sigma_f^a \leq t_1^*,\ M_f^{\mathcal{C}}(\tilde{\sigma}_f^*) = |\psi'|\}$$

Apparently $G_f^{\mathcal{C}} \supseteq \tilde{G}_f^{\mathcal{C}}, f \in \mathcal{J}_{e^*}(T_3^*)$ form independent events and hence

$$P\left(\tilde{G}_f^{\mathcal{C}} | \mathcal{F}_{\sigma_f^a}\right) \geq P\left(M_f^{\mathcal{C}}(\sigma_f^a + t_1^*) = |\psi'| | \mathcal{F}_{\sigma_f^a}\right). \tag{3.8}$$

The right-hand side is just the probability a random walk $S_t$ jumping at rate $\kappa$ starting from 1 get absorbed at $|\psi'|$ before at 0. Let $\tau_{0,|\psi'|} = \inf\{t : S_t = 0 \text{ or } |\psi'|\}$, so using standard results of random walk the RHS in (3.8)

$$\geq P_1(S_{\tau_{0,|\psi'|}} = |\psi'|, \tau_{0,|\psi'|} \leq t_1^*) \geq P_1(S_{\tau_{0,|\psi'|}} = |\psi'|) - P_1(\tau_{0,|\psi'|} \geq t_1^*)$$
$$\geq \frac{1}{|\psi'|} - \frac{E_1\tau_{0,|\psi'|}}{t_1^*} = \frac{1}{|\psi'|} - \frac{(|\psi'|-1)/\kappa}{t_1^*} \geq \frac{1}{2|\psi'|}.$$

for all $f \in \mathcal{J}_{e^*}(T_3^*)$. Taking expectation again we obtain the result. □

*Proof of Lemma 3.6.* In our setting once consensus on $\mathcal{C}$ for at least one issue is reached it will not be diminished, and hence finite Markov chain theory tells that $T_3 < \infty$, leaving the rest to prove (3.6). We continue to choose $\kappa = \delta/2$, and an issue $f$ the dynamics on $e^*$ feeds in an opinion so the consensus on $\psi$ breaks down, and the consensus will be reached with probability bounded away from 0 by time $t_1^*$. For simplicity we give only one chance for each single issue and the recovery process takes $t_1^*$ time with probability of success $\geq \eta_{e^*} > 0$. Using (3.2) we have for all $t \geq T_*$

$$\min\{A_\psi(t), A_{\psi'}(t)\} \geq \delta F - |\mathcal{J}_{e^*}(t)| \geq (\delta/2)F$$

which implies $T_3^* + t_1^* \le \min\{\tau_\psi, \tau_{\psi'}\}$. Thus the branching process upper bound indicates that there exist $G_1, \dots, G_n$ i.i.d. with geometric distribution of probability $e^{-t_1^*}$ such that $|\mathcal{J}_{e^*}(T_3^* + t_1^*)| \le \sum_{i=1}^{(\kappa/2)e^{-t_1^*}F} G_i$. Chebyshev's inequality implies that if $Z, Z_1, \dots, Z_n$ are independent with common distribution geometric$(p)$

$$P\left(\frac{1}{n}\sum_{i=1}^n Z_i \ge 2EZ\right) \le P\left(\left|\frac{1}{n}\sum_{i=1}^n Z_i - EZ\right| \ge EZ\right) \le \frac{\mathrm{var}(Z)}{n(EZ)^2} = \frac{(1-p)p^{-2}}{np^{-2}} \le \frac{1}{n},$$

so with probability $\ge 1 - Ce^{t_1^*}F^{-1}$ we have $|\mathcal{J}_{e^*}(T_3^* + t_1^*)| \le (\kappa/2)e^{-t_1^*}F \cdot 2e^{t_1^*} = \kappa F$ which proves the first inequality in (3.6). Regarding to the second one note as long as $\min\{A_\psi(t), A_{\psi'}(t)\} \ge (\delta/2)F$ we can use the arguments in Section 2.5 and couple with independent random walks indexed by $f = 1, \dots, F$. Lemma 3.7 allows us to couple $A_{\mathcal{C}}(T_3^* + t_1^*)$ with a binomial variable

$$Y \sim \mathrm{Binomial}((\kappa/2)e^{-t_1^*}F, (2|\psi'|)^{-1}), \text{ so that } A_{\mathcal{C}}(T_3^* + t_1^*) \ge Y.$$

Applying Lemma 3.4 with $n = (\kappa/2)e^{-t_1^*}$ and $p = 2\varepsilon = (2|\psi'|)^{-1}$ we obtain $A_{\mathcal{C}}(T_3^* + t_1^*) \ge \delta_{e^*}F$ with probability $\ge 1 - \exp\left(-(\kappa|\psi'|^{-2}/64)e^{-t_1^*}F\right)$ for $\delta_{e^*} := (\kappa/8)|\psi'|^{-1}e^{-t_1^*}$, completing the proof. □

**Lemma 3.8.** *There exists $C(\kappa, \psi') > 0$ such that with high probability as $F \to \infty$ $T_3(\delta_{e^*}) \le C(\kappa, \psi') \log F$.*

证明. The process $A_{\mathcal{C}}(t)$ can be lower bounded by a discrete-time branching process where $\underline{A}_n$ with $\underline{A}_0 = 1$ and offspring distribution $1 + \mathrm{Bernoulli}(\tilde{\eta}_{e^*})$ where $\tilde{\eta}_{e^*} = (1 - e^{-t_1^*})\eta_{e^*}$. Coupling result indicates

$$\underline{A}_n \le A_{\mathcal{C}}(t_1^* \cdot 2n) \tag{3.9}$$

Note if we choose $C_1 > 0$ such that $(1 + \tilde{\eta}_{e^*})^{C_1} = 1.1$, our case that the offspring distribution has no probability mass on 0, finite second moment and the expectation $> 1$, so from [1]

$$W^F = \frac{\underline{A}_{C_1 \log F}}{(1 + \tilde{\eta}_{e^*})^{C_1 \log F}}$$

which converges almost surely (and hence weakly) to a positive variable $W^\infty$. Thus

$$P(\underline{A}_{C_1 \log F} \le \delta_{e^*}F) = P(W^F \le \delta_{e^*}F^{-0.1}) \to 0.$$

Using coupling in (3.9), w.h.p.

$$T_3(\delta_{e^*}) \leq t_1^* \cdot 2C_1 \log F.$$

so to conclude the Lemma one can pick $C = 2t_1^* C_1$ which depends only on $\kappa, \psi'$. $\square$

### 3.3 Surprises in general case and proof of Theorem 4

Our goal in this section is to prove Theorem 4 in the general case. If there are multiple surprises then at time $T_*$ we have several components linked by edges activated by surprises, and the techniques from Section 3.2 can be adapted to this situation. The ultimate conclusion is the same: we end up with agreement for a positive fraction of issues on the final component, and then the process proceeds to fixation.

Let $T_* = (2N^6/\delta)\log\log F$ which is twice the RHS of (2.16). The following Lemma says that at time $T_*$, the number of agreements on any edge activated by a surprise event is very small compared to $F$.

**Lemma 3.9.** *There is a constant $\alpha > 0$ such that $|\mathcal{J}_{e^*}(T_*)| \le (\log F)^\alpha$ with high probability for any edge $e^* \in \mathcal{E}$ with $A_{e^*}(0) = 0$,*

*证明.* Lemma 3.2 indicates that $A_{e^*}(0) = 0$ on the set $H_*(\kappa, F)$, which occurs with probability $\ge 1-\varepsilon$ when $\kappa \le \varepsilon/(\gamma N^6)$. $|\mathcal{J}_{e^*}(t)|$ can be coupled and upper bounded by Yule binary fission at rate 1 starting with the same initial conditions. Computing the expectation gives $E|\mathcal{J}_{e^*}(t)| \le \exp(t)$, so from (2.16)

$$P\left(|\mathcal{J}_{e^*}(T_*)| \ge (\log F)^\alpha\right) \le (\log F)^{-\alpha} \cdot E|\mathcal{J}_{e^*}(T_*)| \le \exp\left(\left(\frac{2N^6}{\delta} - \alpha\right)\log\log F\right)$$

which tends to 0 for fixed $\alpha > (2N^6/\delta)$. □

We will assume in this section that the network is *star-like*, i.e. each component is linked to the giant component with one single edge, and all surprise edges must connect to $\chi$. In the case where $N$ is large this is a reasonable assumption since the giant component scales as $O(N)$ while the second largest is of $O(\log N)$. The competing growth arguments for agreements on several extended components no longer work, since the growth rates can be different and hardly estimated in accuracy, so we turn to a multi-round process. The sketch is as follows. For $e^*$ being a surprise edge at $T^*$ let $\phi_{e^*}$ be the component $\chi$ is connected to via $e^*$, and define $\mathcal{S}_0 = \{e^* \notin \chi : e^* \sim \chi \text{ and } A_{\chi\cup e^*\cup\phi_{e^*}}(T_*) \ge 1\}$ as the set of edges activated by surprises within rescaled time $[0, T_*]$ and have at least one issue of agreement on the new component at time $T_*$. In each round $k \ge 0$ we identify the surprise edges $\{e^* : e^* \in \mathcal{P}\} \subset \mathcal{S}_k$ where the number of agreements over each $e^* \in \mathcal{P}$ is of positive fractions. $\tau_k^o$ is the time that we have $e^*$ and the component it connects to, $\phi_{e^*}$, join the giant component. The arguments in Section 3.2.1 and 3.2.2 can be adapted to allow

$$A_{\chi_k \cup \bigcup_{\mathcal{P}}(\{e^*\}\cup\phi_{e^*})}(\tau_k^o) \geq \delta_{k+1} F$$

for some constant $\delta_{k+1} > 0$. Let $\chi_{k+1} = \chi_k \cup \bigcup_{\mathcal{P}} (\{e^*\} \cup \phi_{e^*})$ and $\mathcal{S}_{k+1} = \mathcal{S}_k \backslash \bigcup_{\mathcal{P}} \{e_k^*\}$, then we are back in the setup where $A_{\chi_{k+1}} \geq \delta_{k+1} F$ and $\max_{e^* \in \mathcal{S}_{k+1}} A_{\phi_{k+1}} < \kappa F$. The procedure is gone through for $K$ rounds until $\mathcal{S}_K = \varnothing$, so agreements on the ultimate component $\phi_K$ has agreements on $\delta_K F$ issues.

To argue for each round $k$, let $\mathcal{C} = \chi \cup \bigcup_{\mathcal{P}} (\{e^*\} \cup \phi_{e^*})$. At time $T_*$ consensus has been reached on $\chi$ and $(\phi_{e^*} : e^* \in \mathcal{S})$ over all issues not disturbed by surprises. We introduce

$$\begin{aligned}
\tau_\chi(\kappa) &= \inf\{t \geq T_* : A_\chi(t) < 0.75F\} \\
T_{\mathcal{P}}(\kappa) &= \inf\{t \geq T_* : \max_{\mathcal{P}} |\mathcal{J}_{e^*}(t)| \geq \kappa F\} \\
T_\Delta(\kappa) &= \inf\{t \geq T_* : A_{\mathcal{C}}(t) \geq \delta_{e^*} F\}
\end{aligned}$$

which are stopping times with regards to the filtration generated by Axelrod's model at rate $F$.

**Lemma 3.10.** *At $T_{\mathcal{S}}$ we have for some $\delta' < \kappa$ positive such that with high probability as $F \to \infty$*

$$\left| \bigcap_{e^* \in \mathcal{P}} \mathcal{J}_{e^*}^F(T_{\mathcal{P}}) \right| \geq \delta' F.$$

证明. Assume $\kappa_{e^*} = A_{e^*}(T_{\mathcal{P}}(\kappa))$ and we have $\mathcal{J}_{e^*} = \mathcal{J}_{e^*}(T_{\mathcal{P}})$ as random subset of $\{1, \ldots, F\}$ of size $\kappa_{e^*} F$. Therefore

$$\begin{aligned}
E\left(\left|\bigcap_{e^* \in \mathcal{P}} \mathcal{J}_{e^*}\right|\right) &= \left(\prod_{e^*} \kappa_{e^*}\right) F \\
\mathrm{var}\left(\left|\bigcap_{e^* \in \mathcal{P}} \mathcal{J}_{e^*}\right|\right) &= \sum_{f \in [F]} \mathrm{var}(\mathbf{1}_{f \in \cap_{e^*} \mathcal{J}_{e^*}}) + \sum_{f,g \in [F], f \neq g} \mathrm{cov}\left(\mathbf{1}_{f \in \cap_{e^*} \mathcal{J}_{e^*}}, \mathbf{1}_{g \in \cap_{e^*} \mathcal{J}_{e^*}}\right) \leq F/4
\end{aligned}$$

where by simple calculations the covariance terms are $\leq 0$. Let $Y = \left|\bigcap_{e^* \in \mathcal{P}} \mathcal{J}_{e^*}\right|$ then by applying Chebyshev's inequality

$$P\left(Y \leq \frac{1}{2} E(Y)\right) \leq \frac{4\mathrm{var}(Y)}{(E(Y))^2} \leq cF^{-1}$$

and hence the desired result follows by taking $\delta' = \frac{1}{2} \prod_{e^* \in \mathcal{P}} \kappa_{e^*}$. □

**Lemma 3.11.** *With constants $\kappa > 0$ appropriately chosen, w.h.p.* $T_\Delta \leq T_{\mathcal{P}} \leq \tau_\chi$.

证明. Since $\kappa$ is small we can assume WLOG $\kappa|\mathcal{P}| \leq 1/4$. The issues in $\bigcup_{e^*\in\mathcal{P}} \mathcal{J}_{e^*}$ are the only ones whose agreement can possibly break down, and thus $I_\chi(t) \supseteq I_\chi(T^*)\backslash \bigcup_{e^*\in\mathcal{P}} \mathcal{J}_{e^*}(t)$ for $t \geq T_*$ which, along with Lemma 3.9, implies

$$A_\chi(t) \geq F - (\log F)^a - \sum_{e^*\in\mathcal{P}} |\mathcal{J}_{e^*}(t)| \geq F - (\log F)^a - \kappa|\mathcal{P}|F$$

for $t \leq T_{\mathcal{P}}$. By definition of $\tau_\chi$ the second inequality is valid. For the first one, we consider for issue $f \in \cap_{e^*} \mathcal{J}_{e^*}(T_{\mathcal{P}})$ each edge activated by surprises $e^* = \{i, j\}$ where $i \in \chi$ and $j \in \phi_{e^*}$. Either

(1) $i$ copies its opinion to $j$ and $f \in I_\chi(t-)$, and the opinion at $j$ immediately spreads over to all members of $\phi_{e^*}$, one extra agreement on $\chi \cup e^* \cup \phi_{e^*}(t)$ is generated;

(2) $i$ copies its opinion to $j$ and $f \in I_\chi(t-)$, but the opinion at $j$ was fastly taken over by the opinion of $\phi_{e^*}$, no increment is caused;

(3) $j$ copies its opinion to $i$ or $f \notin I_\chi(t-)$, no increment is caused.

Let $|\phi_{e^*}|$ be the number of vertices in the component $\phi_{e^*}$. Note from the arguments in Section 3.2.1 and Section 3.2.2, before $\tau_\chi$ case (i) above has probability $\geq (0.75/2)\eta_{e^*}(\kappa, \phi_{e^*})$. Apply Lemma 3.4 with $n = \delta' F$ ($\delta'$ defined in the proof of Lemma 3.10), $\kappa' = \prod_{\mathcal{P}} ((0.75/2)\eta_{e^*}(\kappa, \phi_{e^*}))$, $p = \kappa' = 2\varepsilon$ we have with high probability the agreement $A_{\chi\cup\mathcal{P}\cup\phi_{\mathcal{P}}}(T_{\mathcal{P}})$ is bounded below by $\kappa' F$ which makes the first inequality valid in our settings. □

Theorem 4 follows from Lemma 3.11 and 3.9.

# 参考文献

[1] Athreya, K. B., and Ney, P. (1972) *Branching processes* (Vol. 196). Springer, New York

[2] Axelrod, R. (1997) The dissemination of culture: a model with local convergence and global polarization. *J. Conflict. Resolut.* 41, 203–226

[3] Castellano, C., Fortunato, S., and Loreto, V. (2009) Statistical physics of social dynamics. *Rev. of Mod. Phys.* 81, 591–646

[4] Castellano, C., Marsili, M., and Vespignani, A. (2000) Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Letters.* 85, 3536–3539

[5] Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000) Mixing beliefs among interacting agents. *Adv. Complex Syst.* 3, 87–98

[6] Durrett, R. (1995) Ten lectures on particle systems. Pages 97–201 in *Springer Lecture Notes in Mathematics*, 1608.

[7] Durrett, R. (1996) *Stochastic Calculus: A Practical Introduction.* CRC Press, Boca Raton, FL

[8] Durrett, R. (2007) *Random Graph Dynamics.* Cambridge University Press

[9] Durrett, R. (2010) *Probability: Theory and Examples.* Fourth Edition. Cambridge University Press

[10] Durrett, R., and Mayberry, J. (2011). Traveling waves of selective sweeps. *Ann. Appl. Probab.* 21, 699–744

[11] Ethier, S.N., and Kurtz, T.G. (1986) *Markov Processes: Characterization and Convergence.* John Wiley and Sons, New York

[12] Harris, T. E. (1972) Nearest-neighbor Markov interaction processes on multidimensional lattices. *Adv. in Math.* 9 (1), 66–89.

[13] Kesten, H. (1982) *Percolation Theory for Mathematicians.* Birkahuser, Boston

[14] Kurtz, T. G. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* 49–58.

[15] Lanchier, N. (2009) Opinion dynamics with confidence threshold: an alternative to the Axelrod model. *Alea.* 6, 1–42

[16] Lanchier, N. (2012) The Axelrod model for the dissemination of culture revisited. *Ann. Appl. Probab.* 22, 860–880

[17] Lanchier, N., and Scarlatos, S. (2012) Fixation in the one-dimensional Axelrod model. *Ann. Appl. Probab.* to appear

[18] Lanchier, N., and Schweinsberg, J. (2012) Consensus in the two-state Axelrod model. *Stoch. Proc. Appl.* 122, 3701–3717.

[19] Liggett, T. M. (1985) *Interacting Particle Systems.* Springer.

[20] Vázquez, F., Krapivsky, P.L., and Redner, S. (2003) Constrained opinion dynamics: freezing and slow evolution. *J. Phys. A.* 36, L61–L68

[21] Vázquez, F., and Redner, S. (2003) Ultimate fate of constrained voters. *J. Phys. A.* 37, 8479–8494

[22] Vilone, D., Vespignani, A. and Castellano, C. (2002) Ordering phase transition in the one-dimensional Axelrod model. *Eur. Phys. J. B.* 30, 399–406.