Voter Models on Graphs

by

Ran Huo

Department of Mathematics
Duke University

Date: ____________________________
Approved:

____
Richard Durrett, Supervisor

____
Jianfeng Lu

____
James Nolen

____
Matthew Junge

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2019

Abstract

Voter Models on Graphs

by

Ran Huo

Department of Mathematics
Duke University

Date: ____________________
Approved:

____
Richard Durrett, Supervisor

____
Jianfeng Lu

____
James Nolen

____
Matthew Junge

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2019

Abstract

The voter model which describes the flow of information through interactions between neighbors has been widely studied in the field of probability. In this paper we study two variations of the voter model, one is called the Latent Voter Model and the other is called the Zealot Voter Model. Both models are implemented in a space that is a random graph. In the latent voter model, which models the spread of a technology through a social network, individuals who have just changed their choice have a latent period, which is exponential with rate λ, during which they will not buy a new device. We study site and edge versions of this model on random graphs generated by a configuration model in which the degrees $d(x)$ have $3 \leqslant d(x) \leqslant M$. We show that if the number of vertices $n \to \infty$ and $\log n \ll \lambda_n \ll n$ then the fraction of 1's at time $\lambda_n t$ converges to the solution of $du/dt = c_p u(1-u)(1-2u)$. Using this we show the latent voter model has a quasi-stationary state in which each opinion has probability $\approx 1/2$ and persists in this state for a time that is $\geqslant n^m$ for any $m < \infty$. Thus, even a very small latent period drastically changes the behavior of the voter model, which has a one parameter family of stationary distributions and reaches fixation in time $O(n)$. Inspired by the spread of discontent as in the 2016 presidential election, we consider a voter model in which 0's are ordinary voters and 1's are zealots. Thinking of a social network, but desiring the simplicity of an infinite object that can have a nontrivial stationary distribution, space is represented by a tree. The dynamics are a variant of the biased voter: if x has degree $d(x)$ then at

rate $d(x)p_k$ the individual at x consults $k \geqslant 1$ neighbors. If at least one neighbor is 1, they adopt state 1, otherwise they become 0. In addition at rate p_0 individuals with opinion 1 change to 0. As in the contact process on trees, we are interested in determining when the zealots survive and when they will survive locally, i.e., the root of the tree is in state 1 infinitely often.

This thesis is dedicated to my family, in particular my most loving grandparents.

Contents

List of Figures

Acknowledgements

I would like to thank my family for their unconditional love and confidence in me, especially my dad who is my best pal that supports my every crazy idea. Without them I would never be able to finish this adventure. I would also like to express my deep gratitude to my advisor Prof. Rick Durrett for his academic guidance as well as his full support my in academic endeavors. I would also like to thank Prof. Jianfeng Lu, Prof. Jim Nolen and Prof. Sarah Schott for their advice and letters during my application. Besides, my thanks extends to all the professors I took class with. Moreover I would like to thank my undergraduate research mentor Prof. Benedek Valkó ,not only for his support on my job, but also it was his project that I was motivated to do a PhD in probability.

I am very grateful that I have encountered so many wonderful people in this department, Orsola Capovilla-Searle, Erika Ordog, Gabe Earle, Yu Cao, Ruby Kim, Sarah Ritchey, Dena Zhu, Erin Beckman and many many others . With you guys, my research is a lot easier and more interesting! In particular, I would like to express my infinite gratitude to one of my most diligent office mate, most reliable roommate Weifan Liu who is such a great listener that I can shamelessly complain to, explode my tears and share my epic happiness with. Last and not least, to the most two important friends of my life, as I promised to mention their names–Yuyang Hu and Ruiheng Wen, my thanks to you guys is more than a million words.

1

Introduction

Holley and Liggett introduced introduced the voter model. Here we consider two types of voter models: the Latent voter model and the Zealot voter model.

1.1 the Latent Voter Model

In the first part we will study the latent voter model introduced in 2009 by Lambiotte, Saramaki, and Blondel [16]. In this model each individual owns one of two types of technology, say an iPad or a Microsoft Surface tablet. In the voter model on the d-dimensional lattice, individuals at times of a rate one Poisson process pick a neighbor at random and imitate their opinion. However, in the current interpretation of that model, it is unlikely that someone who has recently bought a new tablet computer will replace it, so we introduce latent states 1^* and 2^* in which individuals will not change their opinion. If an individual is in state 1 or 2 we call them active. Letting f_i be the fraction of neighbors in state i or i^*, the dynamics can be formulated as follows

$$
\begin{array}{ll}
1 \to 2^* \text{ at rate } f_2 & 1^* \to 1 \text{ at rate } \lambda \\
2 \to 1^* \text{ at rate } f_1 & 2^* \to 2 \text{ at rate } \lambda
\end{array}
$$

We will study the system with large λ. That is, although the individual has a latent period after changing device, pretty quickly this person will resume to be active. In [16] the authors showed that if individuals in the population interact equally with all the others then the system converges to a limit in which both technologies have frequency close to 1/2. Surprisingly, the latent voter model have similar behavior on finite graphs generated by the configuration model which we will discuss later, see Theorem 1.1.2 for the result. The difference from [16] is that in the site version of the latent voter model, the population does not interact equally. To study this model, we will construct the system using a graphical representation. Suppose first that the system takes place on $\mathbb{Z}^d$ and that $d \geqslant 3$. For each $x \in \mathbb{Z}^d$ and nearest neighbor y, we have independent Poisson processes $T_n^{x,y}$, $n \geqslant 1$ that come at rate 1 (in the edge version, this equals the degree of x). At each time $t = T_n^{x,y}$ we draw an arrow from $(y,t) \to (x,t)$ to indicate that if the individual at x is active at time t then they will imitate the opinion at y.

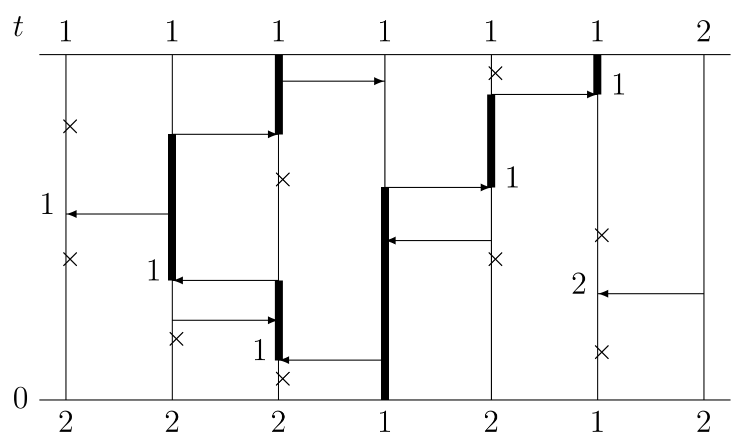
To implement the other part of the mechanism, we introduce for each site x, a Poisson process T_n^x, $n \geqslant 1$ of "wake-up dots" that return the voter to the active state.

- If there is only one voter arrow between two wake up dots, the result is an ordinary voter event.
- If between two wake up dots there are voter arrows to x from two different neighbors, an event of probability $O(\lambda^{-2})$, then x will change its opinion if and only at least one of the two neighbors has a different opinion. To check this, we note that if the first arrow causes a change then the second one is ignored, while if the first arrow comes from a site with the same opinion as the one

at x then there will be a change if and only if the second site has an opinion different from the one at x.

- If t is fixed then at a given site there are $O(\lambda)$ wake-up dots by time t. Thus if we want to see the influence of intervals with two voter arrows then we want to run time at rate λ. The probability of k voter arrows between two wake-up dots is $(1+\lambda)^{-k}$, so in the limit the probability of three of more voter events between two wake-up dots goes to 0 as $\lambda \to \infty$.

The following is an example:



In the figure, the bottom shows the initial states assigned to each individual. t is the current moment and the figure in between shows the timeline of each individual. We use thick line segments to differentiate the time period where the person is in state 1. To explain how it helps us identify the states at any time t, take the third left person as an example. It first looks at its right neighbor and changed to 1 then stays inactive. Since the next arrow from its left comes before a wake-up dot, it is ignored. However, going upward, the third arrows has an effect since this person becomes active after the second wake-up dot. The main idea is that we would be able to find the states of the population at anytime by working upward the graphical representation.

1.1.1 ODE limit of the density

Here we consider the Latent voter model on random graphs with bounded degrees (see Section 2.0.1 in Chapter 2 for details). Let $\xi_t(x)$ denote the state of vertex x at time t. The last detail is to define the density $U^n(t)$. To do this we note that a random walk that jumps to a neighbor chosen at random has stationary distribution $\pi(x) = d(x)/D$, where $D = \sum_y d(y)$, while a random walk that jumps to each neighbor at rate 1 has stationary distribution $\pi(x) = 1/n$. To treat the two cases in one definition we let

$$U^n(t) = \sum_x \pi(x) 1_{\{\xi_{\lambda t}(x)=1\}} \tag{1.1}$$

Our first result is the ODE limit.

Theorem 1.1.1. *Suppose that* $\log n \ll \lambda_n \ll n$. *If* $U^n(0) \to u_0$ *then* $U^n(t)$ *converges in probability and uniformly on compact sets to* $u(t)$*, the solution of*

$$\frac{du}{dt} = c_p u(1-u)(1-2u) \qquad u(0) = u_0. \tag{1.2}$$

where the value of c_p *depends on the degree distribution and the version of the voter model.*

For people who are interested in the voter model perturbation, it turns out that the latent voter model is a voter model perturbation in the sense of Cox, Durrett, and Perkins [7]. In the site version if we let $\lambda = \varepsilon^{-2}$ and let $n_k(x)$ be the number of neighbors in state k or k^* then the rate of flips from 1 to 2 in the latent voter model when the configuration is ξ is:

$$\varepsilon^{-2} c^v_{1,2}(x,\xi) + h_{1,2}(x,\xi) \quad \text{where} \quad c^v_{1,2}(x,\xi) = 1_{\{\xi(x)=1\}} \frac{n_2(x)}{d(x)}$$

If we let $y_1, \dots y_{2d}$ be an enumeration of the nearest neighbors of x, the perturbation is

$$h_{1,2}(x,\xi) = 1_{\{\xi_t(x)=1\}} \frac{2}{d(x)(d(x)-1)} \sum_{1\leqslant k<\ell\leqslant 2d} 1_{\{\xi(y_k) \text{ or } \xi(y_\ell)\in\{2,2*\}\}}$$

Similar formulas hold when the roles of 1 and 2 are interchanged. $h_{1*,j} = h_{2*,j} \equiv 0$. Authors in [7] have extensively studied the voter model perturbation on $\mathbb{Z}^d$ where in above $d(x)$ is replaces by d. Their Theorem 1.2 suggests that under mild assumptions on the perturbation, if we scale space by ε, the density of 1's in the rescaled particle system converges to the solution of the limiting PDE:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \phi(u) \quad \text{with} \quad \phi(u) = \langle h_{2,1}(0,\xi) - h_{1,2}(0,\xi)\rangle_u \tag{1.3}$$

and $\langle \cdot \rangle_u$ denotes the expected value with respect to the voter model with density u.

Intuitively, (2.1) holds because of a separation of time scales. The rapid voting means that the configuration near x looks like the voter model equilibrium with density $u(t,x)$. Later in the paper will show, see (2.11), that in the case of the latent voter model

$$\phi(u) = c_d u(1-u)(1-2u).$$

Similar argument applies to the edge version.

If we consider the latent voter model on a torus with n sites and let $\lambda_n \to \infty$ then the system can be analyzed using ideas from a recent paper of Cox and Durrett [5]. In their case the density of 1's at time t by

$$U^n(t) = \frac{1}{n}\sum_x 1_{\{\xi_{\lambda t}(x)=1\}} \tag{1.4}$$

They showed that

Theorem 1.1.2. *Suppose $n^{2/d} \ll \lambda_n \ll n$. If $U^n(0) \to u_0$ then $U^n(t)$ converges uniformly on compact sets to $u(t)$ the solution of*

$$\frac{du}{dt} = c_d u(1-u)(1-2u) \qquad u(0) = u_0$$

1.1.2 Duality

To explain why Theorems 1.1.1 and 1.1.2 are true, we will introduce a dual process that is the key to the analysis. The dual process was first introduced more than 20 years ago by Durrett and Neuhauser [9], and is the key to work of Cox, Durrett, and Perkins [7]. To do this, we construct the process using a graphical representation that generalizes the one introduced for $\mathbb{Z}^d$. For each $x \in \mathbb{Z}^d$ and neighbor y, we have independent Poisson processes $T_n^{x,y}$, $n \geqslant 1$. At each time $t = T_n^{x,y}$ we draw an arrow from $(y,t) \to (x,t)$ to indicate that if the individual at x is active at time t then they will imitate the opinion at y. In the edge case all these processes have rate 1. In the site case $T_n^{x,y}$, $n \geqslant 0$ has rate $1/d(x)$. To implement the other part of the mechanism, we have for each site x, a rate λ Poisson process T_n^x, $n \geqslant 1$ of "wake-up dots" that return the voter to the active state.

To compute the state of x at time t we start with a particle at x at time t. To be precise $\zeta_0^{x,t} = \{x\}$. As we work backwards in time the particle does not move until the first time s there is an arrow $(y, t-s) \to (x, t-s)$.

- If this is the only voter arrow between the two adjacent wake-up dots then the particle jumps to y.
- If in the interval between the two adjacent wake-up dots there are arrows from k distinct y_i then the state changes to $\{x, y_1, \ldots y_k\}$ since we need to know the current state of all these points to know what change should occur in the process. In the limit as $\lambda \to \infty$ we will only see branchings that add two y_i. We include the case $k > 2$ to have the dual process well-defined.

- We do not need to know the order of the arrows because x will change if at least one of the y_i has a different opinion. When λ is small some of the y_i might change their state during the interval between the two wake-up dots but this possibility has probability zero in the limit.

- To complete the definition of the dual, we declare that if a branching event adds a point already in the set, or if a particle jumps onto an occupied site then the two coalesce to one.

The dual process can be used to compute the state of x at time t. The first step is to work backwards in time to find $\zeta_t^{x,t}$ the set of sites at time 0 that can influence the state of x at time t. We note the states of the sites at time 0 and then work up the graphical representation to determine what changes should occur at the branching points in the dual.

To prove Theorem 1.1.1, following the approach of Cox and Durrett [5], we will show that after a branching event any coalescence between the particle that branched and the two newly created particles will happen quickly, in time $O(1)$ or these particles will need time $O(n)$ to coalesce. Our assumption on the graphs that all vertices have degree $\geqslant 3$ implies that the mixing time for random walks on these graphs is $O(\log n)$. Thus when $\lambda_n \gg \log n$, we have the situation that after a branching event there may be some coalescence in the dual at times $O(1)$ but then the existing particles will come to equilibrium on the graph before the next branching occurs in the dual, so we can forget about the relative location of the particles and we end up with an ODE limit. $\lambda_n \ll n$ is needed for the perturbation to have a nontrivial effect. Otherwise a collection of k random walks might all coalesce before the first branching arrow.

1.1.3 Long time survival

The next result proves the long time survival of the model. The latent voter model has two absorbing states $\equiv 1$ and $\equiv 2$. On a finite graph the latent voter model is a finite state Markov chain, so we know it will eventually reach one of them. However by analogy with the contact process on the torus [20] and on power-law random graphs, [21], this result should hold for times up to $\exp(\gamma n)$ for some $\gamma > 0$. Unfortunately we are only able to prove survival for any power of n. This failure is to due to our estimate of $P(\Omega_1^c)$ which appears on the right-hand side of Theorem 2.0.4.

Theorem 1.1.3. *Suppose that* $\log n \ll \lambda_n \ll n$. *Let* $\varepsilon > 0$ *and* $k < \infty$. *If* $U^n(0) \to u_0 \in (0,1)$ *there is a* T_0 *that depends on the initial density so that for any* $m < \infty$ *if* n *is large then with high probability*

$$|U^n(t) - 1/2| \leqslant 5\varepsilon \qquad \textit{for all } t \in [T_0, n^m].$$

To prove Theorem 2.0.3, we use Theorem 4.2 of Darling and Norris [8], which is Theorem 2.0.4 in the following. To state their result we need to introduce some notation. To make it easier compare with their paper we use their notation even though in some cases it conflicts with ours. Let ξ_t be a continuous time Markov chain with countable state space S and jump rates $q(\xi, \xi')$. In our case ξ_t will be the state of the voter model on the random graph. For their coordinate function $x : S \to \mathbb{R}^d$ we will take $d = 1$ and

$$x(\xi) = \sum_{x \in G_n} \pi(x) 1_{\{\xi(x)=1\}}.$$

We are interested in proving an ODE limit for $X_t = x(\xi_{\lambda_n t})$. To compare with the paper note that our ξ_t is their X_t and our X_t is their $\mathbf{X}_t$.

If we set $U = [0,1]$ in [8] then we always have $x(\xi_t) \in U$ so their condition (2) is not needed. For each $\xi \in S$ we define the drift vector

$$\beta(\xi) = \sum_{\xi' \neq \xi} (x(\xi') - x(\xi)) q(\xi, \xi')$$

We let b be the drift of the proposed deterministic limit x_t:

$$x_t = x_0 + \int_0^t b(x_s)\, ds.$$

In our case $b(y) = cy(1-y)(1-2y)$. To measure the size of the jumps we let $\sigma_\theta(y) = e^{\theta|y|} - 1 - \theta|y|$ and let

$$\phi(\xi, \theta) = \sum_{\xi' \neq \xi} \sigma_\theta(x(\xi') - x(\xi)) q(\xi, \xi').$$

Consider the events $\Omega_0 = \{|X_0 - x_0| \leqslant \eta\}$,

$$\Omega_1 = \left\{ \int_0^t |\beta(\xi_{\lambda s}) - b(X_s)|\, ds \leqslant \eta \right\},$$

$$\text{and } \Omega_2 = \left\{ \int_0^t \phi(\xi_s, \theta)\, ds \leqslant \theta^2 At/2 \right\}.$$

The parameters in these events are coupled by the following relationships. If we let K be the Lipschitz constant of the drift b then $\eta = \varepsilon e^{-Kt_0/3}$ and $\theta = \eta/(At)$ where $A > 0$. We have changed their δ to η because we use δ in a number of our arguments.

Theorem 1.1.4. *Under the conditions above, for each fixed t*

$$P\left(\sup_{s \leqslant t_0} |X_s - x_s| > \varepsilon \right) \leqslant 2e^{-\eta^2/(2At_0)} + P(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c)$$

To bound the probabilities on the right-hand side, we note

- We have jumps that change the density by $1/n$ at times of a Poisson processes at total rate $\leqslant M\lambda n$. If $\theta|y|$ is small $\sigma_\theta(\pm 1/n) \sim \theta^2/2n^2$. So if n is large enough so that $M\lambda/n \leqslant n^{-1/2}t_0/2$ then a standard large deviations estimate for the $P(\Omega_2^c) \leqslant \exp(-c\lambda n)$.

- Our assumption that $U^n(0) \to u_0$ implies that $\Omega_0^c = \varnothing$ for large n.

- The hard work comes in estimating $P(\Omega_1^c)$, i.e., estimating the difference in the drift in the particle system from what we compute on the basis of the current density. We do this in Section 2.2.3 by computing the expected value of the mth moment of the difference $|\beta(\xi_s) - b(X_s)|$ so we end up with estimates that for a fixed time are $\leqslant n^{-m}$.

Once these three steps are done, the remainder of the proof of Theorem 2.0.3 given in Section 2.2.5 is routine. By subdividing the interval into small pieces we can use the single time estimates to control the supremum and hence the integral but only over a bounded time interval. However this is enough since it allows us to show that when the density wanders more than 4ε away from 1/2, we can return it to within 2ε with probability n^{-m}, and in addition never have the difference exceed 5ε.

1.2 The Zealot Voter Model

In the standard (linear) voter model, which was introduced by Holley and Liggett [13], a site flips at a rate equal to the fraction of neighbors that have the other opinion. Cox and Durrett [4] began the study of voter models with non-linear flip rates. One of the most successful ideas from that paper is the threshold-θ voter model in which sites flip at rate 1 if at least θ neighbors have the opposite opinion. Liggett [18] obtained results for coexistence of opinions when $\theta = 1$, while Chatterjee and Durrett [2] showed that the model with $\theta \geqslant 2$ had a discontinuous phase transition on the random r-regular graph when $r \geqslant 3$. Lambiotte and Redner [15] studied the "vacillating voter model" in which a voter looks at the opinions of two randomly chosen neighbors and flips if at least one disagrees. At about the same time, Sturm and Swart considered "rebellious voter models" in one dimension. In the one-sided case $\xi_t(i)$ canges its opinion at rate α if $\xi_t(i+1) \neq \xi_t(i))$ and at an additional rate $1-\alpha$ if $\xi_t(i+1) \neq \xi_t(i+2))$. They also considered a spatially symmetric version. In all these variants of the voter model, the process is symmetric under interchange of 0's and 1's. Our zealot voter model does not have this symmetry.

In our process, space is represented by a tree $\mathcal{T}$ in which the degree of each vertex x satisfies $3 \leqslant d_{min} \leqslant d(x) \leqslant M$. This guarantees that our trees are infinite. Voters can be in state 0 (ordinary voter) or 1 (zealot). Given a probability distribution p_k on $\{0, 1, 2, \dots d_{min}\}$, if $k \geqslant 1$ then at rate $d(x)p_k$ the voter x picks k neighbors without replacement. As in the vacillating voter model the voter becomes 1 if at least one of the chosen neighbors is a 1, otherwise it becomes 0. In addition at rate p_0, voters change their opinion from 1 to 0.

If $p_0 = 0$ then this model is a variant of the biased voter model. In that system, a 0 at x changes to 1 at rate λ times $n_1(x)$ the number of neighbors of x that are in state 1, and a 1 at x changes to 0 at rate $n_0(x)$ the number of neighbors of x that are in state 0. If the degree is constant then the behavior of the process is easy to understand. If we start from finitely many 1's then the number of 1's at time t, N_t^1 decreases by 1 at a rate equal to D_t the number of $(1, 0)$ edges, and increases by 1 at rate λD_t. Thus N_t^1 is a time change of a simple random walk that increases by 1 with probability $\lambda/(\lambda+1)$ and decreases by 1 with probability $1/(\lambda+1)$. Using this observation it is easy to show that the critical value for the survival of 1's $\lambda_c = 1$. In our setting sites do not have constant degree and we have a different type of bias. This makes things more complicated, and it is hard to get precise results on the location of phase transitions.

Our process is additive in the sense of Harris [12] and hence can be constructed on a graphical representation with independent Poisson processes $T_n^{x,i}$, $n \geqslant 1$, $0 \leqslant i \leqslant d_{min}$.

- The $T_n^{x,0}$ have rate p_0. At these times we write a δ at x that will kill a 1 at the site.
- The $T_n^{x,i}$ have rate $d(x)p_i$. At time $T_N^{x,i}$ we write a δ at x that will kill a 1 at the site. In addition we draw oriented arrows to x from i neighbors $y_1, \dots y_i$ chosen at random without replacement from the set of neighbors. If any of the y_i are in state 1, then x will be in state 1. Otherwise it will be in state 0.

We will often use coordinate notation for the process, i.e., $\xi_t(x)$ gives the state of x at time t. However it is also convenient to use the set-valued approach with ξ_t^A giving the set of sites occupied by zealots at time t when the initial set of zealots is A. Intuitively, the process ξ_t^A can be defined by introducing fluid at the sites in A. The fluid flows up the graphical representation, being blocked by δ's, and flowing across edges in the direction of their orientations. The state at time t, ξ_t^A is the set of points that can be reached by fluid at time t starting from some site in A at time 0.

A nice feature of this construction is that it allows us to define a dual process in which fluid flows down the graphical representation, is blocked by δ's and flows across edges in a direction opposite their orientations. We let $\zeta_s^{B,t}$ be the points reachable at time $t - s$ starting from B at time t. It is immediate from the construction that

$$\{\xi_t^A \cap B \neq \varnothing\} = \{A \cap \zeta_t^{B,t} \neq \varnothing\} \tag{1.5}$$

It should be clear from the construction that the distribution of $\zeta_s^{B,t}$ for $0 \leqslant s \leqslant t$ does not depend on t, so we drop the t and write the duality as

$$P(\xi_t^A \cap B \neq \varnothing) = P(A \cap \zeta_t^B \neq \varnothing) \tag{1.6}$$

The dual ζ_t^B is a coalescing branching random walk (COBRA) with the following rules. A particle at x dies at rate p_0 and at rate $d(x)p_k$ it dies after giving birth to offspring that occupy k of the neighboring sites chosen at random without replacement. For more details see Griffeath [11].

In the case $p_0 = 0$ this pair of dual processes has been studied by Cooper, Radzik, and Rivera [3]. In their situation the zealot voter model is called a biased infection with a persistent source (BIPS). The phrase persistent source refers to the fact that the BIPS model has one individual that stays infected forever. Their main interest is in the cover time for COBRA, i.e., the time for the process to visit all of the sites. By duality this is related to time for the BIPS to reach all 1's.

In this paper, when we say that a process *survives* we mean that with positive probability it avoids becoming $\varnothing$. We say a process *survives locally* if with positive probability the root 0 is occupied infinitely many times.

When $A = B = \{0\}$, (1.6) implies

$$P(0 \in \xi_t^0) = P(0 \in \zeta_t^0) \tag{1.7}$$

so local survival of one process implies local survival of the other. Taking one of the sets $= \mathcal{T}$ and the other $= \{0\}$ we get

$$P(\xi_t^0 \neq \varnothing) = P(0 \in \zeta_t^{\mathcal{T}}) \qquad P(\zeta_t^0 \neq \varnothing) = P(0 \in \xi_t^{\mathcal{T}})$$

so survival of one process implies that the other has a nontrivial stationary distribution obtained by letting $t \to \infty$ in $\zeta_t^{\mathcal{T}}$ or $\xi_t^{\mathcal{T}}$. Our first result is very general.

Theorem 1.2.1. *On any tree with degrees $3 \leqslant d(x) \leqslant M$, the zealot voter model survives if*

$$\sum_{k \geqslant 2} (k-1)p_k - p_0 > 0.$$

The result is proved by comparing the growth of the process at the "frontier" with a branching process. For the definition of frontier, see the text before Lemma 3.1.2. Note that the degree distribution does not appear in the condition.

1.2.1 Results for d−regular Trees

Let $\beta = 1 - (d-1)^{-2}$ be the probability that two independent random walks on the d-regular tree that start at distance two never hit. See Lemma 3.2.1 for a proof of this.

Theorem 1.2.2. *On a d-regular tree the COBRA dies out if*

$$d\beta \sum_{k \geqslant 2} (k-1)p_k - p_0 < 0. \tag{1.8}$$

When this holds the zealot voter model does not have a nontrivial stationary distribution.

To explain the condition, note that in the dual, a particle dies at rate p_0 and gives birth to k particles at rate dp_k. To get an upper bound on the growth of the dual (i) we ignore coalescence between individuals that are not siblings, and (ii) if k particles are born we number them $1, 2, \dots k$ and ignore coalescence between particles $i > 1$ and $j > 1$. This gives an upper bound on the dual COBRA.

Theorem 1.2.3. *If* (1.8) *holds then the zealot voter model dies out on a d-regular tree.*

To study the local survival of our voter model, we use (1.7) to change the problem to studying the local survival of the COBRA. Let $\mu = \sum_k kp_k$ be the mean number of offspring in the dual process

Theorem 1.2.4. *Given a d-regular tree T, the zealot voter model dies out locally if*

$$\mu < \frac{d(1-p_0)+p_0}{2\sqrt{d-1}}.$$

If $p_0 = 0$ *this is* $\mu < d/(2\sqrt{d-1})$.

This result is proved by comparing COBRA with a branching random walk by ignoring coalescence. The second bound is sharp for the branching random walk with no death. That is, the corresponding branching random walk visits the root with positive probability if $\mu > d/(2\sqrt{d-1})$ and that the root is visited finitely many times if $\mu < d/(2\sqrt{d-1})$. This result can be found in Pemantle and Stacey [22]. There they studied the branching random walk on trees where each particle gives birth at a rate λ independently onto each neighbor, and dies at rate 1. Since our branching process has simultaneous births and deaths we modify their proof to cover our situation and give the proof in Lemma 3.3.1.

To give sufficient conditions for local survival, we follow a tagged particle in the COBRA. If there is a particle produced on the site closer to the root, we follow this particle; otherwise we follow a new particle chosen uniformly at random from the offspring and ignore the rest. The recurrence of the tagged particle implies the local survival of COBRA.

Theorem 1.2.5. *On a d-regular tree the zealot voter model survives locally if* $p_0 = 0$ *and*

$$\mu > \frac{d}{\sqrt{d-1}+1}.$$

The proof uses some ideas from the proof of Lemma 4.57 in Liggett's 1999 book [17]. Combining this with Theorem 1.2.4, we notice that when $p_0 = 0$ the phase transition of local survival μ_l satisfies

$$\mu_l \in \left[\frac{d}{2\sqrt{d-1}},\ \frac{d}{1+\sqrt{d-1}}\right]$$

1.2.2 Results for Galton-Watson Trees

In a Galton-Watson process with $Z_0 = 1$ each individual in generation n has an independent and identically distributed number of children, which are members of generation $n + 1$. The Galton-Watson tree is the genealogy of this process. The one member of generation 0 is the root. Edges are drawn from each individual in generation n to their children. Let p_k be the probability of k children. We have assumed $p_k = 0$ unless $3 \leqslant d_{min} \leqslant k \leqslant M$, so the tree is infinite with probability 1, and all vertices have at most M children.

To prove an analogue of Theorem 1.2.2, we formulate our model as a voter model perturbation: let $\bar{p}_i = \varepsilon p_i$ when $i \neq 1$ and choose $\bar{p}_1$ to make the $\bar{p}_i$ sum to 1. A random walk that jumps to each neighbor at rate 1 has a reversible stationary distribution that is uniform on the graph. Let π_m be the fraction of vertices in the tree with degree m, and let $\mu_{m,k}$ be the expected number of surviving particles in the dual when we pick k neighbors of a vertex of degree m at random and run the coalescing random walk to time ∞.

Theorem 1.2.6. *Let $\delta > 0$. If ε is small then the COBRA dies out if*

$$\sum_m \pi_m \sum_k k p_k(\mu_{m,k} - 1) - p_0 < -\delta$$

and survives if the last quantity is $> \delta$.

This result can be easily proved using the techniques in [14]. The key idea is that when ε is small most of the steps in the dual are random walk steps, and the random walk is transient, so any coalescence occurs soon after branching, and the dual is essentially a coalescing branching random walk. These ideas go back to [7], where they were used on $\mathbb{Z}^d$ with $d \geqslant 3$. More recent applications include [6, 14, 19]. The zealot voter model has an additive dual, so things are simpler, and we can use the approach of [10]. In Section 4 we will provide more details about the method.

Remark. The last result concerns the survival of the dual, which is the same as the existence of a nontrivial stationary distribution for the zealot voter model.

Our next result concerns local survival. Given any Galton-Watson tree $\mathcal{T}^{GW}$, let M denote its maximal degree, and let $\mathcal{T}^M$ be the tree in which each vertex has M children. Let $\mu_l(G)$ denote the threshold for local survival of the COBRA on graph G. Note the expected number of new born particles at each time are the same on both trees. Since particles on tree $\mathcal{T}^M$ have more tendency to move further away from the root, a simple comparison leads to

$$\mu_l\left(\mathcal{T}^{GW}(\eta_t)\right) \leqslant \mu_l\left(\mathcal{T}^M(\eta_t)\right)$$

where η_t is the BRW without coalescence. The comments under Theorem 1.2.4 says for $p_0 = 0$,

$$\mu_l\left(\mathcal{T}^M(\eta_t)\right) = M/(2\sqrt{M-1})$$

It follows immediately that

Theorem 1.2.7. *If $p_0 = 0$ and $\mu < M/(2\sqrt{M-1})$ then COBRA and the zealot voter model both die out locally.*

Next we look for conditions implying local survival. On a tree we define the level ℓ_x of a vertex x to be its distance to the root. As on $d-$regular trees, our strategy is to follow a tagged particle and seek conditions guaranteeing its recurrence. Let X_t be the level of the tagged particle at time t. If ϕ is a harmonic function for the tagged particle process X_t, i.e. $\phi(X_t)$ is a martingale, then it follows from the optional stopping theorem that if T_0 is the time to hit the root and T_N is the first time the walk hits a site at level N

$$\phi(1) \geqslant \left(\min_{x:\ l_x=N} \phi(x)\right) P_1\left(T_N < T_0\right) \tag{1.9}$$

where the subscript 1 on P indicates that X_0 is at level 1. From (1.9) we see that if $\phi(x)$ goes to ∞ along all paths to ∞ in the tree, then the tagged particles is recurrent. In order for ϕ to be a harmonic function

$$\phi(x+1) - \phi(x) = \frac{p_x}{1-p_x}\left[\phi(x) - \phi(x-1)\right] = \frac{\mu}{d(x)-\mu}\left[\phi(x) - \phi(x-1)\right]$$

where $p_x = \mu/d(x)$ is the probability the tagged particle moves closer to the root. Taking logarithms, then this is

$$\log\left[\phi(x+1) - \phi(x)\right] = \log\left[\phi(x) - \phi(x-1)\right] + \log\left[\frac{\mu}{d(x)-\mu}\right]$$

As we will now explain, there is a natural mapping from the log-increments of the harmonic function to a branching random walk on $\mathbb{R}$. If we consider a particle at level x to be at $\log\left[\phi(x) - \phi(x-1)\right]$ on $\mathbb{R}$ then $d(x)-1$ new particles will be dispersed to

$$\log\left[\phi(x) - \phi(x-1)\right] + \log\left[\frac{\mu}{d(x)-\mu}\right].$$

As a result along any genealogical path, the distance between two consecutive generations is i.i.d with law the same as $\log[\mu/(d(x)-\mu)]$.

This process just described is different from the usual branching random walk in which children are dispersed independently from their parent. However Biggins [1] has proved results for more general branching random walks that contain ours as a special case. Let $F(t) = E(\zeta(-\infty, t])$ be the expected number of children that lie in $(-\infty, t]$ and define the Laplace transform of the mean measure by

$$m(\theta) = \int e^{-\theta t}\, dF(t)$$

Theorem 1.2.8. *If* $\min_{\theta \geqslant 0} m(\theta) < 1$ *then the leftmost particle in the branching random walk goes to* ∞*. This implies* ϕ *goes to* ∞ *along all paths to* ∞ *in the tree and we have local survival.*

To apply this result to our examples, we begin by noting that

$$m(\theta) = \sum_{j \geqslant 3} q_j (j-1) \left(\frac{j-\mu}{\mu} \right)^{\theta}$$

It is not easy to use this formula with Theorem 1.2.8 to get explicit predictions, so we focus on Galton-Watson tree with degrees only 3 and 4. Let $\mu = 3q_3 + 4q_4$ and

$$\nu(0) = \min_{\theta \geqslant 0} m(\theta).$$

We have computed the threshold for various μ in Section 3.3.3. See also Figure 1.1.

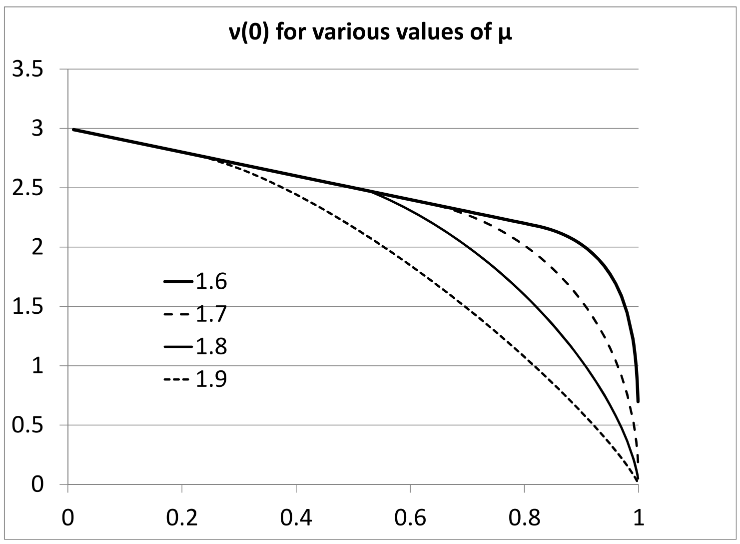


FIGURE 1.1: $\nu(0)$ as a function of q_3. Local survival occurs when $q_3 \geqslant 0.996$ for $\mu = 1.6$; $q_3 \geqslant 0.97$ for $\mu = 1.7$; $q_3 \geqslant 0.91$ for $\mu = 1.8$; and $q_3 \geqslant 0.82$ for $\mu = 1.9$.

2

The Latent Voter Model

In this Chapter we will study the latent voter model introduced in 2009 by Lambiotte, Saramaki, and Blondel [16]. In this model each individual owns one of two types of technology, say an iPad or a Microsoft Surface tablet. In the voter model on the d-dimensional lattice, individuals at times of a rate one Poisson process pick a neighbor at random and imitate their opinion. However, in the current interpretation of that model, it is unlikely that someone who has recently bought a new tablet computer will replace it, so we introduce latent states 1^* and 2^* in which individuals will not change their opinion. If an individual is in state 1 or 2 we call them active. Letting f_i be the fraction of neighbors in state i or i^*, the dynamics can be formulated as follows

$$\begin{array}{ll} 1 \to 2^* \text{ at rate } f_2 & \quad 1^* \to 1 \text{ at rate } \lambda \\ 2 \to 1^* \text{ at rate } f_1 & \quad 2^* \to 2 \text{ at rate } \lambda \end{array}$$

In [16] the authors showed that if individuals in the population interact equally with all the others then the system converges to a limit in which both technologies have frequency close to 1/2. Here, we will study the system with large λ, since in this case it is a voter model perturbation in the sense of Cox, Durrett, and Perkins

[7]. To explain this, we will construct the system using a graphical representation. Suppose first that the system takes place on $\mathbb{Z}^d$ and that $d \geqslant 3$. For each $x \in \mathbb{Z}^d$ and nearest neighbor y, we have independent Poisson processes $T_n^{x,y}$, $n \geqslant 1$. At each time $t = T_n^{x,y}$ we draw an arrow from $(y,t) \to (x,t)$ to indicate that if the individual at x is active at time t then they will imitate the opinion at y.

To implement the other part of the mechanism, we introduce for each site x, a Poisson process T_n^x, $n \geqslant 1$ of "wake-up dots" that return the voter to the active state.

- If there is only one voter arrow between two wake up dots, the result is an ordinary voter event.

- If between two wake up dots there are voter arrows to x from two different neighbors, an event of probability $O(\lambda^{-2})$, then x will change its opinion if and only at least one of the two neighbors has a different opinion. To check this, we note that if the first arrow causes a change then the second one is ignored, while if the first arrow comes from a site with the same opinion as the one at x then there will be a change if and only if the second site has an opinion different from the one at x.

- If t is fixed then at a given site there are $O(\lambda)$ wake-up dots by time t. Thus if we want to see the influence of intervals with two voter arrows then we want to run time at rate λ. The probability of k voter arrows between two wake-up dots is $(1+\lambda)^{-k}$, so in the limit the probability of three of more voter events between two wake-up dots goes to 0 as $\lambda \to \infty$.

If we let $\lambda = \varepsilon^{-2}$ and let $n_k(x)$ be the number of neighbors in state k or k^* then the rate of flips from 1 to 2 in the latent voter model when the configuration is ξ is:

$$\varepsilon^{-2} c^v_{1,2}(x,\xi) + h_{1,2}(x,\xi) \quad \text{where} \quad c^v_{1,2}(x,\xi) = 1_{\{\xi(x)=1\}} \frac{n_2(x)}{2d}$$

If we let $y_1, \dots y_{2d}$ be an enumeration of the nearest neighbors of x, the perturbation is

$$h_{1,2}(x,\xi) = 1_{\{\xi_t(x)=1\}} \frac{2}{(2d)^2} \sum_{1 \leqslant k < \ell \leqslant 2d} 1_{\{\xi(y_k) \text{ or } \xi(y_\ell) \in \{2,2^*\}\}}$$

Similar formulas hold when the roles of 1 and 2 are interchanged. $h_{1^*,j} = h_{2^*,j} \equiv 0$.

If we scale space by ε then Theorem 1.2 of [7] shows that under mild assumptions on the perturbation, the density of 1's in the rescaled particle system converges to the solution of the limiting PDE:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \phi(u) \quad \text{with} \quad \phi(u) = \langle h_{2,1}(0,\xi) - h_{1,2}(0,\xi) \rangle_u \tag{2.1}$$

and $\langle \cdot \rangle_u$ denotes the expected value with respect to the voter model with density u.

Intuitively, (2.1) holds because of a separation of time scales. The rapid voting means that the configuration near x looks like the voter model equilibrium with density $u(t,x)$. Later in the paper will show, see (2.11), that in the case of the latent voter model

$$\phi(u) = c_d u(1-u)(1-2u).$$

If we consider the latent voter model on a torus with n sites and let $\lambda_n \to \infty$ then the system can be analyzed using ideas from a recent paper of Cox and Durrett [5]. Define the density of 1's at time t by

$$U^n(t) = \frac{1}{n} \sum_x 1_{\{\xi_{\lambda t}(x)=1\}} \tag{2.2}$$

Theorem 2.0.1. *Suppose $n^{2/d} \ll \lambda_n \ll n$. If $U^n(0) \to u_0$ then $U^n(t)$ converges uniformly on compact sets to $u(t)$ the solution of*

$$\frac{du}{dt} = c_d u(1-u)(1-2u) \qquad u(0) = u_0$$

Remark 1. *The proof of this result is very similar to that of Theorem 6 in [5] so the proof is omitted. Note that the only thing we assume about the initial state is that the density* $U^n(0) \to u_0$*. Fast voting will turn the initial condition into a voter model equilibrium in a time that is* $o(\lambda_n)$*. If we consider the voter model without a latent period then the limiting differential equation is* $du/dt = 0$*. The last conclusion for the voter model is a very simple special case of the results in [5].*

2.0.1 Random graphs

We will explain the intuition behind Theorem 1.1.2 after we state our new result that replaces the torus by a random graph G_n generated by the *configuration model*. For the rest of the paper we will only consider the latent voter model on G_n. In the configuration model vertices have degree k with probability p_k. To create that graph we assign i.i.d. degrees d_i to the vertices and condition the sum $d_1 + \cdots + d_n$ to be even, which is a necessary condition for the values to be the degrees of a graph. We attach d_i half-edges to vertex i and then pair the half-edges at random. We will assume that

(A0) the graph G_n has no self-loops or parallel edges.

If $\sum_k k^2 p_k < \infty$ then the probability of (A0) is bounded away from 0 as $n \to \infty$. See Theorem 3.1.2 of [**?**]. The reader can consult Chapter 3 of that reference for more on the configuration model.

It seems likely that the results we prove here are true under the assumption that the degree distribution has finite second moment, but the presence of vertices of large degrees causes a number of technical problems. To avoid these we will assume:

(A1) $p_m = 0$ for $m > M$, i.e., the degree distribution is bounded.

In addition, we want a graph that is connected and has random walks with good mixing times, so we will also suppose:

(A2) $p_k = 0$ for $k \leqslant 2$.

The relevance of (A2) for mixing times will be explained in Section 2. Assumptions (A0), (A1) and (A2) will be in force throughout the paper.

On graphs that are not regular there are two versions of the voter model.

(i) The *site version* in which sites change their opinions at rate 1, and imitate a neighbor chosen at random,

$$c^s_{i,j}(x,\xi) = 1_{\{\xi(x)=i\}} \frac{n_j(x)}{d(x)}$$

where $n_j(x)$ is the number of neighbors of x in state j, and $d(x)$ is the degree of x.

(ii) The *edge version* in which each neighbor that is different from x causes its opinion to change at rate 1,

$$c^e_{i,j}(x,\xi) = 1_{\{\xi(x)=i\}} n_j(x).$$

The site version is perhaps the "obvious" generalization of the voter model on $\mathbb{Z}^d$, e.g., it is a special case of the general formulation used in Liggett [**?**]: x imitates y with probability $p(x,y)$, where p is a transition probability. However, the edge version has two special properties. First, in the words of [**?**] "magnetization is conserved," i.e., the number of 1's is a martingale. Second, if we consider the biased version in which after an edge (x,y) is picked a 1 at x always imitate a 2 at y but a 2 at x imitates a 1 at y with probability $\rho < 1$ then the probability a single 2 takes over a system that is otherwise all 1 is the same as the probability a simple random walk that jumps up with probability $1/(1+\rho)$ and down with probability $\rho/(1+\rho)$ never hits 0. This observation is due to Maruyama in 1970 [**?**], but has recently been rediscovered by [**?**], who call this version of the voter model "isothermal".

From our discussion of the graphical representation for latent voter model on $\mathbb{Z}^d$ it should be clear that the latent voter model on G_n is a voter model perturbation.

If we let $y_1, \dots y_{d(x)}$ be an enumeration of the neighbors of x, then in the site version

$$h^s_{1,2}(x,\xi) = 1_{\{\xi_t(x)=1\}} \frac{2}{(d(x))^2} \sum_{1 \leqslant k < \ell \leqslant d(x)} 1_{\{\xi(y_k) \text{ or } \xi(y_\ell) \in \{2,2*\}\}}$$

while in the edge version

$$h^e_{1,2}(x,\xi) = 1_{\{\xi_t(x)=1\}} \cdot 2 \sum_{1 \leqslant k < \ell \leqslant d(x)} 1_{\{\xi(y_k) \text{ or } \xi(y_\ell) = \in \{2,2*\}\}}$$

As before interchanging the roles of 1 and 2 we can define $h^s_{2,1}(x,\xi)$ and $h^e_{2,1}(x,\xi)$ while $h^s_{i,j}(x,\xi) = h^e_{i,j}(x,\xi) = 0$ when $i = 1^*$ or 2^*.

The last detail is to define the density $U^n(t)$. To do this we note that a random walk that jumps to a neighbor chosen at random has stationary distribution $\pi(x) = d(x)/D$, where $D = \sum_y d(y)$, while a random walk that jumps to each neighbor at rate 1 has stationary distribution $\pi(x) = 1/n$. To treat the two cases in one definition we let

$$U^n(t) = \sum_x \pi(x) 1_{\{\xi_{\lambda t}(x)=1\}} \tag{2.3}$$

Theorem 2.0.2. *Suppose that* $\log n \ll \lambda_n \ll n$. *If* $U^n(0) \to u_0$ *then* $U^n(t)$ *converges in probability and uniformly on compact sets to* $u(t)$, *the solution of*

$$\frac{du}{dt} = c_p u(1-u)(1-2u) \qquad u(0) = u_0. \tag{2.4}$$

where the value of c_p *depends on the degree distribution and the version of the voter model.*

Remark 2. *Again in the voter model without a latent period the limit is* $du/dt = 0$. *That result can be easily proved using the arguments for Theorem 1.1.1.*

2.0.2 Duality

To explain why Theorems 1.1.2 and 1.1.1 are true, we will introduce a dual process that is the key to the analysis. The dual process was first introduced more than 20 years ago by Durrett and Neuhauser [9], and is the key to work of Cox, Durrett, and Perkins [7]. To do this, we construct the process using a graphical representation that generalizes the one introduced for $\mathbb{Z}^d$. For each $x \in \mathbb{Z}^d$ and neighbor y, we have independent Poisson processes $T_n^{x,y}$, $n \geqslant 1$. At each time $t = T_n^{x,y}$ we draw an arrow from $(y, t) \to (x, t)$ to indicate that if the individual at x is active at time t then they will imitate the opinion at y. In the edge case all these processes have rate 1. In the site case $T_n^{x,y}$, $n \geqslant 0$ has rate $1/d(x)$. To implement the other part of the mechanism, we have for each site x, a rate λ Poisson process T_n^x, $n \geqslant 1$ of "wake-up dots" that return the voter to the active state.

To compute the state of x at time t we start with a particle at x at time t. To be precise $\zeta_0^{x,t} = \{x\}$. As we work backwards in time the particle does not move until the first time s there is an arrow $(y, t-s) \to (x, t-s)$.

- If this is the only voter arrow between the two adjacent wake-up dots then the particle jumps to y.
- If in the interval between the two adjacent wake-up dots there are arrows from k distinct y_i then the state changes to $\{x, y_1, \ldots y_k\}$ since we need to know the current state of all these points to know what change should occur in the process. In the limit as $\lambda \to \infty$ we will only see branchings that add two y_i. We include the case $k > 2$ to have the dual process well-defined.
- We do not need to know the order of the arrows because x will change if at least one of the y_i has a different opinion. When λ is small some of the y_i might change their state during the interval between the two wake-up dots but this

possibility has probability zero in the limit.

- To complete the definition of the dual, we declare that if a branching event adds a point already in the set, or if a particle jumps onto an occupied site then the two coalesce to one.

The dual process can be used to compute the state of x at time t. The first step is to work backwards in time to find $\zeta_t^{x,t}$ the set of sites at time 0 that can influence the state of x at time t. We note the states of the sites at time 0 and then work up the graphical representation to determine what changes should occur at the branching points in the dual.

To prove Theorem 1.1.2, Cox and Durrett [5] show that after a branching event any coalescence between the particle that branched and the two newly created particles will happen quickly, in time $O(1)$ or these particles will need time $O(n)$ to coalesce. (here we are using the original time scale.) Let $L = n^{1/d}$ be the side length of the torus. When $\lambda_n \gg n^{2/d}$ the particles will come to equilibrium on the torus before the next branching occurs in the dual, so we can forget about the relative location of the particles and we end up with an ODE limit. On the random graph, our assumption that all vertices have degree $\geqslant 3$ implies that the mixing time for random walks on these graphs is $O(\log n)$. Thus when $\lambda_n \gg \log n$, we have the situation that after a branching event there may be some coalescence in the dual at times $O(1)$ but then the existing particles will come to equilibrium on the graph before the next branching occurs in the dual. In both cases $\lambda_n \ll n$ is needed for the perturbation to have a nontrivial effect. Otherwise a collection of k random walks might all coalesce before the first branching arrow.

Remark 3. *There is no reason for having vertices of degree 0 in our graph. If $p_2 > 0$ and we look at the dynamics on the giant component then Theorem 1.1.1 will hold if $\log^2 n \ll \lambda_n \ll n$. The increase in the lower bound is needed to compensate for the fact* *that the mixing time for random walks on the graph is $O(\log^2 n)$. See e.g., Section 6.7 in [**?**]. Allowing $p_1 > 0$ should not change the behavior but, for simplicity, we derive our results under the assumption $d(x) \geqslant 3$.*

2.0.3 Long time survival

The latent voter model has two absorbing states $\equiv 1$ and $\equiv 2$. On a finite graph the latent voter model is a finite state Markov chain, so we know it will eventually reach one of them. However by analogy with the contact process on the torus [20] and on power-law random graphs, [21], this result should hold for times up to $\exp(\gamma n)$ for some $\gamma > 0$. Unfortunately we are only able to prove survival for any power of n. This failure is to due to our estimate of $P(\Omega_1^c)$ which appears on the right-hand side of Theorem 2.0.4.

Theorem 2.0.3. *Suppose that $\log n \ll \lambda_n \ll n$. Let $\varepsilon > 0$ and $k < \infty$. If $U^n(0) \to u_0 \in (0,1)$ there is a T_0 that depends on the initial density so that for any $m < \infty$ if n is large then with high probability*

$$|U^n(t) - 1/2| \leqslant 5\varepsilon \qquad \text{for all } t \in [T_0, n^m].$$

Remark 4. *Here and in what follows "with high probability" means with probability $\to 1$ as $n \to \infty$. Cox and Greven [**?**] have shown that for the nearest neighbor voter model on the torus in $d \geqslant 3$ that if we let θ_t be the fraction of sites in state 1 at time Nt then the configuration at time nt looks like the voter model equilibrium with density θ_t and θ_t changes according to the Wright-Fisher diffusion*

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

with β_d the probability that two random walks starting from adjacent sites never hit.

To prove Theorem 2.0.3, we use Theorem 4.2 of Darling and Norris [8]. To state their result we need to introduce some notation. To make it easier to compare with their paper we use their notation even though in some cases it conflicts with ours. Let ξ_t be a continuous time Markov chain with countable state space S and jump rates $q(\xi, \xi')$. In our case ξ_t will be the state of the voter model on the random graph. For their coordinate function $x : S \to \mathbb{R}^d$ we will take $d = 1$ and

$$x(\xi) = \sum_{x \in G_n} \pi(x) 1_{\{\xi(x)=1\}}.$$

We are interested in proving an ODE limit for $X_t = x(\xi_{\lambda t})$. Here and in what follows we drop the subscript n. To compare with the paper note that our ξ_t is their X_t and our X_t is their $\mathbf{X}_t$.

If we set $U = [0, 1]$ in [8] then we always have $x(\xi_t) \in U$ so their condition (2) is not needed. For each $\xi \in S$ we define the drift vector

$$\beta(\xi) = \sum_{\xi' \neq \xi} (x(\xi') - x(\xi)) q(\xi, \xi')$$

We let b be the drift of the proposed deterministic limit x_t:

$$x_t = x_0 + \int_0^t b(x_s)\, ds.$$

In our case $b(y) = cy(1 - y)(1 - 2y)$. To measure the size of the jumps we let $\sigma_\theta(y) = e^{\theta|y|} - 1 - \theta|y|$ and let

$$\phi(\xi, \theta) = \sum_{\xi' \neq \xi} \sigma_\theta(x(\xi') - x(\xi)) q(\xi, \xi').$$

Consider the events $\Omega_0 = \{|X_0 - x_0| \leqslant \eta\}$,

$$\Omega_1 = \left\{ \int_0^t |\beta(\xi_{\lambda s}) - b(X_s)|\, ds \leqslant \eta \right\},$$

$$\text{and } \Omega_2 = \left\{ \int_0^t \phi(\xi_s, \theta)\, ds \leqslant \theta^2 A t/2 \right\}.$$

The parameters in these events are coupled by the following relationships. If we let K be the Lipschitz constant of the drift b then $\eta = \varepsilon e^{-Kt_0}/3$ and $\theta = \eta/(At_0)$ where $A > 0$. We have changed their δ to η because we use δ in a number of our arguments in Section 3.

Theorem 2.0.4. *Under the conditions above, for each fixed t*

$$P\left(\sup_{s\leqslant t_0} |X_s - x_s| > \varepsilon\right) \leqslant 2e^{-\eta^2/(2At_0)} + P(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c)$$

In our application t_0 will be fixed and K is independent of n so η does not depend on n. To make the first term vanish in the limit we will take $A = n^{-1/2}$. To bound the probabilities on the right-hand side, we note

- We have jumps that change the density by $1/n$ at times of Poisson processes at total rate $\leqslant M\lambda n$. If $\theta|y|$ is small $\sigma_\theta(\pm 1/n) \sim \theta^2/2n^2$. By assumption $\lambda n \to 0$ so if n is large enough so that $M\lambda/n \leqslant \varepsilon A t_0/2$ then a standard large deviations estimate for the Poisson shows that $P(\Omega_2^c) \leqslant \exp(-c\lambda n)$.

- Our assumption that $U^n(0) \to u_0$ implies that $\Omega_0^c = \varnothing$ for large n.

- The hard work comes in estimating $P(\Omega_1^c)$, i.e., estimating the difference in the drift in the particle system from what we compute on the basis of the current density. We do this in Section 3.3-3.4 by computing the expected value of the mth moment of the difference $|\beta(\xi_s) - b(X_s)|$ so we end up with estimates that for a fixed time are $\leqslant n^{-m}$.

Once these three steps are done, the remainder of the proof of Theorem 2.0.3 given in Section 3.5 is routine. By subdividing the interval into small pieces we can use the single time estimates to control the supremum and hence the integral but only over a bounded time interval. However this is enough since it allows us to show that when the density wanders more than 4ε away from $1/2$, we can return it to within 2ε with probability n^{-m}, and in addition never have the difference exceed 5ε.

Theorem 1.1.1 is proved in Section 2 and Theorem 2.0.3 in Section 3. These results hold for other voter model perturbations such as the evolutionary games considered in [5]. However, the main obstacle to proving a general result is to find a formulation that works well on graphs with variable degrees. The arguments in the first proof closely parallel arguments in [5] but now use estimates for random walks on random graphs.

The keys to the second proof are results concerning the behavior of coalescing random walks (CRWs). There have been a number of studies of the time it takes for CRWs starting from every site of a random graph to coalesce to 1. Cooper et al [**?**, **?**] and Oliveira [**?**] considered coalescing random walks with one particle at each site and obtained results on the time needed for all particles to coalesce to 1. In [**?**] sufficient conditions were given for the number of particles in the coalescing random walk to converge to Kingman's coalescent. However, here we need estimates on the number of coalescences that can occur by time $C \log n$. This is done in Section 3.2.

2.1 Proof of Theorem 1.1.1

2.1.1 Mixing times for random walks

Bounds for the mixing times come from studying the conductance

$$Q(x,y) = \pi(x)q(x,y)$$

where π is the stationary distribution and $q(x,y)$ is the rate of jumping from x to y. In the site version $q(x,y) = 1/d(x)$ while $\pi(x) = d(x)/D$ when y is a neighbor of x, $y \sim x$, so $Q(x,y) = 1/D$ when $y \sim x$. In the edge version, $q(x,y) = 1$ if $y \sim x$, while $\pi(x) = 1/n$ where n is the number of vertices, so $Q(x,y) = 1/n$ when $y \sim x$. When degrees are bounded, the two conductances are the same up to a constant.

Define the isoperimetric constant by

$$h = \min_{\pi(S)\leqslant 1/2} \frac{Q(S,S^c)}{\pi(S)}$$

where $\pi(S) = \sum_{x\in S}\pi(x)$ and $Q(S,S^c) = \sum_{x\in S, y\in S^c} Q(x,y)$. Cheeger's inequality, see e.g. Theorem 6.2.1. in [?], implies that the spectral gap of Q, $\beta = 1-\lambda_1$ has

$$\frac{h^2}{2} \leqslant \beta \leqslant 2h \tag{2.5}$$

Using Theorem 6.1.2 in [?] we see that if $p_t(x,y)$ is the transition probability associated with Q

$$\Delta(t) \equiv \max_{x,y}\left|\frac{p_t(x,y)}{\pi(y)} - 1\right| \leqslant \frac{e^{-\beta t}}{\pi_{min}} \tag{2.6}$$

where $\pi_{min} = \min \pi(x)$.

Gkantsis, Mihail, and Saberi [?] have shown, see Theorem 6.3.2. in [?]:

Theorem 2.1.1. *Consider a random graph in which the minimum degree is $\geqslant 3$. There is a constant α_0 so that with high probability $h \geqslant \alpha_0$.*

Combining the last result with (2.5), (2.6), and the fact that $\pi_{min} \geqslant 1/(C_0 n)$ for large n, we see that

$$\Delta(t) \leqslant C_0 n e^{-\gamma t} \qquad \text{where } \gamma = \alpha_0^2/2.$$

If we let $C_1 = (6/\alpha_0^2)$ then for n large we have for $t \geqslant C_1 \log n$

$$\Delta(t) \leqslant 1/n \tag{2.7}$$

2.1.2 Our random graph is (almost) locally a tree

Recall that to construct our random graph we let $d_1, d_2, \dots d_n$ be i.i.d. from the degree distribution conditioned on $d_1 + \cdots + d_n$ to be even and then we pair the half-edges at random. Given a vertex x with degree $d(x)$, we let $y_1(x) \dots y_{d(x)}(x)$ be its neighbors. To grow the graph we let $V_0 = \{x\}$. On the first step we draw edges from x to $y_1(x) \dots y_{d(x)}(x)$ and let $V_1 = \{y_1(x), \dots, y_{d(x)}(x)\}$ which we consider to be an ordered list. If V_t has been constructed we let x_t be the first element of V_t and draw edges from x_t to $y_1(x_t) \dots y_{d(x_t)}(x_t)$. To create V_{t+1} we remove x_t and add the members of $y_1(x_t) \dots y_{d(x_t)}(x_t)$ not already in V_t.

We stop when we have determined the neighbors of all vertices at distance $<$ $(1/5)\log_M n$ from x. A simple calculation using branching processes shows that the total number of neighbors within that distance of x is $\leqslant n^{1/5} \log n$ for large n. The $\log n$ takes care of the limiting random variable. Thus in the construction we will generate $\leqslant M n^{1/5} \log n$ connections. We say that a collision occurs at time t if we connect to a vertex already in V_t. The probability of a collision on single connection is $\leqslant M n^{-4/5} \log n$. The expected number of collisions involving the first $n^{1/5} \log n$ sites is $\leqslant C M n^{-3/5} \log^2 n$, so for most starting points (but not all) the graph will be a tree. To get a conclusion that applies to all starting points we note that the probability of two collisions in the construction starting from one site is

$$\leqslant \binom{CMn^{1/5}\log n}{2}(n^{-4/5}\log^2 n)^2 = O(n^{-6/5}\log^6 n)$$

As we build up the graph we first find all of the neighbors of vertices at distance 1 from x then distance 2, etc. Thus when a collision occurs it will connect a vertex at distance k with one at distance k or to one at distance $k+1$ that already has a neighbor at distance k. As we will explain after the proof of the next lemma, this makes very little difference.

2.1.3 Results for hitting times

Lemma 2.1.2. *Once two particles are at distance r_n with $1 \leqslant r_n \leqslant (1/25)\log n$ then with probability $\geqslant 1 - 2^{1-r_n}$, they will reach a distance $5r_n$ before hitting each other.*

Proof. For the proof we will pretend that the graph is exactly a tree up to distance $5r_n$. We return to this issue in a remark after the proof. Let Z_t be the distance between these two particles and let T_m be the first time the distance is m. Note that on each jump, with probability $p \geqslant 2/3$, the particles get 1 step further apart, while with probability $\leqslant 1/3$, the particles get one step closer. This implies that $\phi(z) = (1/2)^z$ is a supermartingale, so

$$\phi(r_n) \geqslant P_{r_n}(T_0 < T_{5r_n})\phi(0) + (1 - P_{r_n}(T_0 < T_{5r_n}))\phi(5r_n).$$

Rearranging gives

$$P_{r_n}(T_0 < T_{5r_n}) \leqslant \frac{\phi(5r_n) - \phi(r_n)}{\phi(5r_n) - \phi(0)} \leqslant \frac{2^{-r_n}}{1 - 2^{-5}} \leqslant 2^{1-r_n} \tag{2.8}$$

as $n \to \infty$ which proves the desired result. □

Remark 5. *As noted after the construction, when a collision occurs it will connect a vertex at distance k with one at distance k or to one at distance $k+1$ that already* *has a neighbor at distance k. In the first case at distance k the comparison chain moves towards x with probability $\leqslant 1/3$, the chain stays at the same distance with probability $\leqslant 1/3$ and moves further away with probability $\geqslant 1/3$. In the second case at distance $k+1$ the comparison chain moves toward the root with probability $\leqslant 2/3$ and further away with probability $\geqslant 1/3$.*

If we have a birth and death chain X_n that jumps $p(k,k+1) = p_k$, $p(k,k) = r_k$ and $p(k,k-1) = q_k$ then

$$\phi(k+1) - \phi(k) = \frac{q_k}{p_k}[\phi(k) - \phi(k-1)]$$

recursively defines a function ϕ so that $\phi(X_n)$ is a martingale. In our comparison chain $q_k/p_k = 1/2$ for all but one value of k, which has $q_k/p_k \leqslant 1$, so calculations like the one in (2.8) will work but give a slightly larger constant. Because of this, we will suppress these annoying details by assuming the graph is exactly a tree up to distance $(1/5)\log n$.

To prepare for the next result we need

Lemma 2.1.3. *If S_k is the sum of k independent mean one exponentials then*

$$P(S_k \leqslant ak) \leqslant \left(\frac{ae}{1+a}\right)^k$$

Remark 6. *This holds for all a but is only useful when $ae/(1+a) < 1$, which holds if $a < 1/2$.*

Proof. Let $\theta > 0$ and note $\int_0^\infty e^{-\theta x} e^{-x}\,dx = 1/(1+\theta)$. Using Markov's inequality we have

$$e^{-\theta ak} P(S_k \leqslant ak) \leqslant (1+\theta)^{-k}$$

Taking $\theta = 1/a$ and rearranging gives the desired result. □

Lemma 2.1.4. *Suppose two particles are a distance $r_n = 2\log_2 \log n$. Then with high probability the two particles will not collide by time $\log^2 n$.*

Proof. By Lemma 2.1.2 the probability of hitting $5r_n$ before 0 starting from r_n is

$$\geqslant 1 - 2/(\log n)^2 \tag{2.9}$$

A particle must make $4r_n$ jumps to go from distance $5r_n$ to r_n. Since jumps occur at rate 1 in the site model and at rate $\leqslant M$ in the edge model, the last lemma implies that the probability of $k = r_n$ jumps in time $\leqslant ar_n/M$ is

$$\leqslant (ae)^{r_n} \leqslant 1/(\log^3 n)$$

for large n if a is small enough. If a particle makes $2M(\log^2 n)/ar_n$ attempts to reach 0 before $5r_n$ starting from r_n then (2.9) implies that with high probability it will not be successful, while the last bound implies that this number of attempts will take time $\geqslant 2\log^2 n$ with high probability. □

Lemma 2.1.5. *Suppose two particles are a distance $r_n = 2\log_2 \log n$ and let $s_n/n \to 0$. Then with high probability the two particles will not hit by time s_n.*

Proof. Lemma 2.1.4 takes care of times up to $\log^2 n$. The result in (2.7) implies that if n is large then for $t \geqslant \log^2 n$, $p_t(x, y) \leqslant 2/n$. Summing we see that if the two particle move independently the expected amount of time the two particles spend at the same site at times in $[\log^2 n, s_n]$ is $\leqslant 2s_n/n \to 0$. Since the jump rates are bounded above this implies the desired result. □

In the proof of Lemma 2.2.10 we need the following result for hitting times.

Lemma 2.1.6. *Let $L = (1/5)\log_M n$. Suppose two particles performing independent continuous time (site or edge) random walks start at points are separated by distance k. Then there is a constant so that the probability the two particles hit by time $C_1 \log n$ is $\leqslant C2^{-(k\wedge L/2)}$.*

Proof. Using the bound in Lemma 2.1.3 as in the proof of Lemma 2.1.4, if we replace C_1 by $C_2 = 4MC_1$ then we can treat the discrete time random walk in which on each step, one of the particles is picked at random to jump. L is chosen so that arguments in Section 2.2 show that when either particle looks at the ball of radius L around them, what they see differs from like a tree by at most one edge. If we ignore the extra edge, which is justified by Remark 5, and if we let D_n be the distance between the two particles after n jumps, then (2.8) implies that if $k \leqslant L/2$

$$P_k(T_0 < T_L) \leqslant \frac{2^{-k} - 2^{-L}}{1 - 2^{-L}} \leqslant 2 \cdot 2^{-k}$$

if n is large.

Suppose now that the distance between the two points is $\geqslant L/2$. The particles cannot hit until they first reach a distance of $L/2$, at which point the previous estimate can be applied. The journey from $k \leqslant L/2$ to L takes at least $L/2$ steps. Thus if a particle makes K cycles from $L/2$ to L it has used up $KL/2$ units of time which is larger than $C_2 \log n$ if K is chosen large enough. This shows that the estimate holds with $C = 2K$. □

2.1.4 Results for the dual process

In this section we will consider the dual process starting from a single site on its original time scale, i.e., jumps occur at rate $O(1)$. In either version of the model, the rate at which branching occurs is $\leqslant L/\lambda$ where $L = M^2$. (Here we are using the fact in the edge model the degree is bounded.) Let R_n be the time of the nth branching. If $t_n = c_2 \log n$ for some constant $c_2 > 0$ then

$$P(R_1 \leqslant t_n) \to 0 \qquad \textit{as } n \to \infty$$

Let $N(t)$ be the number of branching events by time λt. Comparing with a branching process we have $EN(t) \leqslant e^{Lt}$. The expected number of branchings in the interval

$[\lambda t - t_n, \lambda t]$ is $\leqslant e^{Lt}(c_2 \log n)/\lambda$ so as $n \to \infty$,

$$P(\lambda t - R_{N(t)} \leqslant t_n) \to 0 \tag{2.10}$$

In the next three results C_1 is the constant defined in (2.7) and we make the following assumption:

($\star$) Suppose there are k particles in the dual at time 0, and each pair are separated by a distance $r_n = 2 \log_2 \log n$.

Lemma 2.1.7. *Suppose that at time 0, the first particle encounters a branching event. By time $C_1 \log n$, there may be coalescences between new born particles or with their parent, but with high probability there will be no other coalescences.*

Proof. This follows from Lemma 2.1.4. □

Lemma 2.1.8. *At time $C_1 \log n$ all the particles are almost uniformly distributed on the graph with the bound on the total variation distance uniform over all configurations allowed by ($\star$).*

Proof. It follows from Lemma 2.1.4 that with high probability no pair of particles will collide by time $C_1 \log n$. If we consider k independent random walks then the bound in (7) implies that the total variation distance between their joint distribution and the k-fold product of π is $\leqslant k/n$. Combining the last two observations with the triangle ineuqality gives the desired result □

Lemma 2.1.9. *After time $C_1 \log n$, with high probability there is no coalescence between particles before the next branching event, and right before the next branching event, all the particles are r_n apart away from each other.*

Proof. The claim about coalescence follows from Lemma 2.1.5. The branching time is random but it is independent of the movement of the particles, so the result about the separation between particles follows from (2.7), since under the k-fold product of π the probability some pair of particles is within distance r_n is $\leqslant \binom{k}{2} M^{2\log_2 \log n} M/n \to 0$. $\square$

Together with (2.10), Lemma 2.1.9 implies that there is no coalescence in the dual $[R_{N(t)}, \lambda t]$ and particles are at least r_n apart right before $R_{N(t)}$. According to Lemma 2.1.8, the coalescences between new born particles and their parents can only happen before $R_{N(t)} + C_1 \log n$, with no other coalescences. Lemma 2.1.8 tells us at times $\geqslant R_{N(t)} + C_1 \log n$, all the particles are almost uniformly distributed over the graph. Thus when we feed values into the dual process to begin to compute the state of x at time t the values are independent and equal to 1 with probability u.

Lemma 2.1.10. *For fixed t, $EU^n(t)$ converges to a limit $u(t)$.*

Proof. Let $Z(s)$, $s \leqslant t$ be the number of particles in the dual process, when we impose the rule that the number of particles is not increased until time $(C_1 \log n)/\lambda$ after a branching event. Our results imply that $Z(s)$ converges to a branching process. The last result shows that when we use the dual to compute the state of x at time t we put independent and identically distributed values at the $Z(t)$ sites. The result now follows from results in [7]. $\square$

Lemma 2.1.11. *For fixed t, $U^n(t) - EU^n(t)$ converges in probability to 0.*

Proof. It follows from Lemma 2.1.5 that if $|x - y| > r_n$ then there will with high probability be no collisions between particles in the dual processes starting from x and y, and hence the values we compute for x and y are almost independent. In this case if n is large

$$\text{var}\left(\sum_x 1_{\{\xi_\lambda t(x)=1\}}\right) \leqslant \varepsilon n^2 + nM^{2\log_2 \log n}$$

Since ε is arbitrary, an application of Chebysev's inequality gives the desired result. $\square$

It is straightforward to show that the last two conclusions remain valid when $t_n \to t$. Once this is done we have convergence uniform on compact sets. To see this suppose that there is an ϵ, a seqeunce $n(k) \to \infty$ and a sequence of points $s_k \in [0, T]$ with $|EU_{n(k)}(s_k) - u(s_k)| > \varepsilon$ [or $P(|U_{n(k)}(s_k) - EU_{n(k)}(s_k)| > \varepsilon) > \varepsilon$] then by finding a convergent subsequence of the s_k we would have a contradiction.

2.1.5 Computation of the reaction term

The final step is to show that $u(t)$ satisfies the differential equation. To warm up for the real proof, we begin by doing this on $\mathbb{Z}^d$ If ν_u is the voter model stationary distribution with density u and v_1 and v_2 are randomly chosen neighbors of x then

$$\langle h_{1,2}(x, \xi) \rangle_u = \nu_u(\xi(x) = 1, \xi(v_1) = 2 \text{ or } \xi(v_2) = 2)$$

The right-hand side can be computed using the duality between the voter model and coalescing random walk. Following the approach in Section 4 of [**?**] if we let $p(x|y|z)$ be the probability the random walks starting from x, y, and z never hit and $p(x|y, z)$ be the probability y and z coalesce but don't hit x then

$$\nu_u(\xi(x) = 1, \xi(y) = 2 \text{ or } \xi(z) = 2) = p(x|y|z)u(1 - u^2) + q(x, y, z)u(1 - u)$$

where $q(x, y, z) = p(x|y, z) + p(x, y|z) + p(x, z|y)$

Using this identity we can compute the reaction term defined in (2.1)

$$\begin{aligned} \phi(u) &= \langle h_{2,1}(x, \xi) - h_{1,2}(x, \xi) \rangle_u \\ &= p(x|v_1|v_2)(1 - u)(1 - (1 - u)^2) + q(x, v_1, v_2)u(1 - u) \\ &\qquad - [p(x|v_1|v_2)u(1 - u^2) + q(x, v_1, v_2)u(1 - u)] \\ &= p(x|v_1|v_2)[(1 - u)u(2 - u) - u(1 - u)(1 + u)] \\ &= p(x|v_1|v_2)u(1 - u)(1 - 2u) \end{aligned} \tag{2.11}$$

The computations for the random graph are similar but in that setting we have to take into account the degree of x and what the graph looks like locally seen from

x. Let q_k be the size-biased distribution kp_k/μ where $\mu = \sum_k p_k$ is the mean degree. Let $\mathbb{P}_k$ be a Galton Watson tree in which the root has degree k and the other vertices have j children with probability q_{j+1}.

In the site version a dual random walk path will spend a fraction $\pi^s(k) = q_k$ at vertices with degree k so

$$\langle h^s_{2,1} - h^s_{1,2} \rangle_u = \sum_k q_k \mathbb{P}_k(x|v_1|v_2) u(1-u)(1-2u)$$

where v_1 and v_2 are randomly chosen neighbors of the root. In the edge version $\pi^e(k) = p_k$ so

$$\langle h^e_{2,1} - h^e_{1,2} \rangle_u = \sum_k p_k \mathbb{P}_k(x|y|z) u(1-u)(1-2u)$$

2.2 Proof of Theorem 2.0.3

Recall that the density in the time-rescaled latent voter model is given by:

$$X_t = \sum_{x \in G_n} \pi(x) 1_{(\xi_{\lambda t}(x)=1)}. \tag{2.12}$$

To complete the proof of Theorem 2.0.3 using the result of Darling and Norris [8] given in Theorem 2.0.4 we need to estimate the probability of

$$\Omega_1^c = \left\{ \int_0^t |\beta(X_s) - b(X_s)|\, ds > \eta \right\} \tag{2.13}$$

where $\beta(\xi) = \sum_{\xi' \neq \xi} (x(\xi') - x(\xi)) q(\xi, \xi')$ is the drift in the particle system and $b(u) = c_p u(1-u)(1-2u)$ is the drift in the ODE.

To begin to do this, we define $\tilde{\xi}_s$ to be the same as ξ_s for time $s \leqslant \lambda t - C_1 \log n$, while on the time interval $(\lambda t - C_1 \log n, \lambda t]$, $\tilde{\xi}$ only follows the paths from voter events of ξ, ignoring those from branching events. Let

$$\tilde{X}_t = \sum_{x \in G_n} \pi(x) 1_{\{\tilde{\xi}_{\lambda t}(x)=1\}}$$

be the density of this new process $\tilde{\xi}$. In order to determine $\tilde{\xi}_{\lambda t}$, we run the coalescing random walks backward in time, starting from time λt and stopping at time $\lambda t - C_1 \log n$.

We will first outline the proof then go back and fill in the details. In Section 3.1 we will prove

Lemma 2.2.1. *For any $\varepsilon > 0$ and $m < \infty$ the probability that more than εn sites are changed by branching arrow is $\leqslant n^{-m}$ for large n.*

Let $\tilde{u} = x(\xi_{\lambda t - C_1 \log n})$. To bound $P(\Omega_1^c)$ we will prove:

Lemma 2.2.2. *Suppose* $\log n \ll \lambda_n \ll n$ *and* $m > 0$*. There is a* $\delta > 0$ *independent of* m *and constants* C_m *so that for any* $\varepsilon > 0$ *if* $n \geqslant n_0(m)$

$$P\left(|\tilde{X}_t - \tilde{u}| > \epsilon | \mathcal{F}_{t-(C_1 \log n)/\lambda}\right) \leqslant \frac{C_m}{\varepsilon^{2m} n^{\delta m}} \tag{2.14}$$

This will follow from Chebyshev's inequality once we have a suitable estimate on the $2m$th moment (see Lemma 2.2.3 below). Using (2.12) we have

$$\tilde{X}_t - \tilde{u} = \sum_{x \in G_n} \pi(x)[1_{(\xi_{\lambda t}(x)=1)} - \tilde{u}],$$

so if we let $Y(x) = 1_{(\xi_{\lambda t}(x)=1)} - \tilde{u}$ then

$$E(\tilde{X}_t - \tilde{u})^{2m} = \sum_{x_1, \ldots x_{2m}} \pi(x_1) \cdots \pi(x_{2m}) \cdot Y(x_1) \cdots Y(x_{2m}).$$

We will use π^k to denote the distribution $\pi \times \cdots \times \pi$ on G_n^k. If we introduce the dual coalescing random walks $W_1, \ldots, W_{2m}$ starting from distribution π^{2m} and let $r = C_1 \log n$ then we can write this as

$$E[Y(W_1(r)) \ldots Y(W_{2m}(r))]$$

To estimate probabilities for coalescing random walks we introduce independent random walks $W_1', \ldots, W_{2m}'$ starting from distribution π^{2m}, and use these to construct the $W_1, \ldots, W_{2m}$ by dropping the higher number after collisions. To simplify notation let $Z_i = Y(W_i(C_1 \log n))$ and $Z_i' = Y(W_i'(C_1 \log n))$.

Lemma 2.2.3. *There is* $\delta > 0$ *so that for each* m *we have*

$$|E[Z_1 \ldots Z_{2m}]| \leqslant C_m n^{-\delta m}$$

This will be proved in Section 3.3 after we bound the coalescence probabilities in Section 3.2.

2.2.1 Ignoring branching

To prove Lemma 2.2.1 we begin by noting that since the branching rate is $1/\lambda$ we can suppose without loss of generality that $\lambda \leqslant n^{1/2}$. In order for a site to be changed a branching arrow must hit the dual process for the site but one branching arrow can change multiple sites. Consider the coalescing random walk starting from all sites occupied and let T_k be the amount of space time in $[0, C_1 \log n]$ occupied by k-particles, i.e., a particle that is a coalescence of k particles. Clearly

$$\sum_{k=1}^{\infty} kT_k = C_1 n \log n \tag{2.15}$$

Let Π_k be the number of branching arrows that hit k-particles. A standard large deviations result (see e.g., (2.6.2) and Exercise 3.1.4 in [?]) shows that there is a constant C_2 so that if $Z_\mu = \text{Poisson}(\mu)$ then

$$P(Z_\mu > 2\mu) \leqslant \exp(-C_2\mu)$$

If $T_k \geqslant n^{2/3}$ then $P(\Pi_k > 2T_k/\lambda) \leqslant \exp(-C_2 n^{2/3}/\lambda)$, so if $\mathcal{K} = \{k : T_k \geqslant n^{2/3}\}$ then using (2.15)

$$P\left(\sum_{k\in\mathcal{K}} k\Pi_k > 2C_1(n\log n)/\lambda\right) \leqslant |\mathcal{K}| \exp(-C_2 n^{1/6})$$

To control the contribution from particles of large weight, we will get a bound on the largest particle weight seen. Let $N_x(s)$ be the size of the cluster containing the particle that started at x at time t when we run the coalescing random walk to time $t-s$ and let $N_{max}(s)$ be the size of the largest cluster. We will show

Lemma 2.2.4. *If $\alpha > 0$ and $m < \infty$ and $t = C_1 \log n$ then for large n*

$$P(N_{max}(t) > n^\alpha) \leqslant n^{-m}.$$

When $k \notin \mathcal{K}$ monotonicity implies $P(\Pi_k > 2n^{2/3}/\lambda) \leqslant \exp(-c_2 n^{1/6})$, so if n is large

$$P\left(\sum_{k \notin \mathcal{K}} k\Pi_k > n^{2\alpha} n^{2/3}/\lambda\right) \leqslant n^{-m} + n^{\alpha} \exp(-c_2 n^{1/6})$$

which proves Lemma 2.2.1. It remains then to prove Lemma 2.2.4.

We begin by considering the edge model.

Lemma 2.2.5. *If* $s \geqslant 1/2M$ *then* $\mathbb{E}(N_x(s) - 1) \leqslant 4Mes$*.*

Proof. Let $y \neq x$ and W^y be the edge random walk starting from y. Noting that when W^x and W^y hit, they stay together for a time $\geqslant 1/2M$ with probability e^{-1} gives

$$\mathbb{P}(W^x \textit{ and } W^y \textit{ hit by time } s) \times \frac{1}{2Me} \leqslant \int_0^{s+1/2M} \sum_z p_r(x,z) p_r(y,z)\, dr$$

Since the edge random walks are reversible with respect to the uniform distribution, the transition probability is symmetric

$$\int_0^{s+1/2M} \sum_z p_r(x,z) p_r(y,z)\, dr = \int_0^{s+1/2M} \sum_z p_r(x,z) p_r(z,y)\, dr \tag{2.16}$$

$$= \int_0^{s+1/2M} p_{2r}(x,y)\, dr \tag{2.17}$$

Using this we have

$$EN_x(s) = \sum_y \mathbb{P}(W^x \textit{ and } W^y \textit{ hit by time } s) \leqslant 2Me \int_0^{s+1/2M} dr \leqslant 4Mes$$

where in the last step we have used $s \geqslant 1/2M$ □

Our next step is to bound the second moment of $N_x(t)$.

Lemma 2.2.6. *If* $s \geqslant 1/2M$ *then* $\mathbb{E}(N_x(s) - 1)(N_x(s) - 2) \leqslant 3(4Mes)^2$*.*

Proof. We begin by observing that

$$\mathbb{E}(N_x(s)-1)(N_x(s)-2) = \sum_{x_1,x_2} P(x_1, x_2 \in N_x(s)).$$

where the sum is over $x_i \neq x$ and $x_1 \neq x_2$. We first consider the case in which x and x_1 are the first to collide, and we bound

$$\sum_{x_1,x_2,y,z} \int_0^{s+1/2M} p_r(x,y)p_r(x_1,y)p_r(x_2,z)P(z \in N_{y,r}(s))\,dr$$

where $N_{y,r}(s)$ is the cluster at time s of the random walk that starts at y at time r. As in the previous proof $2Me$ times this quantity will bound the desired hitting probability. By symmetry $\sum_{x_2} p(x_2, z) = \sum_{x_2} p(z, x_2) = 1$. Using Lemma 2.2.5

$$\sum_z P(z \in N_{y,r}(s)) \leqslant 4Mes$$

Using reversibility we can write what remains of the sum as

$$\sum_{x_1,y} \int_0^{s+1/2M} p_r(x,y)p_r(y,x_1)\,dr = \sum_{x_1} \int_0^{s+1/2M} p_{2r}(x,x_1)\,dr \leqslant 2s \tag{2.18}$$

The second case to consider is when x_1 and x_2 are the first to collide, and we bound

$$\sum_{x_1,x_2,y,z} \int_0^{s+1/2M} p_r(x_1,y)p_r(x_2,y)p_r(x,z)P(z \in N_{y,r}(s))\,dr$$

Using symmetry $p_r(x_1,y)p_r(x_2,y) = p_r(y,x_1)p_r(y,x_2)$ then summing over x_1, x_2 we have

$$\leqslant \sum_{y,z} \int_0^{s+1/2M} p_r(x,z)P(z \in N_{y,r}(s))\,dr$$

We have $P(z \in N_{y,r}(s)) = P(y \in N_{z,r}(s))$ because either event says y and z coalesce in $[r, s]$, so summing over y and using Lemma 2.2.5 the above is

$$\leqslant (4Mes) \sum_z \int_0^{s+1/2M} p_r(x, z)\, dr \leqslant (4Mes) \cdot 2s \tag{2.19}$$

Combining our calculations proves the desired result. □

Lemma 2.2.7. *If* $s \geqslant 1/2M$ *then* $\mathbb{E}[(N_x(s) - 1) \cdots (N_k(s) - k)] \leqslant C_k(4Mes)^k$ *and hence*

$$\mathbb{E}N_x^m(s) \leqslant C_{m,M}(1 + s)^m$$

Proof. The second result follows easily from the first since

$$x^m = 1 + \sum_{k=1}^{m} c_{m,k}(x - 1) \cdots (x - k)$$

The first case is

$$\sum_{\substack{x_1,\ldots,x_k,\\ y,z_1,\ldots z_{k-1}}} \int_0^{s+1/2M} p_r(x, y) p_r(x_1, y) p_r(x_2, z_1) \ldots p_r(x_k, z_{k-1}) P(z_1, \ldots z_{k-1} \in N_{y,r}(s))\, dr$$

Using symmetry and summing over $x_2, \ldots, x_k$ removes the $p_r(x_2, z_1) \ldots p_r(x_k, z_{k-1})$ from the sum. Next we sum over $z_1, \ldots z_{k-1}$ (which are distinct) and use induction to bound the sum by $C_{k-1}(4Mes)^{k-1}$. Finally we finish up by applying (2.18).

The second case is

$$\sum_{\substack{x_1,\ldots,x_k,\\ y,z_1,\ldots z_{k-1}}} \int_0^{s+1/2M} p_r(x_1, y) p_r(x_2, y) p_r(x_3, z_1) \ldots p_r(x_k, z_{k-2})$$
$$p_r(x, z_{k-1}) P(z_1, \ldots z_{k-1} \in N_{y,r}(s))\, dr$$

Using symmetry and summing over $x_1, \ldots, x_k$ removes the

$$p_r(x_1, y) p_r(x_2, y) p_r(x_3, z_1) \ldots p_r(x_k, z_{k-2}).$$

As in the previous proof $P(z_1, \dots z_{k-1} \in N_{y,r}(s)) = P(z_1, \dots z_{k-2}, y \in N_{z_{k-1},r}(s))$, so summing over $z_1, \dots, z_{k-2}, y$ and using induction we can bound the sum by $C_{k-1}(4Mes)^{k-1}$. Finally we finish up by applying (2.19) with $z = z_{k-1}$ □

Remark 7. *To extend to the site case where we do not have symmetry, we note that reversibility of this model with respect to* $\pi(y) = d(y)/D$ *implies*

$$p_r(y, z) \leqslant d(y)p_r(y, z) = d(z)p_r(z, y) \leqslant Mp_r(z, y)$$

so the proof works as before but we accumulate a factor of M *each time we use symmetry.*

Now we are ready to give an upper bound on the size of the maximal cluster $N_{max}(t)$ at time λt. Here and for the rest of the proof of Lemma 2.2.2, we only use moment bounds so the proof is the same for the edge and site models

Proof of Lemma 2.2.4. By Chebyshev's inequality

$$n^{\alpha k} P(N_x(t) > n^\alpha) \leqslant C_{k,M}(1+t)^k$$

If we pick $k > (m+1)/\alpha$ then

$$P\left(\max_x N_x(t) > n^\alpha\right) \leqslant \frac{n}{n^{k\alpha}} C_{k,M}(2\log^2 n)^k = o(n^{-m})$$

which proves the desired result. □

2.2.2 Bounds on coalescence probabilities

Recall that $W_1, \ldots, W_{2m}$ are coalescing random walks starting from distribution π^{2m}

Lemma 2.2.8. *Let H_{12} be the event that W_1 and W_2 hit by time $C_1 \log n$.*

$$P(H_{12}) \leqslant (C \log n)/n.$$

Proof. Let $\Delta = \{(v, v) : 1 \leqslant v \leqslant n\} \subset G_n \times G_n$ be the "diagonal." Let W_1' and W_2' be independent random walks. Since $(W_1'(t), W_2'(t)) =_d \pi^2$ for all $t \geqslant 0$, the expected occupation time of Δ is $(C_1 \log n)\pi^2(\Delta)$. In the edge case π is uniform so $\pi^2(\Delta) = 1/n$. In the site case if we let $d(x)$ be the degree of x and $D = \sum_x d(x)$ then $\pi(x) = d(x)/D$ so

$$\pi^2(\Delta) = \sum_x \frac{d(x)^2}{D^2} \leqslant \frac{M^2}{n}$$

since $d(x) \leqslant M$, $D \geqslant n$, and $|G_n| = n$.

The jump rate for $(W_1'(t), W_2'(t))$ is 2 in the site case and $\leqslant 2M$ in the edge case, so when W_1' and W_2' hit the expected time they spend together is $\geqslant 1/2M$, and we have

$$P(H_{12}) \leqslant (C_1 \log n)\frac{M^2}{n} \cdot 2M$$

which proves the desired result. □

Remark 8. *In what follows we will prove the result only for the site case, since the time change argument in the last paragraph of the proof can be used to extend the argument to the edge case.*

The computation of higher order coalescence probabilities is made complicated by the fact that if particles 1 and 2 are the first to coalesce at time $T_{1,2}$ then the joint distribution of (W_1, W_3, W_4) at time T_{12} is not π^3. To avoid some of these difficulties, we will estimate the probability that coalescences occur in a specific pattern, and ignore collisions not consistent with the pattern. For example let $H_{12,34}$ be the event that

- 1 and 2 coalesce at T_{12}. We ignore collisions involving particles 3 and 4 before that time.

- At time $T_{12,34} > T_{12}$ particles 3 and 4 coalesce. We ignore particle 1 on $[T_{12}, T_{12,34}]$.

We say particle 1 in the second bullet because, in our coupling, at time T_{12} we drop W_2' and use W_1' to move the coalesced particle.

The act of ignoring particles may look odd but the reasoning is legitimate. Doing this enlarges the event that the coalescences occurred in the indicated pattern leading to an over estimate of the probabilities of interest.

Lemma 2.2.9. *$P(H_{12,34}) \leqslant C(\log^2 n)/n^2$.*

Proof. By Lemma 2.2.8, the probability of the event H_{12} that 1 and 2 hit by time $C_1 \log n$ is $\leqslant C(\log n)/n$. Since we are ignoring particles 3 and 4 up to time T_{12}, their joint distribution at time T_{12} conditional on $H_{1,2}$ is π^2. This implies that

$$P(H_{12,34}) \leqslant P(H_{12}) \cdot (C \log n)/n$$

which proves the desired result. □

Consider now the event $H_{12,3}$

- 1 and 2 coalesce at time T_{12}. We ignore collisions involving particles 3 and 4 before that time.

- 3 coalesces with particle 1 at $T_{12,3} > T_{12}$. We ignore particle 4 during $[T_{12}, T_{12,3}]$.

Lemma 2.2.10. *There is an $\delta > 0$ so that $P(H_{12,3}) \leqslant C \log n/n^{1+\delta}$*

Proof. As in the previous argument the probability that 1 and 2 hit is $\leqslant C(\log n)/n$ and at time T_{12} the location of W_3' has distribution π and is independent of W_1'. Unfortunately W_1' does not have distribution π since it may be easier for particles 1 and 2 to coalesce at some points than others.

Let π_{min} be the minimal value of $\pi(x)$ and let $L = (1/5)\log_M n$. Using the observations that (i) since the max degree is $M \geqslant 3$, the number of vertices at distance k from a fixed vertex is $\leqslant M^k$ and (ii) by Lemma 2.1.6 the probability two particles separated by $k \leqslant L$ hit by time $C_1 \log n$ is $\leqslant C2^{-k}$, and (iii) two particles that are separated by more that L must come to within distance L before they hit.

$$P(H_{12,3}|H_{12}) \leqslant \sum_{k=1}^{L/2} 2^{-k} M^k \pi_{min} + 2^{-L/2}(n - M^{L/2})\pi_{min}$$

$$\leqslant 2^{-L/2} n \pi_{min} \sum_{j=0}^{\infty} (2/M)^j \leqslant \frac{C2^{-(1/10)\log_M n}}{1 - 2/3} \leqslant Cn^{-\delta}$$

which proves the desired result. □

Lemmas 2.2.9 and 2.2.10 contain all the ideas needed to estimate general coalescence patterns. The hardest part of doing this in general is to find appropriate notation to enumerate the possibilities. To do this we will use notation used to describe phylogenetic trees. For example

$$H_{(((12)5)9),(34),((67)8)}$$

means that coalescence reduces the 9 particles to 2. First 1 and 2 coalesce, then they coalesce with 5 and with 9; 3 and 4 coalesce and then do not coalesce with any other particle; 6 and 7 coalesce, and then 8 coalesces with them. In defining these events we ignore collisions between particles that have already coalesced with another one. In the example under consideration we have

$$P(H_{(((12)5)9),(34),((67)8)}) \leqslant C \left(\frac{\log n}{n}\right)^3 \cdot n^{-3\delta} \leqslant Cn^{-9\delta/2}$$

We say a random walk W_i is *isolated* if W_i does not coalesce with other random walks by time $C_1 \log n$. Suppose s is the number of isolated random walks among $W_1, \ldots W_{2m}$ and define

$$G^{NI}_{i_1,\ldots,i_k} = \{\text{None of } W_{i_1}, \ldots, W_{i_k} \text{is isolated}\}$$
$$G^{iso}_{i_1,\ldots,i_k} = \{\text{All } W_{i_1}, \ldots, W_{i_k} \text{ are isolated}\}$$

Since $G^{NI}_{i_1,\ldots,i_k}$ is contained in a union of H events involving k particles.

Lemma 2.2.11. *Given any k coalescing random walks $W_{i_1}, \ldots, W_{i_k}$ starting from the stationary distribution, then there exist constants $C > 0$ and $\delta > 0$ such that the probability of no isolated random walk is bounded by*

$$P\left(G^{NI}_{i_1 \ldots i_k}\right) \leqslant C/n^{\delta k/2} \tag{2.20}$$

2.2.3 Moment estimates

We begin with second moments.

Lemma 2.2.12. *If n is large $E[Z_1 Z_2] \leqslant (C \log n)/n$.*

Proof. On H_{12},

$$Y(W_1(r))Y(W_2(r)) = Y^2(W_1(r)) \leqslant 1$$

so the contribution to the expected value is $\leqslant C(\log n)/n$. Since $W_1'(r)$ and $W_2'(r)$ are independent and have distribution π

$$E[Y(W_1'(r))Y(W_2'(r))] = 0$$

Using our coupling and this we get

$$|E[Y(W_1(r))Y(W_2(r)); H_{12}^c]| = |E[Y(W_1'(r))Y(W_2'(r)); H_{12}^c]|$$
$$= |E[Y(W_1'(r))Y(W_2'(r)); H_{12}]| \leqslant P(H_{12}) \leqslant \frac{C \log n}{n}$$

and the desired result follows. □

Turning now to 4th moments, let S be the number of isolated random walks among W_1, W_2, W_3 and W_4. Then

$$E[Z_1 Z_2 Z_3 Z_4] = \sum_{s=1}^{4} E\left[Z_1 Z_2 Z_3 Z_4; S = s\right] \tag{2.21}$$

Case 1: s=1. First by symmetry, we have

$$E\left[Z_1 Z_2 Z_3 Z_4; S = 1\right] = 4E\left[Z_1 Z_2 Z_3 Z_4; G_{123}^{NI} \cap G_4^{iso}\right]$$

where 4 comes from the choices of the only isolated random walk. Now we couple W_4 to an independent random walk W_4'. Precisely, $(W_1, W_2, W_3, W_4) = (W_1, W_2, W_3, W_4')$

until $W_4 = W_4'$ hits the trajectory of W_i for some $i \leqslant 3$. When this occurs, the first vector becomes (W_1, W_2, W_3, W_i). From this coupling, we know

$$E\left[Z_1Z_2Z_3Z_4; G_{123}^{NI} \cap G_4^{iso}\right] = E\left[Z_1Z_2Z_3Z_4'; G_{123}^{NI} \cap G_4^{iso}\right]$$

Note that Z_4' is independent of Z_1, Z_2, Z_3. Hence $E\left[Z_1Z_2Z_3Z_4'; G_{123}^{NI}\right] = 0$. This implies

$$\begin{aligned}&\left|E\left[Z_1Z_2Z_3Z_4'; G_{123}^{NI} \cap G_4^{iso}\right]\right| = \left|E\left[Z_1Z_2Z_3Z_4'; G_{123}^{NI} \cap G_4^{iso,c}\right]\right| \\ \leqslant& 3\left|E\left[Z_1Z_2Z_3Z_4'; G_{123}^{NI} \cap G_{14}\right]\right| \leqslant 3P(G_{123}^{NI} \cap G_{14}) \leqslant \frac{C}{n^\delta}P(H_{12,3}) \leqslant C\log n/n^{1+2\delta}\end{aligned} \tag{2.22}$$

Case 2: s=2. Similarly, symmetry tells us

$$E\left[Z_1Z_2Z_3Z_4; S=2\right] = 6E\left[Z_1Z_2Z_3Z_4; G_{12}^{NI} \cap G_{34}^{iso}\right] = 6E\left[Z_1Z_2Z_3Z_4; G_{12} \cap G_{34}^{iso}\right]$$

Now couple (W_3, W_4) to independent random walks (W_3', W_4'). Precisely, (W_1, W_2, W_3, W_4) and (W_1, W_2, W_3', W_4') are the same until W_3 or W_4 hits others. Then

$$E\left[Z_1Z_2Z_3Z_4; G_{12} \cap G_{34}^{iso}\right] = E\left[Z_1Z_2Z_3'Z_4'; G_{12} \cap G_{34}^{iso}\right]$$

Note that Z_3' and Z_4' are independent of W_1 and W_2. Hence $E\left[Z_1Z_2Z_3'Z_4'; G_{12}\right] = 0$. This implies

$$\begin{aligned}&\left|E[Z_1Z_2Z_3'Z_4'; G_{12} \cap G_{34}^{iso}]\right| = \left|E[Z_1Z_2Z_3'Z_4'; G_{12} \cap G_{34}^{iso,c}]\right| \\ \leqslant& C\left(P(H_{12,3}) + P(H_{12,34})\right) \leqslant C\log n/n^{1+\delta}\end{aligned}$$

The first inequality holds because on G_{12} if W_3' is not isolated, it can either hit W_1 or W_2, which has probability bounded by $CP(H_{12,3})$ where the constant C takes care of the order of coalescent; or it can hit W_4' without hitting W_1 or W_2, which has probability bounded by $P(H_{12,34})$. The same argument applies to W_4'. Therefore, we have obtained

$$E\left[Z_1Z_2Z_3Z_4; S=2\right] \leqslant C\log n/n^{1+\delta} \tag{2.23}$$

Case 3: s=3. Impossible

Case 4: s=4. Couple W_i, $1 \leqslant i \leqslant 4$ to independent random walks W'_i, $1 \leqslant i \leqslant 4$. Precisely, they agree until there is a coalescence. Then

$$E\left[Z_1Z_2Z_3Z_4; S=4\right] = E\left[Z_1Z_2Z_3Z_4; G^{iso}_{1234}\right]$$
$$= E\left[Z'_1Z'_2Z'_3Z'_4; G^{iso}_{1234}\right] = -E\left[Z'_1Z'_2Z'_3Z'_4; G^{iso,c}_{1234}\right]$$

Now $G^{iso,c}_{1234}$ is a disjoint union of events

$$G^{NI}_{A^c} \cap G^{iso}_{A}$$

where $A \subsetneqq \{1,2,3,4\}$ and $A^c = \{1,2,3,4\}\backslash A$. Combining this with (2.21), (2.22), and (2.23) gives Lemma 2.2.3 for $m = 2$.

General m. To compute $E[Z_1Z_2\ldots Z_{2m}]$, let S denote the number of isolated random walks among $W_1, \ldots, W_{2m}$ at time $C_1 \log n$. Then

$$E[Z_1Z_2\ldots Z_{2m}] = \sum_{s=0}^{2m} E[Z_1Z_2\ldots Z_{2m}; S=s] \tag{2.24}$$

For any $1 \leqslant s \leqslant 2m$, first symmetry gives us

$$E[Z_1Z_2\ldots Z_{2m}; S=s] = \binom{2m}{s} E[Z_1Z_2\ldots Z_{2m}; G^{NI}_{1\ldots 2m-s} \cap G^{iso}_{2m-s+1\ldots 2m}] \tag{2.25}$$

Hence we just need to focus on the case where the last s random walks are isolated while no isolated random walk appears in the first $2m-s$ random walks. Now couple W_j with $2m-s+1 \leqslant j \leqslant 2m$ to independent random walks W'_j with $2m-s+1 \leqslant j \leqslant 2m$. Precisely, $(W_1, \ldots, W_{2m})$ and $(W_1, \ldots, W_{2m-s}, W'_{2m-s+1}, \ldots, W'_{2m})$ are identical until any W_i with $2m-s < i \leqslant 2m$ hits another particle. Then

$$I := E\left[Z_1\ldots Z_{2m}; G^{NI}_{1\ldots 2m-s} \cap G^{iso}_{2m-s+1\ldots 2m}\right]$$
$$= E\left[Z_1\ldots Z_{2m-s}Z'_{2m-s+1}\ldots Z'_{2m}; G^{NI}_{1\ldots 2m-s} \cap G^{iso}_{2m-s+1\ldots 2m}\right] \tag{2.26}$$

As before, note that $Z'_{2m-s+1}, \ldots, Z'_{2m}$ are all independent of $Z_1, \ldots, Z_{2m-s}$. This implies that $E\left[Z_1 \ldots Z_{2m-s} Z'_{2m-s+1} \ldots Z'_{2m}; G^{NI}_{1\ldots 2m-s}\right] = 0$ and that

$$E[Z_1 \ldots Z'_{2m}; G^{NI}_{1\ldots 2m-s} \cap G^{iso}_{2m-s+1\ldots 2m}] = -E[Z_1 \ldots Z'_{2m}; G^{NI}_{1\ldots 2m-s} \cap G^{iso,c}_{2m-s+1\ldots 2m}] \tag{2.27}$$

Note that $G^{NI}_{1\ldots 2m-s} \cap G^{iso,c}_{2m-s+1\ldots 2m}$ is a union of disjoint sets of the form

$$G^{NI}_{A^c} \cap G^{iso}_{A}$$

where $A \subsetneqq \{2m - s + 1, \ldots, 2m\}$ and $A^c = \{1, 2, \ldots, 2m\} \backslash A$. Combining this with (2.26) and (2.27) gives

$$I \leqslant \sum_{A \subsetneqq \{2m-s+1,\ldots,2m\}} \left|E[Z_1 \ldots Z'_{2m}; G^{NI}_{A^c} \cap G^{iso}_{A}]\right|$$

Since $|A| \leqslant s-1$, we have reduced the number of isolated random walks by at least 1. Now apply similar argument to each $G^{NI}_{A^c} \cap G^{iso}_{A}$. Note that $Z'_i,\ i \in A$ are independent of the $Z_i\ i \in A^c$. Hence $E[Z_1 \ldots Z'_{2m}; G^{NI}_{A^c}] = 0$. This implies that

$$E[Z_1 \ldots Z'_{2m}; G^{NI}_{A^c} \cap G^{iso}_{A}] = -E[Z_1 \ldots Z'_{2m}; G^{NI}_{A^c} \cap G^{iso,c}_{A}]$$

We can further subdivide each $G^{NI}_{A^c} \cap G^{iso,c}_{A}$ into disjoint sets of the form $G^{NI}_{B^c} \cap G^{iso,c}_{B}$ where $B \subsetneqq A \subsetneqq \{1, \ldots, 2m\}$ and thus $|B| \leqslant s - 2$. This leads to

$$I \leqslant C \sum_{B \subsetneqq \{2m-s+1,\ldots,2m\},\ |B| \leqslant s-2} \left|E[Z_1 \ldots Z'_{2m}; G^{NI}_{B^c} \cap G^{iso}_{B}]\right|$$

where the constant C takes care of the possible repetition of $G^{NI}_{B^c} \cap G^{iso}_{B}$ from those subdivisions. We can keep doing this and decrease the number of isolated random walks by at least 1 at each step until we have no isolated random walk, i.e. when all those A (or B)given above have $|A| = 0$ (or $|B| = 0$). We will eventually have

$$I \leqslant C \left|E[Z_1 \ldots Z'_{2m}; G^{NI}_{1,2,\ldots,2m}]\right|$$

Since $\left|Z_1 \ldots Z_{2m-s} Z'_{2m-s+1} \ldots Z'_{2m}\right| \leqslant 1$, then by Lemma 3.1.2 we have

$$I \leqslant CP\left(G^{NI}_{12\ldots 2m}\right) \leqslant C/n^{\delta m} \tag{2.28}$$

and the proof of Lemma 2.2.3 is complete.

2.2.4 Bounding the drift

The drift

$$\beta(\xi_{\lambda t}) = \sum_{x \in G_n} \pi(x) \sum_{y \sim x} \sum_{z \sim x, z \neq y} [1_{\{\xi_{\lambda t}(x)=2,\ \xi_{\lambda t}(y)=1 \text{ or } \xi_{\lambda t}(z)=1\}}$$
$$- 1_{\{\xi_{\lambda t}(x)=1,\ \xi_{\lambda t}(y)=2 \text{ or } \xi_{\lambda t}=2\}}]$$

We want to show

Lemma 2.2.13. *There is a $\gamma > 0$ so that for any m there is a constant C_m so that*

$$P(|\beta(\xi_{\lambda t}) - b(X_t)| \geqslant \epsilon | \mathcal{F}_{t-(C_1 \log n)/\lambda}) \leqslant \frac{C_m}{\varepsilon^{2m} n^{2m\gamma}}. \tag{2.29}$$

Proof. We prove the result for the edge case and leave the straightforward extension to the site case to the reader. If we let $\mathbb{1}(x|y|z)$ is the indicator function of the event that the dual random walks starting from x, y, and z at time t do not hit by time $t - C_1(\log n)/\lambda$ and $p(x|y|z) = E\mathbb{1}(x|y|z)$ then

$$E[\beta(\xi_t)|\mathcal{F}_{t-C_1(\log n)/\lambda}] = \frac{1}{n} \sum_{x \in G_n} \sum_{y \sim x} \sum_{z \sim x, z \neq y} \mathbb{1}(x|y|z) \tilde{u}(1-\tilde{u})(1-2\tilde{u}) \tag{2.30}$$

$$b(X_t) = \frac{1}{n} \sum_{x \in G_n} \sum_{y \sim x} \sum_{z \sim x, z \neq y} p(x|y|z) \tilde{u}(1-\tilde{u})(1-2\tilde{u}) \tag{2.31}$$

The random variables $\mathbb{1}(x|y|z)$ are dependent if the triples (x, y, x) and (x', y', z') overlap or if the associated random walks coalesce. To simplify things we will let $\hat{\mathbb{1}}(x|y|z)$ be the event none of the walks r-coalesce, i.e., the pair collides before either of them exits $B(x,r)$, where $r = \varepsilon \log n$ and ε will be chosen later to be small enough. Let

$$Y_{x,y,z} = \hat{\mathbb{1}}(x|y|z) - p(x|y|z) \quad \text{and} \quad \hat{Y}_{x,y,z} = \hat{\mathbb{1}}(x|y|z) - \hat{p}(x|y|z)$$

where $\hat{p}(x|y|z) = E(\hat{\mathbb{1}}(x|y|z))$. As before we will compute $2m$th moments of the sum over $x \in G_n$ and neighbors y, $z \neq y$ of x. The calculation is simpler here than in Sections 3.2–3.3, since we are only concerned whether the particles coalesce and not how they are spread over the graph at time $C_1 \log n$.

Lemma 2.2.14. *There is a $\beta > 0$ so that for any m we have*

$$E\left(\sum_{x,y,z} \hat{Y}_{x,y,z}\right)^{2m} \leqslant C_m n^{1+\beta}.$$

Proof. The sum has $K = \sum_x d(x)(d(x)-1)$ terms. The $2m$th moment of the sum has terms of the form.

$$\hat{Y}_{x_1,y_1,z_1} \cdots \hat{Y}_{x_{2m},y_{2m},z_{2m}}$$

If some x_i has distance $3r$ from all of the other x_j then Y_{x_i,y_i,z_i} is independent of the product of the rest of the random variables and the expected value is 0.

Suppose now that for each x_i there is at least one x_j that is within distance $3r$. Create a graph D (for dependency) where there is an edge between i and j if $d(x_i,x_j) < 3r$. Let κ be the number of components in the graph. The number of points within distance $3r = 3\varepsilon \log n$ of a given x is $\leqslant M^{3\varepsilon \log n} \equiv n^\beta$. If the dependency graph has κ components the number of terms $\leqslant A_D n^\kappa n^{\beta(2m-\kappa)}$. When D has no singletons $\kappa \leqslant m$. Since $E|Y_{x_1,y_1,z_1} \cdots Y_{x_{2m},y_{2m},z_{2m}}| \leqslant 1$ the desired result follows. □

To bound the sum of the $Y_{x,y,z}$ we will write

$$Y_{x,y,z} = \hat{Y}_{x,y,z} + (\hat{p}(x|y|z) - p(x|y|z)) + (\mathbb{1}(x|y|z) - \hat{\mathbb{1}}(x|y|z))$$

To control the middle term note that Lemma 2.1.4 implies that with high probability the two walks that are separated by $r = \varepsilon \log n$ will not hit before they are separated by $5r$ is $\leqslant 2^{1-\varepsilon \log n} \equiv 2n^{-\alpha}$. Using this result repeatedly we see the probability they do not hit by $C_1 \log n$ is $\leqslant Cn^{-\alpha}$. Thus

$$\sum_{x,y,z} |p(x|y|z) - \hat{p}(x|y|z)| \leqslant cn^{1-\alpha}$$

To control the $2m$th moment of the sum of the third term, suppose we are given $1 \leqslant K \leqslant 2m$ distinct (x_i, y_i, z_i) where y_i and z_i are different neighbors of x_i. Note that $Z = \prod_{i=1}^{K} \left(\mathbb{1}(x_i|y_i|z_i) - \hat{\mathbb{1}}(x_i|y_i|z_i)\right) > 0$ if and only if $\mathbb{1}(x_i|y_i|z_i) - \hat{\mathbb{1}}(x_i|y_i|z_i) > 0$ for all $1 \leqslant i \leqslant K$. That is, for any i, there exist a pair among $(W^{x_i}, W^{y_i}, W^{z_i})$ such that they do not $r-$coalesce but coalesce after exiting $B(x, r)$. As before, we only focus on coalescent in a specific pattern and consider the event H as the following:

- Suppose $T_0 = 0$. Some particles from $(W^{x_1}, W^{y_1}, W^{z_1})$ coalesce at time T_1 but do not $r-$coalesce. We ignore collisions involving other particles and let them do independent random walks before that time.

- $(W^{x_1}, W^{y_1}, W^{z_1})$ are ignored immediately after time T_1 except for the ones who appear in $(W^{x_2}, W^{y_2}, W^{z_2})$. At time $T_2 > T_1$, some particles from $(W^{x_2}, W^{y_2}, W^{z_2})$ coalesce after exiting $B(x_2, r)$. We ignore all other particles on $[T_1, T_2]$.

- In general, at time $T_k > T_{k-1}$, some particles from $(W^{x_k}, W^{y_k}, W^{z_k})$ coalesce after exiting $B(x_k, r)$. We ignore all other particles on $[T_{k-1}, T_k]$.

This again enlarges the probability of our interest up to a constant factor from permutation. Moreover by Lemma 2.1.6, each step has probability $\leqslant C/n^{-\alpha}$ to occur

. Hence

$$P(Z > 0) \leqslant CP(H)$$

$$\leqslant C \prod_{k=1}^{2m} P\left(T_k < \infty | T_{k-1}\infty\right) \leqslant C/n^{-K\alpha}$$

Note that in the expansion of $\left(\sum_{x,y,z} \mathbb{1}(x|y|z) - \hat{\mathbb{1}}(x|y|z)\right)^{2m}$, the number of terms consisting of such K distinct (x_i, y_i, z_i) is $\leqslant n^K M(M-1)$.

This implies that

$$E\left(\sum_{x,y,z} \mathbb{1}(x|y|z) - \hat{\mathbb{1}}(x|y|z)\right)^{2m} \leqslant \sum_{K=1}^{2m} n^K M(M-1) \times C/n^{-K\alpha} \leqslant C_m n^{2m(1-\alpha)}.$$

Combining our results and using Chebyshev's inequality, the proof of Lemma 2.2.13 is complete.

2.2.5 *Final details*

To extend Lemma 2.2.13 to bound the probability of

$$\Omega_1^c = \left\{ \int_0^t |\beta(X_s) - b(X_s)|\, ds \geqslant \eta \right\}$$

we subdivide the interval $[0,t]$ into subintervals of length $1/\lambda n^{1/2}$. Within each interval the probability that more than $2n^{1/2}$ sites will flip is $\leqslant \exp(-c\sqrt{n})$. From this it follows that if $2t\varepsilon \leqslant \eta$ then

$$P(\Omega_1^c) \leqslant t\lambda n^{1/2} \left[\frac{C_{m,\varepsilon}}{n^{m/2}} + \exp(-c\sqrt{n}) \right] \tag{2.32}$$

The last bound only works for fixed t. To get long time survival we will iterate. Let

$$T_0 = \inf\{t : |x_t - 1/2| < \varepsilon\}$$

and note that x_t is the solution of the ODE so this is not random. Theorem 2.0.4 implies that at this time $|X_t - 1/2| \leqslant 2\varepsilon$ with very high probability, i.e., with an error of less that $Cn^{-(m-1)/2}$. Let

$$T_1 = \inf\{t > T_0 : |X_t - 1/2| \geqslant 4\varepsilon\}$$

and note that on $[T_0, T_1]$ we have $|X_t - 1/2| \leqslant 4\varepsilon$. There is a constant t_0 so that if $x(0) = 1/2 + 4\varepsilon$ or $x(0) = 1/2 - 4\varepsilon$ then $|x(t_0) - 1/2| \leqslant \varepsilon$. Let $S_1 = T_1 + t_0$. Theorem 2.0.4 implies that with high probability $|X(S_1) - 1/2| \leqslant 2\varepsilon$ and $|X_t - 1/2| \leqslant 5\varepsilon$ on $[T_1, S_1]$. For $k \geqslant 2$ let

$$T_k = \inf\{t > S_{k-1} : |X_t - 1/2| \geqslant 4\varepsilon\} \quad \text{and} \quad S_k = T_k + t_0.$$

We can with high probability iterate the construction $n^{(m-2)/2}$ times before it fails. Since each cycle takes at least t_0 units of time, the proof of Theorem 2.0.3 is complete.

3

The Zealot Voter Model

Since we have described the model in the introduction. In this Chapter, we will focus on the results and in particular the proofs.

3.1 Proof of Theorem 1.2.1

Recall that our first result works for all bounded trees.

Theorem 3.1.1 (Theorem 1.2.1). *On any tree with degrees* $3 \leqslant d(x) \leqslant M$*, the zealot voter model survives if*

$$\sum_{k \geqslant 2} (k-1)p_k - p_0 > 0.$$

There are four steps in the proof.

- We begin by deriving a differential equation for the expected number of occupied sites.
- We define the frontier and the external boundary of a set of occupied sites and prove lower bounds on their sizes.

- Combining the first two steps we obtain differential equations that lower bound the number of occu[pied sites and the size of the frontier.

- We prove Theorem 1.2.1 by showing that the set of occupied sites dominates a supercritical branching walk.

3.1.1 Step 1: Derivation of the ODE

Let $d_k(x) = (d(x)-1)\cdots(d(x)-(k-1))$. Note that $d_k(x)$ is the number of ways of picking $k-1$ things out of $d(x)-1$ when the order of the choices is important. Using $x*(k-1) \neq y_k$ to indicate that we sum over all ordered choices of $k-1$ different neighbors $y_1, ..., y_{k-1}$ of x that are not $\neq y_k$.

$$\frac{d}{dt}\sum_x P(\xi_t(x)=1) = -p_0\sum_x P(\xi_t(x)=1)$$

$$+\sum_{k>1}\sum_{x,y_k\sim x}\frac{p_k}{d_k(x)}\sum_{x*(k-1)\neq y_k}\left[P\left(\xi_t(x)=1,\xi_t(y_k)=0\right)-P\left(\xi_t(x)=1,\text{all }\xi_t(y_i)=0\right)\right] \tag{3.1}$$

$$+\sum_{k>1}\sum_{x,y_k\sim x}\frac{p_k}{d_k(x)}\sum_{x*(k-1)\neq y_k}P(\xi_t(x)=\xi_t(y_k)=0,\xi_t(y_i)=1\text{ for some }i<k)$$

Note that the second and third terms are $\geqslant 0$.

Proof. Breaking things down according to the value of k, treating births and deaths separately, and noting that in the last four terms jumps occur at rate $d(x)$

$$\begin{aligned}
\frac{d}{dt}P(\xi_t(x)=1) &= -p_0 P(\xi_t(x)=1) \\
&- p_1 \sum_{y\sim x} P(\xi_t(x)=1, \xi_t(y)=0) \\
&+ p_1 \sum_{y\sim x} P(\xi_t(x)=0, \xi_t(y)=1) \qquad (3.2) \\
&- \sum_{k>1}\sum_x \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x)=1, \xi_t(y_i)=0 \text{ for all } 1\leqslant i\leqslant k\right) \\
&+ \sum_{k>1}\sum_x \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x)=0, \xi_t(y_i)=1 \text{ for some } 1\leqslant i\leqslant k\right)
\end{aligned}$$

If we sum over x then the second and third terms cancel. We now fix k and split the last term into two

$$\begin{aligned}
&= -\sum_x \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x)=1,\ \text{all } \xi_t(y_i)=0\right) \\
&+ \sum_x \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x)=0, \xi_t(y_k)=1\right) \qquad (3.3) \\
&+ \sum_x \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x)=\xi_t(y_k)=0, \xi_t(y_i)=1 \text{ for some } 1\leqslant i<k\right)
\end{aligned}$$

Recalling the definition of $d_k(x)$, the first sum can be written as

$$\begin{aligned}
&\sum_{x,y_k\sim x} \frac{p_k}{d_k(x)} \sum_{x*(k-1)\neq y_k} P\left(\xi_t(x)=0, \xi_t(y_k)=1\right) \\
&= \sum_{x,y_k\sim x} p_k P\left(\xi_t(x)=0, \xi_t(y_k)=1\right) \\
&= \sum_{y_k,x\sim y_k} p_k P\left(\xi_t(x)=0, \xi_t(y_k)=1\right) \\
&= \sum_{y_k,x\sim y_k} \frac{p_k}{d_k(y_k)} \sum_{y_k*(k-1)\neq x} P\left(\xi_t(x)=0, \xi_t(y_k)=1\right)
\end{aligned}$$

Interchanging the role of x and y_k, the above

$$= \sum_{x, y_k \sim x} \frac{p_k}{d_k(x)} \sum_{x*(k-1) \neq y_k} P\left(\xi_t(x) = 1,\ \xi_t(y_k) = 0\right)$$

Then (3.3) can be reformulated as

$$= -\sum_{x} \frac{p_k}{d_k(x)} \sum_{x*k} P\left(\xi_t(x) = 1, \xi_t(y_i) = 0 \text{ for all } 1 \leqslant i \leqslant k\right)$$

$$+ \sum_{x, y_k \sim x} \frac{p_k}{d_k(x)} \sum_{x*(k-1) \neq y_k} P\left(\xi_t(x) = 1,\ \xi_t(y_k) = 0\right)$$

$$+ \sum_{x, y_k \sim x} \frac{p_k}{d_k(x)} \sum_{x*(k-1) \neq y_k} P(\xi_t(x) = \xi_t(y_k) = 0, \xi_t(y_i) = 1 \text{ for some } \ i < k)$$

Combining the first two summations and summing over $k > 1$ gives the desired result. □

3.1.2 Step 2: Frontier lower bounds

Pick a vertex from the tree to be the root and call it x_0. Given a vertex x in the tree we say that x' is a child of x if it is a neighbor of x and further away from the root than x is. We define the **subtree generated by** x', $S(x')$, to be all of the vertices that can be reached from x' without going through x. By definition, $x' \in S(x')$. For any finite set on the tree A, define its **frontier** $F(A)$ as the set of sites $x \in A$ that have a child x' such that the subtree $S(x') \cap A = \varnothing$ and define the **exterior boundary of A**, $H(A)$ to be the set of all such children x'. That is, $x' \in H(A)$ if and only if $S(x') \cap A = \varnothing$ and the parent of x' is in $F(A)$. to help visualize the definitions, see Figure 3.1.2. Our next step is to lower bound the sizes of the sets we just defined.

Lemma 3.1.2. *$|H(A)| \geqslant |A|$ and $|F(A)| \geqslant |A|/(M-1)$.*

Proof. We prove the first result by induction on the cardinality of $|A|$. If $|A| = 1$, the result is trivial as $|H(A)| \geqslant d(x) - 1 \geqslant 2$. Suppose now that the result is true for all B with $|B| \leqslant n-1$ and let $|A| = n$. Let $x \in A$ be the point with the largest distance to the root and let $B = A \backslash \{x\}$. Then by induction $|H(B)| \geqslant n-1$. Since none of the descendents of x are in A, but x might be in $H(B)$.

$$|H(A)| \geqslant |H(B)| - 1 + d(x) - 1 \geqslant (n-1) - 1 + 2 = n$$

The second result follows from the first since $|H(A)| \leqslant (M-1)|F(A)|$. □

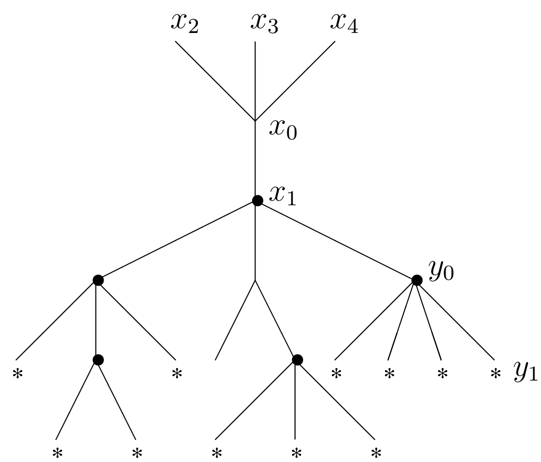


FIGURE 3.1: For simplicity we have only drawn the edges from vertices within distance 3 of the root that are relevant to the definitions. $\bullet$ indicates sites in A. All the $\bullet$s are in $F(A)$ except for x_1. $*$s mark the points in $H(A)$.

3.1.3 ODE lower bounds

Let $A_t = \{x : \xi_t(x) = 1\}$ Our next step is

Lemma 3.1.3. *Let* $\gamma = -p_0 + \sum_k p_k(k-1)$ *(which is* > 0 *by assumption).*

$$\frac{d}{dt} E|A_t| \geqslant \gamma E|A_t|$$

Proof. Let l_x be the distance of x from the root. The expression on the second line in (3.1) is $\geqslant 0$. The third line is

$$= \sum_{x,y_k \sim x} \frac{p_k}{d_k(x)} \sum_{x*(k-1) \neq y_k} P(\xi_t(x) = \xi_t(y_k) = 0,\ \text{some}\ \xi_t(y_i) = 1)$$

$$\geqslant E \sum_{x \in H(A_t), l_{y_k} > l_x} \frac{p_k}{d_k(x)} \sum_{y_i \in F(A_t)\ for\ some\ 1 \leqslant i \leqslant k-1} 1$$

$$= E \sum_{x \in H(A_t), l_{y_k} > l_x} \frac{p_k}{d_k(x)} \cdot (d(x) - 1) \cdot (k-1) \cdot \binom{d(x)-2}{k-1}$$

$$= (k-1) p_k E|H(A_t)| \geqslant (k-1) p_k |A_t|$$

In the third line, $d(x) - 1$ gives the choices for y_k. $k - 1$ is because we have $k - 1$ choices from $y_1, ..., y_{k-1}$ to be on the frontier. Suppose y_1 is chosen to be in the frontier, then the number of choices for $y_2, \ldots y_{k-1}$ is $\binom{d(x)-2}{k-1}$. The final inequality comes from Lemma 3.1.2 □

Choose a neighbor x_1 of the root x_0. (See Figure 3.1.2 for a picture.) Set all the sites outside of $S(x_1) \cup \{x_0\}$ to be always equal to 0. Let $\bar{\xi}_t$ be the process restricted to $S_1 \equiv S(x_1) \cup \{x_0\}$. Let

$$\bar{A}_t = \{x : \bar{\xi}_t(x) = 1\} \qquad A_t^* = \bar{A}_t \cap S(x_1)$$

$$H^*(\bar{A}_t) = H(\bar{A}_t) \cap S(x_1) \qquad F^*(\bar{A}_t) = F(\bar{A}_t) \cap S(x_1)$$

Lemma 3.1.4.

$$\frac{d}{dt} E|\bar{A}_t| \geqslant \gamma E|\bar{A}_t| - (\gamma + 1)(M - 1)$$

Proof. We repeat the proof of Lemma 3.1.3. The differential equation in (3.2) remains valid but when we make the transition to (3.3) there is a term with $k = 1$ that does not cancel:

$$-p_1[d(x_0) - 1] P(\bar{\xi}_t(x_0) = 1)$$

Note that if $x_0 \in \bar{A}_t$, then $x_2, \ldots, x_{d(x_0)} \in H(\bar{A}_t)$ so

$$|H^*(\bar{A}_t)| \geqslant |H(\bar{A}_t)| - [d(x_0) - 1] \geqslant |\bar{A}_t| - (d(x_0) - 1)$$

where the last inequality follows from Lemma 3.1.2. Using $d(x_0) \leqslant M$ the desired result follows. □

Let $L = 2(\gamma + 1)(M - 1)/\gamma$. Lemma 3.1.4 implies that once $E|\bar{A}_t| \geqslant L$ it grows exponentially with rate $\geqslant \alpha = (\gamma + 1)(M - 1)$.

Lemma 3.1.5. *There exists $\varepsilon_0 > 0$ such that*

$$P(|\bar{A}_1| \geqslant L) \geqslant \varepsilon_0 \tag{3.4}$$

for all trees $\mathcal{T}$ with $3 \leqslant d_{min} \leqslant d(x) \leqslant M$

Proof. Let G_L be the event that at time 1 there is an occupied path from the root x_0 to distance $L - 1$. It is easy to see that there exists $\varepsilon_0 > 0$ such that $P(G_L) \geqslant \varepsilon_0$ for all trees $\mathcal{T}$ with $3 \leqslant d_{min} \leqslant d(x) \leqslant M$. To see this note that the worst case occurs when all sties have degree 3 but offspring are sent across an edge with probability $1/M$. □

Lemma 3.1.6. *There exists $t_0 > 0$ such that*

$$E|F^*(\bar{A}_{t_0})| \geqslant 2 \tag{3.5}$$

for all trees with $3 \leqslant d_{min} \leqslant d(x) \leqslant M$

Proof. Conditioning on $\{|\bar{A}_t| \geqslant L\}$ it follows that for all trees with $3 \leqslant d_{min} \leqslant d(x) \leqslant M$.

$$E|\bar{A}_t| \geqslant \varepsilon_0 e^{\alpha(t-1)]}$$

Now $|F^*(\bar{A}_t)| \geqslant |F(\bar{A}_t)| - 1$ with equality if x_0 is in state 1, so by Lemma 3.1.2

$$|F^*(\bar{A}_t)| \geqslant |F(\bar{A}_t)| - 1 \geqslant \frac{1}{M-1}|\bar{A}_t| - 1$$

and the desired result follows. □

3.1.4 Step 4: Lower bounding BRW

Now define a lower bounding branching random walk Z_n. Let $Z_0 = \{x_0\}$, where x_0 is the root. Let $Z_1 = F^*(\bar{A}_{t_0}) = F(\bar{A}_{t_0}) \cap S(x_1)$. Inductively, given Z_n, note that for any $x \in Z_n$, x has a child x' such that $S(x') \cap \bar{A}_{nt_0} = \varnothing$. Let $F^*(\bar{A}^{x,0}_{t_0})$ be the children of x, where the superscript 0 means that to obtain $\bar{A}^{x,0}_t$, we enforce 0-boundary condition on sites above x. Hence all the neighbors of x except for x' are in state 0 during $[nt_0,\ (n+1)t_0]$. Therefore all $\bar{A}^{x,0}_t$ $\forall x \in Z_n$ are independent and $E|F^*(\bar{A}^{x,0}_{t_0})| > 1$ $\forall x \in Z_n$ by Lemma 3.1.5. Define the $n+1$ th generation by

$$Z_{n+1} = \bigcup_{x \in Z_n} F^*(\bar{A}^{x,0}_{t_0})$$

Lemma 3.1.7. *There exists $C_v > 0$ such that for all trees $\mathcal{T}$ with $3 \leqslant d(x) \leqslant M$ for all x*

$$E[Z_{n+1}|Z_n] \geqslant 2Z_n \tag{3.6}$$

$$Var(Z_{n+1}|Z_n) \leqslant C_v Z_n \tag{3.7}$$

Proof. Given any tree $\mathcal{T}$, note that $Z_{n+1} = \bigcup_{x \in Z_n} F^*(\bar{A}^x_{t_0})$ and $F^*(\bar{A}^x_{t_0}) \cap F^*(\bar{A}^y_{t_0}) = \varnothing$ if $x \neq y$. Then

$$E[Z_{n+1}|Z_n, \mathcal{T}] = \sum_{x \in Z_n} E\left[|F^*(\bar{A}^x_{t_0})||Z_n,\ T\right] > (1+\epsilon)Z_n \tag{3.8}$$

To prove (3.7), let η_t be a branching process where $\eta_0 = 1$ and every particle gives a birth at rate M without death. Then given any tree, $|\bar{A}_t|$ is stochastically bounded by $|\eta_t|$. So now, let $\mathcal{T}^x$ denote the subtree consisting of x and its descendents. By independence

$$\text{var}\ (Z_{n+1}|Z_n, \mathcal{T}) = \sum_{x \in Z_n} \text{var}\ \left(|\bar{A}^{x,0}_{t_0}||\mathcal{T}^x\right)$$

$$\leqslant \sum_{x \in Z_n} E\left[|\bar{A}^{x,0}_{t_0}|^2\ |\mathcal{T}^x\right] \leqslant \sum_{x \in Z_n} E|\eta_{t_0}|^2 = C_v Z_n$$

Since T is arbitrary, we have completed the proof. $\square$

The next Lemma completes the proof of Theorem 1.2.1.

Lemma 3.1.8. *With positive probability*

$$\liminf_{n\to\infty} \frac{Z_n}{(3/2)^n} \geqslant 1. \tag{3.9}$$

Proof. First by Lemma 3.1.7 and Chebyshev's Inequality,

$$\begin{aligned}
&P\left(Z_{n+1} < (3/2)^{n+1} | Z_n \geqslant (3/2)^n\right) \\
&\leqslant P\left(|Z_{n+1} - E[Z_{n+1}|Z_n]| > Z_n/2 | Z_n \geqslant (3/2)^n\right) \\
&\leqslant E\left(\frac{C_v Z_n}{(Z_n/2)^2} \middle| Z_n \geqslant (3/2)^n\right) \leqslant 4C_v \cdot (2/3)^n \equiv \delta_n
\end{aligned}$$

Pick n_0 large enough so that $\delta_{n_0} < 1$. It follows from the proof of Lemma 3.1.5 that $P(Z_{n_0} \geqslant (3/2)^{n_0}) > 0$. Since δ_n is decreasing, we have

$$P\left(\liminf_{n\to\infty} \frac{Z_n}{(3/2)^n} \geqslant 1 \middle| Z_{n_0} \geqslant (3/2)^{n_0}\right) \geqslant \prod_{n=n_0}^{\infty} (1 - \delta_n) > 0$$

which proves the desired result. $\square$

3.2 Results for d-regular trees

3.2.1 Extinction

The first result is elementary but a proof is cincluded for completeness.

Lemma 3.2.1. *Let $h(x)$ be the probability two continuous time random walks separated by x on a d-regular tree will hit.*

$$h(x) = \left(\frac{1}{d-1}\right)^x$$

Proof. If the two particles are at distance $x > 0$ then the probability they are at distance $x+1$ after the first jump is $(d-1)/d$, while they are at distance $x-1$ with probability $1/d$.

$$\begin{aligned}\left(\frac{1}{d-1}\right)^x &= \frac{d-1}{d}\left(\frac{1}{d-1}\right)^{x+1} + \frac{1}{d}\left(\frac{1}{d-1}\right)^{x-1} \\ &= \frac{1}{d}\left(\frac{1}{d-1}\right)^x + \frac{d-1}{d}\left(\frac{1}{d-1}\right)^x = \left(\frac{1}{d-1}\right)^x\end{aligned}$$

i.e., if X_t is the distance between two coalescing random walks on a d-regular tree then $((d-1)^{-X_t}$ is a martingale. Since $h(0) = 1$, $h(x) \leqslant 1$ for $x \geqslant 0$ and $h(x) \to 0$ as $x \to \infty$ the desired result follows from the optional stopping theorem. □

Let β be the probability two newborn particles in the dual do not coalesce. Since two newborn particles are at distance two from each other

$$\beta = 1 - \frac{1}{(d-1)^2}.$$

Theorem 1.2.2 *On a d-regular tree the COBRA dies out if*

$$d\beta \sum_{k \geqslant 2} (k-1)p_k - p_0 < 0$$

Proof. In COBRA, a particle dies at rate p_0 and gives birth to k particles at rate dp_k. To make the dual process more like a branching random walk, when a particle dies and gives birth to a positive number of particles, we declare that the particle did not die but jumped to the location of particle 1. If no offspring were produced then the particle dies. To get an upper bound on the growth of the dual (i) we ignore coalescence between the lineages that are not siblings, and (ii) if k particles are born we ignore coalescence between particles $i > 1$ and $j > 1$. Note that particles $2, \dots k$ each have probability $\geqslant \beta$ of not coalescing with 1. Thus the expected number of the particles that do not coalesce with 1 is $(k-1)\beta$. If we use η_t^0 to denote the resulting system starting from a single particle then

$$\frac{d}{dt} E\eta_t^0 = \left[-p_0 + d \sum_k p_k (k-1)\beta \right] E\eta_t^0$$

It is immediate that if $d\beta \sum_{k\geqslant 2}(k-1)p_k - p_0 < 0$ then $E|\zeta_t^0| \leqslant E|\eta_t^0| \to 0$. □

3.2.2 Local Survival

Recall that $\mu = \sum_k kp_k$ is the mean number of offspring in the dual.

Theorem 1.2.4 *Given a d-regular tree T, the zealot voter model dies out locally if*

$$\mu < \frac{d(1-p_0)+p_0}{2\sqrt{d-1}}. \tag{3.10}$$

If $p_0 = 0$ *this is* $\mu < d/(2\sqrt{d-1})$.

Proof. Using a superscript 0 to denote the process starting with only the root occupied, we need to show

$$P(\xi_t^0 \cap \{0\} \neq \varnothing) \longrightarrow 0 \quad \text{as } t \to \infty. \tag{3.11}$$

By duality,

$$P(\xi_t^0 \cap \{0\} \neq \varnothing) = P(\zeta_t^0 \cap \{0\} \neq \varnothing) \tag{3.12}$$

Let $\eta_t^0 \supset \zeta_t^0$ be the BRW in which particles die at rate p_0 and at rate dp_k die and give birth onto k neighbors chosen without replacement.

Lemma 3.2.2. *Let $m(t,x) = E\eta_t^0(x)$ be the expected number of particles on site x at time t. Then $m(t,x)$ satisfies the equation*

$$\frac{d}{dt}m(t,x) = -\alpha m(t,x) + \sum_{y\sim x} m(t,y)\mu \quad \text{where } \alpha = d(1-p_0)+p_0$$

The solution is given by

$$m(t,x) = e^{(\mu-\alpha)dt}P(S_t^0 = x) \tag{3.13}$$

where S_t^0 is the random walk on tree T starting from the root that jumps at rate $d\mu$ to a neighbor chosen uniformly at random.

Proof. To check (3.13), note that using RHS for the right-hand side of the equation

$$\begin{aligned}\frac{d}{dt}RHS &= (\mu-\alpha)de^{(\mu-\alpha)dt}P(S_t^0=x) + e^{(\mu-\alpha)dt}\left[-d\mu P(S_t^0=x) + \sum_{y\sim x} d\mu \times \frac{1}{d}P(S_t^0=y)\right] \\ &= -\alpha de^{(\mu-\alpha)dt}P(S_t^0=x) + \sum_{y\sim x}\mu e^{(\mu-\alpha)dt}P(S_t^0=y) \\ &= -\alpha m(t,x) + \sum_{y\sim x} m(t,y)\mu\end{aligned}$$

which gives the desired result. □

Let $X_t = |S_t^0|$ be the distance from the root. We couple X_t to a simple random walk $\hat{X}_t$ on $\mathbb{Z}$ that jumps to the left at rate μ and to the right at rate $(d-1)\mu$ by using the following recipe: $\hat{X}_t$ follows the move of X_t if $X_t \neq 0$; when X_t jumps from 0 to 1, $\hat{X}_t$ jumps to the left with probability $1/d$. Clearly,

$$\hat{X}_t \leqslant X_t \qquad \forall t \geqslant 0$$

and hence

$$P(S_t^0 = 0) = P(X_t = 0) \leqslant P(\hat{X}_t \leqslant 0) \tag{3.14}$$

Note that if $\theta \leqslant 0$ then

$$\begin{aligned}
P(\hat{X}_t \leqslant 0) \leqslant Ee^{\theta \hat{X}_t} &= \sum_{k=0}^{\infty} e^{-d\mu t} \cdot \frac{(d\mu t)^k}{k!} \left(\frac{1}{d} e^{-\theta} + \frac{d-1}{d} e^{\theta} \right)^k \\
&= \exp\left\{ -d\mu t \left[1 - \left(\frac{1}{d} e^{-\theta} + \frac{d-1}{d} e^{\theta} \right) \right] \right\} \\
&= \exp\left\{ -\mu t [d - (e^{-\theta} + (d-1)e^{\theta})] \right\}
\end{aligned}$$

To optimize this bound we maximize the term in square brackets. To do this, we set

$$0 = \frac{d}{d\theta}[d - (e^{-\theta} + (d-1)e^{\theta})] = e^{-\theta} - (d-1)e^{\theta}$$

Solving we have $e^{2\theta} = 1/(d-1)$ or $e^{\theta} = 1/\sqrt{d-1}$, which leads to the bound

$$P(\hat{X}_t \leqslant 0) \leqslant \exp\left\{ -(d - 2\sqrt{d-1})\mu t \right\}$$

Using this with (3.13) and (3.14) we have

$$\begin{aligned}
m(t, 0) &= e^{(\mu - \alpha)dt} P(S_t^0 = 0) \\
&\leqslant \exp\left\{ \left[\left(d - (d - 2\sqrt{d-1}) \right) \mu - d\alpha \right] t \right\} \\
&= \exp\left\{ \left(2\sqrt{d-1}\mu - d\alpha \right) t \right\}
\end{aligned}$$

Since $\alpha = p_0 + d(1 - p_0)$ our assumption (3.10) implies the exponent is negative. We have completed the proof. □

Theorem 1.2.5. *On a d-regular tree the zealot voter model survives locally if*

$$p_0 = 0 \quad \text{and} \quad \mu > \frac{d}{\sqrt{d-1} + 1}.$$

Proof. Choose a self-avoiding path $\{e_n, -\infty < n < \infty\}$ in T^d such that $e_0 = 0$ is the root and $|e_n - e_{n+1}| = 1$. This gives an embedding of $\mathbb{Z}$ into T^d. Now define

$$u(n) = P\left(e_n \in \zeta_t \text{ for some } t\right)$$

for $n \geqslant 0$. By the strong Markov property, for all $n.m \geqslant 0$

$$u(n+m) \geqslant u(n)u(m)$$

i.e., the sequence is supermultiplicative. This implies that

$$\beta(\mu) \equiv \lim_{n\to\infty}[u(n)]^{1/n} = \sup_{m\geqslant 1}[u(m)]^{1/m}.$$

Let $S(e_0)$ denote the subtree starting from e_0 that does not include e_{-1}. Consider a lower bound $\bar{\zeta}_t$ on the dual COBRA where we birth are only allowed in $S(e_0)$. Our next step is to state a result from the contact process. This is Lemma 4.53 in [17] but the proof also works for our COBRA.

Lemma 3.2.3.

$$\lim_{n\to\infty}\left[\sup_t P(e_n \in \bar{\zeta}_t)\right]^{1/n} = \beta(\mu) \tag{3.15}$$

Since $\bar{\zeta}_t \subset \zeta_t$, the desired result follows from the next two Lemmas.

Lemma 3.2.4. *If $\beta(\mu) > 1/\sqrt{d-1}$, then $\inf_t P\left(e_0 \in \bar{\zeta}_t\right) > 0$.*

Lemma 3.2.5. *If $\mu > d/(\sqrt{d-1}+1)$, then $\beta(\mu) > 1/\sqrt{d-1}$.*

□

Proof of Lemma 3.2.4. The proof here is almost identical to the one on pages 99-100 of Liggestt [17]. According to Lemma 3.2.3 and our assumption, we can fix constants $a > 1/\sqrt{d-1}$, $n \geqslant 1$ and $s > 0$ such that

$$P\left(e_n \in \bar{\zeta}_s\right) = a^n \tag{3.16}$$

We now follow the proof of Proposition 4.57 in Liggett [17] to construct an embedded branching process. Let $B_0 = \{e\}$ and $B_1 = \{x \in \bar{\zeta}_s : |x - e| = n\}$. We ignore all the births outside $S(x)$ and apply the same rules leading from B_0 to B_1 to obtain a random subset $B(x)$ of $\{y \in S(x) \cap \bar{\zeta}_{2s} : |y - e| = 2n\}$. Let $B_2 = \cup_{x \in B_1} B(x)$. We repeat the same rule to construct a branching process B_j. Note $B_j \subset \bar{\zeta}_{js}$. Moreover B_j is supercritical since by (3.16) the offspring distribution has mean $(d-1)^n a^n > 1$. Then

$$\lim_{j \to \infty} \frac{|B_j|}{((d-1)^n a^n)^j}$$

exists and is positive with positive probability. As a result, we can find an ϵ such that for all sufficiently large j,

$$P\left(|B_j| > \epsilon\, ((d-1)a)^{nj}\right) > \epsilon$$

We will show particles from the subchain $\{B_{ji}\}_{i=0}^{\infty}$'s are sufficient to make the process survive locally. Since it takes time ijs to get to B_{ji}, we let

$$r_i = P(0 \in \bar{\zeta}_{2ijs}) \tag{3.17}$$

It follows from the strong Markov property that

$$r_{i+1} \geqslant P(x \in \bar{\zeta}_{2(i+1)js} \text{ for some } x \in B_j) P(e_{nj} \in \bar{\zeta}_{js}) \tag{3.18}$$

Let $\lfloor y \rfloor$ be the largest integer $\leqslant y$ and let $N = \lfloor \epsilon((d-1)a)^{nj} \rfloor$. This is

$$\begin{aligned} &\geqslant P(|B_j| > N)[1 - (1 - r_i)^N] P(e_{nj} \in \bar{\zeta}_{js}) \\ &\geqslant \epsilon[1 - (1 - r_i)^N] P(e_{nj} \in \bar{\zeta}_{js}) \end{aligned} \tag{3.19}$$

Using the strong Markov property on the last probability gives

$$P(e_{nj} \in \bar{\zeta}_{js}) \geqslant \left[P(e_n \in \bar{\zeta}_s)\right]^j = a^{nj} \tag{3.20}$$

Let $f(r) = \epsilon[1-(1-r)^N]a^{nj}$. Combining (3.19), (3.20) and (3.18) gives

$$r_{i+1} \geqslant f(r_i)$$

Note $f(r)$ is increasing over $[0,1]$ with $f(0) = 0$. Moreover, $f'(r) = \epsilon a^{nj} N(1-r)^{N-1}$. So using the definition of N,

$$\begin{aligned} f'(0) &= \epsilon a^{nj} N \\ &\geqslant \epsilon a^{nj}[\epsilon((d-1)a)^{nj} - 1] \\ &= \epsilon^2[a^2(d-1)]^{nj} - \epsilon a^{nj} \end{aligned}$$

Since $a > 1/\sqrt{d-1}$ this is > 1 if j is chosen large enough. Thus $f(r)$ has a fixed point $r^* \in (0,1]$. We will prove by induction that

$$r_i \geqslant r^*, \ \forall i \tag{3.21}$$

When $i = 0$, the inequality is trivial since $r_0 = 1$. Suppose $r_i \geqslant r^*$. By the monotonicity of $f(r)$, we have

$$r_{i+1} \geqslant f(r_i) \geqslant f(r^*) = r^*$$

To generalize (3.21) to all time t. Note particles die at rate d. Precisely

$$P(e \in \bar{\zeta}_t | e \in \bar{\zeta}_{2ijs}) \geqslant e^{-d(t-2ijs)}$$

In particular

$$P(e \in \bar{\zeta}_t) \geqslant e^{-djs} r_i, \text{ if } 2ijs < t < 2(i+1)js$$

We have completed the proof. □

Proof of Lemma 3.2.5. Consider a simple random walk on $\mathbb{Z}$ which takes steps

$$\begin{cases} +1 & \text{with probability } \frac{d-\mu}{d} \\ -1 & \text{with probability } \frac{\mu}{d} \end{cases}$$

Repeating the proof of Lemma 3.2.1 shows that $\phi(x) = \left(\frac{\mu}{d-\mu}\right)^x$ is a martingale. The stopping time theorem for martingales shows

$$P_n(T_0 < \infty) = \left(\frac{\mu}{d-\mu}\right)^n$$

Note

$$P(e_n \in \zeta_t \text{ for some } t > 0) \geqslant P_{e_n}(T_e < \infty) \geqslant \left(\frac{\mu}{d-\mu}\right)^n$$

where the second one is the probability that the COBRA initiated at e_n ever visits the root. Then

$$[u(n)]^{\frac{1}{n}} \geqslant \frac{\mu}{d-\mu}$$

Since $\mu > \frac{d}{\sqrt{d-1}+1}$, by assumption we have

$$\begin{aligned}\beta(\mu) = \lim_{n\to\infty}[\mu(n)]^{1/n} &\geqslant \frac{d}{d-\mu} - 1 \\ &> \frac{d}{d - d/(\sqrt{d-1}+1)} - 1 = \frac{1}{\sqrt{d-1}}\end{aligned}$$

which completes the proof of Lemma 3.2.5 and hence the proof of Theorem 1.2.5. □

3.3 Results for Galton-Watson Trees

3.3.1 Survival of COBRA

We will now prove Theorem 1.2.6. To lead up to that we will describe the proof of the main result in [10] in dimensions $d \geqslant 3$. The model under consideration there is a biased voter model with small bias. Jumps at x from $0 \to 1$ occur st rate $(1+\varepsilon)f_1(x)$, where $f_i(x)$ is the fraction of neighbors in state i, while jumps from $1 \to 0$ occur at rate $f_0(x)$. Suppose for concreteness that the neighborhood consists of the $2d$ nearest neighbors. As in the case of the zealot voter model the process is additive in the sense of Harris [12] and can be constructed from a graphical representation with independent Poisson processes, $T_n^{x,i}$, $n \geqslant 1$ for $i = 1, 2$. Let $e_1, \ldots, e_{2d}$ be an enumeration of the nearest neighbors of 0.

- The $T_n^{x,1}$ have rate 1 and have associated independent random variables $U_n^{x,1}$ that are uniform on $\{1, 2, \ldots 2d\}$. At time $T_n^{x,i}$ we write a δ at x that will kill a 1 at x and draw an arrow from $x + e(U_n^{x,i})$ to x. By considering the four cases for the states of $x + e(U_n^{x,i})$ and x we can easily check that this gadget causes x to imitate its neighbor.

- The $T_n^{x,2}$ have rate ε and have associated independent random variables $U_n^{x,2}$ that are uniform on $\{1, 2, \ldots 2d\}$. At time $T_n^{x,i}$ we draw an arrow from $x+e(U_n^{x,i})$ to x which will cause x to be 1 if $x + e(U_n^{x,i})$ is.

Since branching occurs at rate ε in dual, the suggests that we should run time at rate $1/\varepsilon$ and scale space by $1/\sqrt{\varepsilon}$ to mkae the dual process converge to a branching Brownian motion. One complication is that new born particles will coalesce with their parent with a probability γ which is the probability a random walk started at e_1 returns to 0. It is not hard to show that the probability such a coalescence will occur after time $1/\sqrt{\varepsilon}$ tends to 0. Thus in order for the sequence of processes to be tight, we do not add the newly born particle until time $1/\sqrt{\varepsilon}$ has elapsed. Other estimates in the proof show that it is unlikely for particle to coalesce with another particle that is not its parent, so the sequence of rescaled processes converges to a branching random walk in which new particle are born at rate γ.

In [10] this observation is combined with a block construction to prove the existence of a stationary distribution in a "hybrid zone" in which the process on $x_1 \geqslant 0$ is a biased voter model that favors 1 and on $x_1 < 0$ the process is a biased voter model favoring 0. Things are simpler for the zeaalot voter model on trees. If we only want to prove survival of the dual it is enough to prove that when time is run at rate

$1/\varepsilon$ the size of the dual converges to a supercritical branching process. Taking into account the fraction of time a random walk spends at vertices of degree k we arrive at:

Theorem 7. *Let $\delta > 0$. If $\varepsilon > 0$ is small enough then the COBRA dies out if*

$$\sum_k p_k(\mu_{m,k} - 1) - p_0 < -\delta$$

and survives if the last quantity is $> \delta$.

3.3.2 Local Survival

Proof of Theorem 1.2.7

Since the branching random walk gives an upper bound for the COBRA, it suffices to show

Lemma 3.3.1. *If $p_0 = 0$, then on a d-regular tree, the threshold for the local survival of η_t^0 satisfies*

$$\mu_l(\mathcal{T}^d) = d/(2\sqrt{d-1})$$

where η_t^0 is the branching random walk starting with 1 particle at the root

To prove this result, define $M(v,\ n)$ to be the number of oriented loops of length n starting from vertex v. It is well-known that the limit $L = \lim_{n\to\infty} M(v,\ 2n)^{1/2n} = \sup_{n\to\infty} M(v,\ 2n)^{1/2n}$ exists for all graphs, independent of the choice of vertex. This follows from a simple supermultiplicativity argument. Furthermore, define an *evolutionary walk* from vertex u to vertex v to be a sequence $0 \leqslant T_{n_0}^{x_0,i_0} < T_{n_1}^{x_1,i_1} < \cdots < T_{n_m}^{x_m,i_m} < \infty$ with $x_0 = u,\ x_m = v$. Precisely, this corresponds to a path in the graphical representation such that the fluid can flow from u to v. By definition, for a fixed path of length n on the tree, the expected number of evolutionary walks on this path is $(\mu/d)^n$. This is because when a branching event occurs, the expected number of particles landing on a certain neighbor is μ/d. We will show

Lemma 3.3.2. *Suppose* $L = \lim_{n\to\infty} M(v,\ 2n)^{1/2n}$. *Then* $\mu_l(\mathcal{T}^d) = d/L$.

Proof. Let X_n be the number of evolutionary walks of length n starting and ending at the root e_0. Note $\{X_{2nk}\}$ dominates a branching process with offspring distribution given by X_n. In particular,

$$EX_n \geqslant (\mu/d)^{2n} M(e_0, 2n)$$

So if $\mu > d/L$, this branching process is supercritical if n is sufficiently large. Choose n so that the above expectation is > 1. Note $\forall T > 0$, the expected number evolutionary walks of length n by time T is

$$\leqslant \Gamma(d,n) = \frac{d^n}{(n-1)!}\int_0^T e^{-s}s^{n-1}ds \leqslant \frac{(dT)^n}{n!}$$

The $\Gamma(d,n)$ comes from a sum of n exponential distributions with parameter d. Note this is summable with respect to n. By Borel-Cantelli Lemma, the maximal length of evolutionary walks within any finite time T is bounded and thus the root has to be visited infinitely often. For the other direction, note the expected number of evolutionary walks traversing e_0 is bounded by

$$\sum_{n=1}^{\infty} (\mu/d)^n M(e_0, n) < \infty, \text{ if } \mu < d/L$$

The proof is complete. ∎

Proof of Lemma 3.3.1. It remains to compute L. This is given by Pemantle and Stacey [22]. To summarize, note for an oriented loop of length $2n$, n steps are up (i.e. closer to e_0) n steps are down. At each step, there are $d-1$ choices to move farther away. Hence

$$M(0, 2n) = \binom{2n}{n}(d-1)^n$$

Use Stirling formula the desired result follows. ∎

Condition for Local Survival

We will now prove Theorem 1.2.8. In what follows, assume $p_0 = 0$. Let $p_x = \mu/d(x)$ be the probability that the particle at x will be replaced by a child moving closer to the root. Define a harmonic function depending on the distance to the root by

$$\phi(x) = p_x\phi(x-1) + (1-p_x)\phi(x+1) \tag{3.22}$$

Note (3.22) is equivalent to

$$\begin{aligned}\phi(x+1) - \phi(x) &= \frac{p_x}{1-p_x}\left[\phi(x) - \phi(x-1)\right] \\ &= \frac{\mu}{d(x)-\mu}\left[\phi(x) - \phi(x-1)\right]\end{aligned}$$

Suppose $0 = x_0, x_1, \ldots, x_n = x$ is the path from the root to x. We have

$$\phi(x_n) - \phi(x_{n-1}) = \prod_{k=1}^{n-1} \frac{\mu}{d(x_k)-\mu}\left[\phi(x_1) - \phi(0)\right] \tag{3.23}$$

This recursion allows us to impose function $\phi(x)$ on each vertex $x \in G(V, E)$. By Theorem 6.4.8 in [?] , if $\phi(x) \to \infty$ for all $l_x \to \infty$ then the dual survives locally. However, it is more convenient to pursue conditions such that the log increment $\log[\phi(x) - \phi(x-1)] \to \infty$ instead as we will see later.

Taking log of the recursion formula (3.23) gives

$$\log\left[\phi(x_n) - \phi(x_{n-1})\right] = \log\left[\phi(x_1) - \phi(0)\right] + \sum_{k=1}^{n-1} \log \frac{\mu}{d(x_k)-\mu}$$

Suppose the Galton-Watson tree has degree distribution $\{q_j\}$. Now consider a branching random walk on $\mathbb{R}$ which has an initial particle at the origin. With probability q_j, it gives birth to $j-1$ particles at $\log \frac{\mu}{j-\mu}$ and this forms a point process Z. The location of the first generation is denoted as $\{z_r^1\}$ where r is the index of each individual. For each particle x in the first generation, it generates new particles in a similar way. The location of its children has the same distribution as $\{z_r^1 + x\}$. We obtain the second generation by taking all the children of the first generation. Let $\{z_r^2\}$ be the locations of the second generation. The following generations are produced under the same manner. Denote $\{z_r^n\}$ as the location of the nth generation individuals. Let $F(t) = E\left[Z(-\infty, t]\right]$ be the expected number of points in Z to the left of t. Define

$$m(\theta) = \int_{-\infty}^{\infty} e^{-\theta t} dF(t) = \sum_{j\geqslant 3} q_j(j-1)\left(\frac{j-\mu}{\mu}\right)^{\theta}$$

To avoid notational confusion, we use $\nu(a) = \inf\{e^{\theta a} m(\theta) : \theta \geqslant 0\}$. This is (2.1) defined in [1]. It follows from Corollary (3.4) in [1] that $\nu(0) < 1$ implies $\log[\phi(x) - \phi(x-1)] \to \infty$ for all $l_x \to \infty$. Hence $\nu(0) < 1$ is a sufficient condition for local survival.

Remark. On the $d-$regular tree,

$$m(\theta) = (d-1)\left(\frac{d-\mu}{\mu}\right)^{\theta}$$

Then $v(0) < 1$ iff $\mu > d/2$, which gives another proof of Theorem **??**.

3.3.3 Degree = 3 and 4

Our recursion is

$$\phi(x+) - \phi(x) = \frac{\mu}{d(x) - \mu}(\phi(x) - \phi(x-))$$

$x-$ is neighbor closer to root. x^+ is any neighbor further away

$$m(\theta) = 2q_3\left(\frac{3-\mu}{\mu}\right)^{\theta} + 3q_4\left(\frac{4-\mu}{\mu}\right)^{\theta} \tag{3.24}$$

$$m'(\theta) = 2q_3\left(\frac{3-\mu}{\mu}\right)^{\theta}\log\left(\frac{3-\mu}{\mu}\right) + 3q_4\left(\frac{4-\mu}{\mu}\right)^{\theta}\log\left(\frac{4-\mu}{\mu}\right) \tag{3.25}$$

Reacall $\nu(0) = \min\{m(\theta) : \theta \geqslant 0\}$.

$\mu > 2$

Since $\mu/(3-\mu)$ and $\mu/(4-\mu)$ are both > 1, $\phi(x_n) \to \infty$ along any path $x_n \to \infty$, so the process survives strongly.

$\mu \leqslant 3/2$

Since $\mu/(3-\mu)$ and $\mu/(4-\mu)$ are both < 1, $\phi(x_n) \to 0$ along any path $x_n \to \infty$. However this only tells us that the proof fails.

$3/2 < \mu \leqslant 2$

Case 1. Note that if $2q_3 > 1$ there is a path to ∞ (which may not start at the root) along which we take the products of $\mu/(3-\mu)$ and hence $\phi(x_n) \to 0$, so the proof fails.

Case 2. $(4-\mu)/\mu > (3-\mu)/\mu$ so if

$$m'(0) = 2q_3 \log\left(\frac{3-\mu}{\mu}\right) + 3q_4 \log\left(\frac{4-\mu}{\mu}\right) > 0$$

then $m'(\theta) > 0$ for all $\theta > 0$ and the minimum occurs at 0. $m(0) = 2q_3 + 3q_4 \geqslant 2$, so again the proof fails. let $q_3 = p$ and $q_4 = 1-p$. For fixed μ, $m'(0)$ is linear in p so the condition holds when

$$p < p_c = \frac{3\log((4-\mu)/\mu)}{3\log((4-\mu)/\mu) + 2\log(\mu/(3-\mu))}$$

See Figure 3.2 for various μ.

Case 3. If $m'(0) < 0$ then a minimum at $\bar{\theta} > 0$ exists. Using (3.25) we want

$$2q_3 \left(\frac{3-\mu}{\mu}\right)^{\theta} \log\left(\frac{\mu}{3-\mu}\right) = 3q_4 \left(\frac{4-\mu}{\mu}\right)^{\theta} \log\left(\frac{4-\mu}{\mu}\right)$$

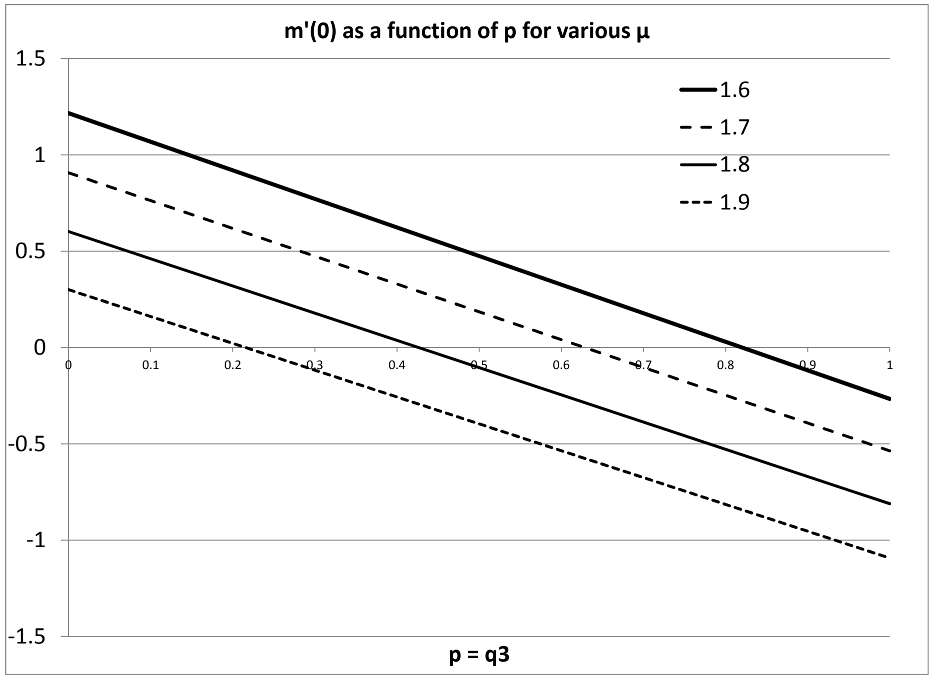


FIGURE 3.2: Local survival is possible only if $\mu'(0) < 0$. This is $q_3 > 0.85$ for $\mu = 1.6$; $q_3 > 0.65$ for $\mu = 1.7$; $q_3 > 0.45$ for $\mu = 1.8$ and $q_3 > 0.25$ for $\mu = 1.9$.

Cross multiplying

$$\left(\frac{4-\mu}{3-\mu}\right)^{\theta} = \frac{2q_3 \log(\mu/(3-\mu))}{3q_4 \log((4-\mu)/\mu)}$$

Let A be the numerator and B be the denominator of the fraction. $m'(0) < 0$ implies $A > B$. The LHS is 1 at $\theta = 0$ and increases $\to \infty$ as $\theta \to \infty$ so a solution exists. Taking logs

$$\theta \log((4-\mu)/(3-\mu)) = \log(A) - \log(B)$$

so we have

$$\bar{\theta} = \frac{\log(A) - \log(B)}{\log((4-\mu)/(3-\mu))} \tag{3.26}$$

There does not seem to be a good formula for $m(\bar{\theta})$. To compute it numerically, we choose $\mu = 1.6,\ 1.7,\ 1.8$ and 1.9 for

$$\nu(0) = m(\bar{\theta}) = 2q_3 \exp(\bar{\theta}\log((3-\mu)/\mu) + 3q_4 \exp(\bar{\theta}\log((4-\mu)/\mu)$$

It shows from the table that the phase transition occurs at $q_3 = 0.996,\ 0.97,\ 0.91$ and 0.82 respectively.

q_3	$\mu = 1.6$	$\mu = 1.7$	$\mu = 1.8$	$\mu = 1.9$
0.8	2.2	2.014149722	1.597069414	1.074539921
0.81	2.19	1.979137551	1.549560204	**1.030929768**
0.82	2.179999993	1.942042353	1.500601354	**0.896331678**
0.83	2.169035962	1.902724759	1.450116162	0.941977346
0.88	2.079229445	1.666138794	1.171184237	0.708137684
0.89	2.052259329	1.608953028	1.109102792	0.659079066
0.9	2.021286727	1.547575636	**1.044453965**	0.609119574
0.91	1.985646496	1.481425138	**0.976943138**	0.558174592
0.92	1.944464056	1.409753116	0.906197761	0.506138592
0.93	1.896552634	1.331568175	0.831734216	0.452876792
0.94	1.840236437	1.245508443	0.752903513	0.398211752
0.95	1.773023132	1.14961228	0.668796138	0.341900422
0.96	1.690934707	**1.040865809**	0.578059943	0.283591365
0.97	1.586938026	**0.914185487**	0.478505588	0.222735055
0.98	1.446391322	0.759622966	0.366073412	0.158358633
0.99	1.227510494	0.551306428	0.23102478	0.088284998
0.995	**1.037752234**			
0.996	**0.98268267**			
0.997	0.915774891			
0.998	0.828879261			

4

Conclusions

The voter model behaves differently when a very small latent period is introduced. Instead of having a one parameter stationary distribution and obtain the fixation configuration by time $O(n)$, the latent voter model has a quasi-stationary state 1/2 on the configuration model with n vertices. Indeed, each opinion has probability approximately 1/2 and the fraction of each opinion persists to it for a time is $\geqslant n^m$ for any $m < \infty$. However as we mentioned, this time of persistence is expected to be $\exp(-\gamma n)$, and this could be a future direction following this project.

In the zealot voter model, we obtained the sufficient condition for the model to survive on any tree with bounded degrees. Moreover, a necessary condition was found on $d-$regular trees. Furthermore, on $d-$trees, the phase transition of local survival, in terms of μ which the average number of chosen neighbors each time, occurs in the interval $[\frac{d}{2\sqrt{d-1}}, \frac{d}{1+\sqrt{d-1}}]$. The condition on Galton-Watson trees is more involed however. Hence simpler conditions for $GW-$trees is interesting to think about. Moreover, a more accurate phase transition for $d-$trees could be another future work.

Bibliography

[1] J. Biggins. Chernoff's theorem in the branching random walk. *Journal of Applied Probability*, 14(3):630–636, 1977.

[2] S. Chatterjee and R. Durrett. A first order phase transition in the threshold θ? 2 contact process on random r-regular graphs and r-trees. *Stochastic Processes and their Applications*, 123(2):561–578, 2013.

[3] C. Cooper, T. Radzik, and N. Rivera. The coalescing-branching random walk on expanders and the dual epidemic process. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 461–467. ACM, 2016.

[4] J. Cox and R. Durrett. Nonlinear voter models. In *Random walks, Brownian motion, and interacting particle systems*, pages 189–201. Springer, 1991.

[5] J. T. Cox and R. Durrett. Evolutionary games on the torus with weak selection. *Stochastic Processes and their Applications*, 126(8):2388–2409, 2016.

[6] J. T. Cox and R. Durrett. Evolutionary games on the torus with weak selection. *Stochastic Processes and their Applications*, 126(8):2388–2409, 2016.

[7] J. T. Cox, R. Durrett, and E. A. Perkins. *Voter model perturbations and reaction diffusion equations*, volume 349. Societé mathématique de France, 2013.

[8] R. Darling, J. R. Norris, et al. Differential equation approximations for markov chains. *Probability surveys*, 5:37–79, 2008.

[9] R. Durrett and C. Neuhauser. Particle systems and reaction-diffusion equations. *The Annals of Probability*, pages 289–333, 1994.

[10] R. Durrett and I. Zähle. On the width of hybrid zones. *Stochastic Processes and their Applications*, 117(12):1751–1763, 2007.

[11] D. Griffeath. *Additive and cancellative interacting particle systems*, volume 724. Springer, 2006.

[12] T. E. Harris. On a class of set-valued markov processes. *The Annals of Probability*, pages 175–194, 1976.

[13] R. Hollcy and T. Liggett. Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab*, 3:643–663, 1975.

[14] R. Huo and R. Durrett. Latent voter model on locally tree-like random graphs. *Stochastic Processes and their Applications*, 128(5):1590–1614, 2018.

[15] R. Lambiotte and S. Redner. Dynamics of vacillating voters. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(10):L10001, 2007.

[16] R. Lambiotte, J. Saramäki, and V. D. Blondel. Dynamics of latent voters. *Physical Review E*, 79(4):046107, 2009.

[17] T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324. springer science & Business Media, 2013.

[18] T. M. Liggett et al. Coexistence in threshold voter models. *The Annals of Probability*, 22(2):764–802, 1994.

[19] R. Ma, R. Durrett, et al. A simple evolutionary game arising from the study of the role of igf-ii in pancreatic cancer. *The Annals of Applied Probability*, 28(5):2896–2921, 2018.

[20] T. Mountford. A metastable result for the finite multidimensional contact process. *Canadian mathematical bulletin*, 36(2):216–226, 1993.

[21] T. Mountford, J.-C. Mourrat, D. Valesin, and Q. Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Processes and their Applications*, 126(7):1974–2013, 2016.

[22] R. Pemantle and A. M. Stacey. The branching random walk and contact process on galton-watson and nonhomogeneous trees. *Annals of probability*, pages 1563–1590, 2001.

Biography

Ran Huo graduated with distinction from Nanjing University, China in 2014, with a Bachelor's degree in Pure Mathematics and Applied Mathematics. In 2013, she spent two semesters at the University of Wisconsin Madison which sparked her interest in pursuing graduate studies in probability theory. She will obtain her Ph.D in Mathematics and Master's degree in Statistics in May 2019. After that, she will be a postdoc at the University of Melbourne, Australia.

Thanks for reading her thesis !