Evolving voter models on thick graphs

Anirban Basak, Rick Durrett, Yuan Zhang

The Model

Individuals have one of two opinions (called 0 and 1). In the discrete time formulation, oriented edges (x, y) are picked at random. If x and y have the same opinion no change occurs.

If x and y have different opinions then: with probability $1 - \alpha$, the individual at x imitates the opinion of the one at y; otherwise, i.e., with probability α, the link between them is broken and x makes a new connection to an individual z chosen at random (i) from those with the same opinion ("rewire-to-same"), or (ii) from the network as a whole ("rewire-to-random").

The evolution of the system stops at time τ when there are no "discordant" edges that connect individuals with different opinions.

Holme and Newman (2006)

were the first to consider a model of this type. They chose option (i), rewire-to-same, and initialized the graph with large number K of opinions so that N/K remained bounded as the number of vertices $N \to \infty$. They argued that there was a critical value α_c so that

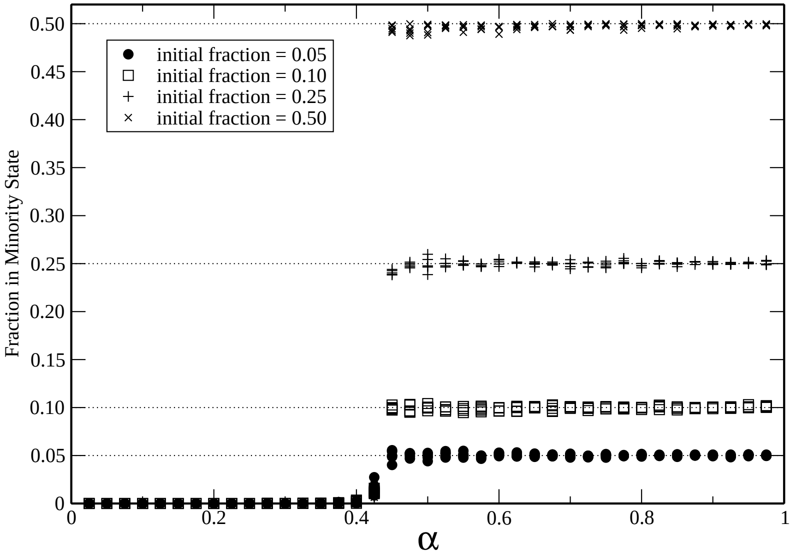
- for $\alpha > \alpha_c$, the graph rapidly disconnects, in time $O(N \log N)$, into a large number of small components,
- if $\alpha < \alpha_c$, the system runs for time $O(N^2)$ and at the end there is a "giant community of like-minded individuals" of size $O(N)$.

PNAS 109 (2012) 3682–3687

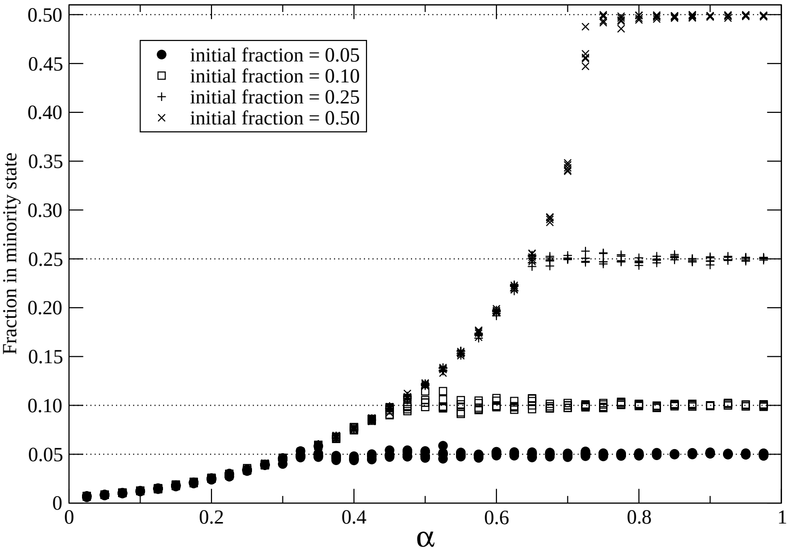
There are two opinions. We start with product measure with density p. Let π be the minority fraction at time τ. Through a combination of simulation and heuristics, Durrett, Gleeson, Lloyd, Mucha, Shi, Sivakoff, Soclolar, and Varghese argued that

- In case (i), rewire-to-same, there is a critical value α_c which does not depend on p, with $\pi \approx p$ for $\alpha > \alpha_c$ and $\pi \approx 0$ for $\alpha < \alpha_c$.
- In case (ii), rewire-to-random, the transition point $\alpha_c(p)$ which depends on the initial density p. For $\alpha > \alpha_c(p)$, $\pi \approx p$, but for $\alpha < \alpha_c(\rho)$ we have $\pi(\alpha, p) = \pi(\alpha, 1/2)$.

Rewire-to-same



Rewire-to-random



Basu and Sly, Ann. Appl. Probab., to appear

Proved the existence of a phase transition for the dynamics on the dense Erdős-Rényi graph $G(N, 1/2)$ with voter events occurring with probability $1-\alpha = \nu/N$. Let τ be the first time there are no discordant edges. Let $N_*(t)$ be the number of vertices holding the minority opinion at time t and for $0 < \epsilon < 1/2$

Theorem 1. Consider the efficient version of the model in which only discordant edges are chosen at random for updating, started from product measure with density. There is a ν_0 so that for all $\nu < \nu_0$ and any $\eta > 0$

$$P\left(\tau < 10N^2,\ N_*(\tau) \geq \frac{1}{2} - \eta\right) \to 1 \quad \text{as } N \to \infty.$$

Prolonged persistence

The next theorem is the main result of their paper and has a very long and difficult proof. Let $\tau_*(\epsilon) = \min\{t : N_*(t) \le \epsilon n\}$.

Theorem 2. Let $\epsilon' \in (0, 1/2)$ be given. There is a $\nu_*(\epsilon')$ so that for $\nu > \nu_*(\epsilon')$ we have $\tau_*(\epsilon') \le \tau$ with high probability and

$$\lim_{c \downarrow 0} \liminf_{N \to \infty} P(\tau > cN^3) = 1.$$

In continuous time where each oriented edge updates at times of a rate 1 Poisson process, N^3 here and Holme and Newman's N^2 are both N.

Rewire to random

Theorem 3. Let $\nu > 0$ be fixed. For the rewire-to-random model, there is an $\epsilon_*(\nu)$ so that $\tau < \tau_*(\epsilon_*)$ with high probability.

This implies that at time τ the minority fraction is $\geq \epsilon_*$. is consistent with the simulation for sparse graphs shown earlier, but is believed to be false for rewire to same.

Thick graphs

The process on $G(N, 1/2)$ is ugly because it quickly develops a large number of parallel edges. We consider Erdös-Renyi graphs with average degree $L = N^a$ with $0 < a < 1$ and forbid the creation of parallel edges. As in Basu and Sly voting rate $1 - \alpha = \nu/L$. We define "finite dimensional distributions"

$$N_i = \sum_x 1_{\{\xi(x)=i\}},$$
$$N_{ij} = \sum_{x,y\sim x} 1_{\{\xi(x)=i,\xi(y)=j\}},$$
$$N_{ijk} = \sum_{x,y\sim x,z\sim y,z\neq x} 1_{\{\xi(x)=i,\xi(y)=j,\xi(z)=k\}},$$

N_{10} versus N_1 when $N = 2500$, $L = 50$, $\nu = 2.5$.

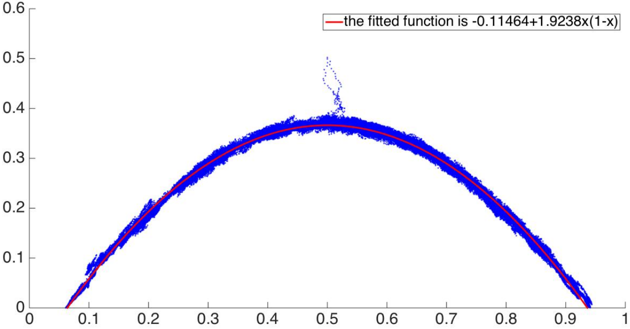


Figure: The arch has endpoints $(\alpha(\nu), 1 - \alpha(\nu))$. Here $(0.0737, 0.09263)$

N_{100} versus N_1 when $N = 2500$, $L = 50$, $\nu = 2.5$.

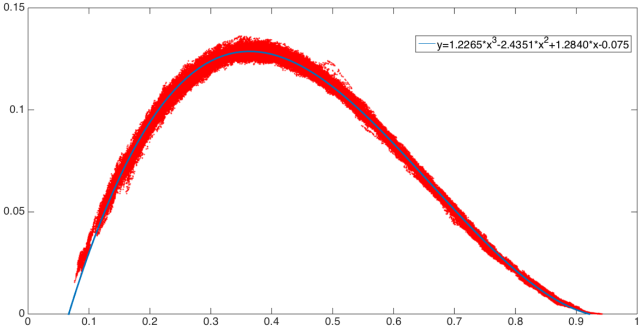


Figure: Same end points as in previous fit. Now the function is cubic.

What the simulations tell us

The fraction of 1's, $\theta_t = N_1(t)/N$, determines the values of all the other statistics, i.e., there is a one-parmeter family of quasi-stationary distributions. Cox and Greven proved this for the voter model on the torus in $d \geq 3$ and that θ_t follows the Wright-Fisher diffusion

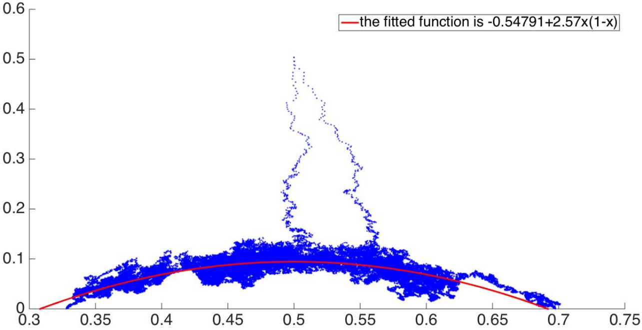
$$d\theta_t = \sqrt{\beta_d \cdot 2\theta_t(1-\theta_t)}$$

In the first part of the simulation the density goes straight down (i.e., θ_t does not change) so

$$\nu_c(p) = \inf\{\nu : p \in (a(\nu), 1-a(\nu))\}$$

If the curve hits the arch it diffuses along it until one endpoint is reached, and hence the ending density does not depend on the starting density.

N_{10} versus N_1 when $N = 2500$, $L = 50$, $\nu = 1$.



Evolution equations

$$\frac{dN_{10}}{dt} = -N_{10} + \frac{\nu}{L}[N_{100} - N_{010} + N_{110} - N_{101}] \qquad (1)$$
$$\frac{1}{2}\frac{dN_{11}}{dt} = pN_{10} + \frac{\nu}{L}[N_{101} - N_{011}] \qquad (2)$$
$$\frac{1}{2}\frac{dN_{00}}{dt} = (1-p)N_{10} + \frac{\nu}{L}[N_{010} - N_{100}] \qquad (3)$$

Note that $N_{ij} = O(NL)$ while $N_{ijk} = O(NL^2)$ so the terms on the right-hand side of (1)-(3) are of the same order of magnitude. In writing these equations we have omitted terms such as $(\nu/L)N_{ij}$ since they are $O(N)$. Since $\sum_{ij} N_{ij} = NL$ the three equations add up to 0.

Pair approximation

let J_i and K_i be the average number of 1 neighbors and 0 neighbors of a vertex in state i. The pair approximation is

$$N_{101} = \sum_{x:\xi(x)=1} \sum_{y:\xi(y)=0} j_0(y) \approx N_{10} J_0.$$

Applying similar reasoning for the other N_{ijk}'s we have

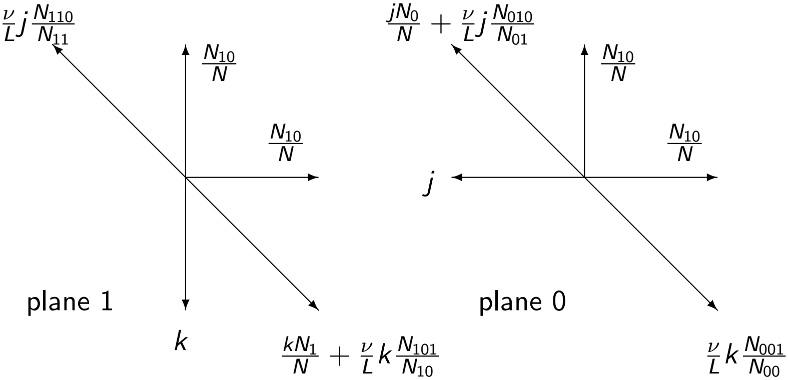
$$\begin{aligned}
\frac{1}{2}\frac{dN_{11}}{dt} &\approx pN_{10} + \frac{\nu}{L}[N_{10}J_0 - N_{01}J_1], \\
\frac{1}{2}\frac{dN_{00}}{dt} &\approx (1-p)N_{10} + \frac{\nu}{L}[N_{01}K_1 - N_{10}K_0].
\end{aligned}$$

This predicts that $J_0^* = Lp(1 - [p^2 + (1-p^2]/\nu)$ and hence $\nu_c(p) = p^2 + (1-p)^2$, but simulations show $\nu_c(1/2) > 0.8$.

Approximate Master Equation

We visualize our system as N particles, one for each vertex, moving in two planes. A point at (i, j, k) means that the state of the vertex is i, there are j neighbors in state 1, and k in state 0.

Voting events at the focal vertex x cause jumping from $(1, j, k) \to (0, j, k)$ at rate $\nu k/L$ and from $(0, j, k) \to (1, j, k)$ at rate $\nu j/L$.



Here the rates on horizontal and vertical edges which come from rewiring are exact. On the diagonal arrows kN_1/N and jN_0/N are exact but the others come from e.g., using N_{ijk}/N_{ij} to compute the expected number of neighbors of z in state k when x is in state i and y is in state j.

To study this system, we will introduce $q = 1 - p$,

$$\alpha = \frac{N_{101}}{N_{10}}, \quad \beta = \frac{N_{110}}{N_{11}}, \quad \eta = \frac{N_{10}}{N} \quad \delta = \frac{N_{010}}{N_{01}}, \quad \epsilon = \frac{N_{001}}{N_{00}}.$$

and analyze the system in general. The infinitesmial mean saisfies

$$\begin{aligned}
\text{plane 1} \quad \frac{dj_1}{dt} &= \eta \qquad\quad + pk_1 + \frac{\nu}{L}\alpha k_1 - \frac{\nu}{L}\beta j_1 \\
\frac{dk_1}{dt} &= \eta - k_1 - pk_1 - \frac{\nu}{L}\alpha k_1 + \frac{\nu}{L}\beta j_1 \\
\text{plane 0} \quad \frac{dj_0}{dt} &= \eta - j_0 - qj_0 - \frac{\nu}{L}\delta j_0 + \frac{\nu}{L}\epsilon k_0, \\
\frac{dk_0}{dt} &= \eta \qquad\quad + qj_0 + \frac{\nu}{L}\delta j_0 - \frac{\nu}{L}\epsilon k_0.
\end{aligned}$$

If we let $N \to \infty$ scale space by L, and suppose

$$\frac{\alpha}{L} \to \bar{\alpha}, \quad \frac{\beta}{L} \to \bar{\beta}, \quad \frac{\eta}{L} \to \bar{\eta}, \quad \frac{\delta}{L} \to \bar{\delta}, \quad \frac{\epsilon}{L} \to \bar{\epsilon}.$$

then in the limit we get a system in which single particles that moves according to the following differential equations

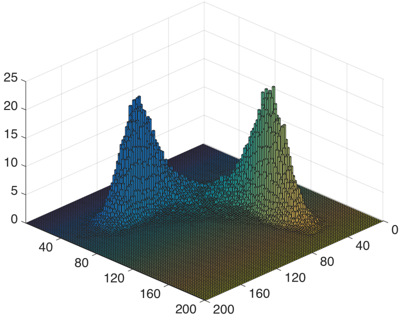
$$\begin{aligned}
\text{plane 1} \quad \frac{dx_1}{dt} &= \bar{\eta} \qquad\quad + py_1 + \nu\bar{\alpha}y_1 - \nu\bar{\beta}x_1, \\
\frac{dy_1}{dt} &= \bar{\eta} - y_1 - py_1 - \nu\bar{\alpha}y_1 + \nu\bar{\beta}x_1. \\
\text{plane 0} \quad \frac{dx_0}{dt} &= \bar{\eta} - x_0 - qx_0 - \nu\bar{\delta}x_0 + \nu\bar{\epsilon}y_0, \\
\frac{dy_0}{dt} &= \bar{\eta} \qquad\quad + qx_0 + \beta\bar{\delta}x_0 - \beta\bar{\epsilon}y_0.
\end{aligned}$$

and jump between the planes.

Using techniques of Lawley, Mattingly, and Reed we can show.

Theorem. Fix $\nu > 0$, $p \in (0,1)$ and let $q = 1 - p$. For any $\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\epsilon}, \bar{\eta} > 0$ The two plane system has a unique stationary distribution that is the limit starting from any initial configuration.

Proof. Start at time $-n$ and run to time 0. As $n \to \infty$ the state at time 0 converges to a limit almost surely.



25
20
15
10
5
0
40
80
120
160
200
0
40
80
120
160
200

Following Silk et al, one can write partial differential equations for the moment generating functions for the limit measures on the two planes

$$0 = \nu\bar{\beta}(b-a)U_a + \nu V_a + ([p+\bar{\alpha}\nu](a-b) - b - \nu)U_b + \bar{\eta}aU + \bar{\eta}bU,$$
$$0 = \nu\bar{\epsilon}(a-b)V_b + \nu U_b + ([p+\bar{\delta}\nu](b-a) - a - \nu)V_a + \bar{\eta}aV + \bar{\eta}bV.$$

Since derivatives are moments, we can extract some information from this

ν	U_b sim	U_{ab} sim	calc	U_{bb} sim	calc	U_{aa} sim	calc
2	0.1666	0.1025	0.1041	0.0604	0.0625	0.2336	0.2208
1.6	0.1371	0.0907	0.0900	0.0466	0.0471	0.2859	0.2574
1.44	0.1216	0.0827	0.0819	0.0394	0.0397	0.3115	0.2810
1.32	0.1094	0.0757	0.0754	0.0343	0.0340	0.3310	0.3047
1.2	0.0896	0.0641	0.0635	0.0264	0.0261	0.3735	0.3351
1	0.0454	0.0339	0.0341	0.0132	0.0113	0.4690	0.4129

Table: Simulation of evolving voter model compared with computations for the approximate mater equation. The caclulated values of U_{ab} and U_{bb} differ by 1% from simulation but U_{aa} is off by 10%.