

## The $q$ -voter model on the torus\*

Pooja Agarwal<sup>†</sup>      Mackenzie Simper<sup>‡</sup>      Rick Durrett<sup>§</sup>

### Abstract

In the  $q$ -voter model, the voter at  $x$  changes its opinion at rate  $f_x^q$ , where  $f_x$  is the fraction of neighbors with the opposite opinion. Mean-field calculations suggest that there should be coexistence between opinions if  $q < 1$  and clustering if  $q > 1$ . This model has been extensively studied by physicists, but we do not know of any rigorous results. In this paper, we use the machinery of voter model perturbations to show that the conjectured behavior holds for  $q$  close to 1. More precisely, we show that if  $q < 1$ , then for any  $m < \infty$  the process on the three-dimensional torus with  $n$  points survives for time  $n^m$ , and after an initial transient phase has a density that it is always close to  $1/2$ . Readers familiar with long time survival results for the contact process and other particle systems might expect the conjecture to say survival occurs for time  $\exp(\gamma n)$  with  $\gamma > 0$ , however we show persistence does not hold for  $\exp(n^\beta)$  with  $\beta > 1/3$ . If  $q > 1$ , then the process rapidly reaches fixation on one opinion. It is interesting to note that in the second case the limiting ODE (on its sped up time scale) reaches 0 at time  $\log n$  but the stochastic process on the same time scale dies out at time  $(1/3) \log n$ .

**Keywords:** voter model perturbation; ODE limit; renormalization.

**MSC2020 subject classifications:** 60K35.

Submitted to EJP on June 4, 2020, final version accepted on August 6, 2021.

## 1 Introduction

In the linear voter model, the state at time  $t$  is  $\xi_t : \mathbb{Z}^d \rightarrow \{0, 1\}$ , where 0 and 1 are two opinions. The individual at  $x$  changes opinion at a rate equal to the fraction  $f_x$  of its neighbors with the opposite opinion. For the last decade physicists have studied the  $q$ -voter model, in which the flip rate at  $x$  is  $f_x^q$ . When  $q$  is an integer, the dynamics may be thought of as: select  $q$  neighbors of  $x$  uniformly at random, and change the opinion of  $x$  if all  $q$  neighbors disagree with  $x$ . However, there is no reason to restrict  $q$  to be an integer.

---

\*PA was partially supported by the NSF Grant DMS 1713032. MS was supported by a National Defense Science & Engineering Graduate Fellowship. RD was partially supported by NSF grant DMS 1809967 from the probability program.

<sup>†</sup>University of California, San Diego. E-mail: [pooja\\_agarwal@alumni.brown.edu](mailto:pooja_agarwal@alumni.brown.edu).

<sup>‡</sup>Stanford University. E-mail: [msimper@stanford.edu](mailto:msimper@stanford.edu).

<sup>§</sup>Duke University. E-mail: [rtd@math.duke.edu](mailto:rtd@math.duke.edu).

Abrams and Strogatz [1] introduced this system in 2003 as a model of language death, and argued based on data on languages in 42 regions that  $q = 1.31 \pm 0.25$ . In the physics literature there have been many studies of the system on lattices, complex networks, and even on graphs that co-evolve with the state of individuals. See [6, 17, 21, 24, 25, 27, 28] and references therein. According to [24], for finite but large systems, the process with  $q < 1$  can remain in a dynamically active phase for observation times that grow exponentially with  $n$ , while for  $q > 1$  the transition into an absorbing state is ‘abrupt’.

The difference between  $q < 1$  and  $q > 1$  is due to the different types of frequency dependence in the two models. When  $q < 1$ , rare opinions spread more rapidly compared to the voter model, while for  $q > 1$ , they spread more slowly. A more quantitative viewpoint is provided by mean field theory. This analysis is often done by writing an equation by pretending sites are always independent of each other. Here, we will instead consider the system on the complete graph in which each site interacts equally with all the others. In this case, the frequency of 1’s,  $u$ , satisfies

$$du/dt = -u(1-u)^q + (1-u)u^q = u(1-u)g(u)$$

where  $g(u) = u^{q-1} - (1-u)^{q-1}$ . This system has three fixed points: 0,  $1/2$  and 1.

- If  $q < 1$ ,  $g(u)$  decreases from  $\infty$  to  $-\infty$  as  $u$  increases from 0 to 1. So the fixed points 0 and 1 are unstable and the interior one is attracting. In this case it is expected that coexistence occurs.
- If  $q > 1$ ,  $g(u)$  increases from  $-1$  to 1 as  $u$  increases from 0 to 1. So the fixed points 0 and 1 are stable and the interior one is unstable. In this case it is expected that clustering occurs. That is, we will see larger and large regions occupied by one type.

For more on the heuristics that lead to these conclusions, see the 1994 paper by Durrett and Levin [12]. In most of the papers in the physics literature, the analysis is done by using the pair-approximation, which is equivalent to supposing that the state of the system is always a Markov chain.

Recently, Vasconcelos, Levin, and Pinheiro [29] have considered a version of the  $q$ -voter in which the powers  $q_1$  and  $q_0$  for flipping to 1 and 0 can be different. They did this to study complex contagions which have been used to model the spread of idioms and hashtags on Twitter [26] and in many other situations, see the book by Centola [7]. When  $q_1 \neq q_0$ , situations arise when one opinion dominates the other, see Figure 2a in [29], but the situation with  $q_1 = q_0$  seems to capture of all of the interesting behavior.

## 1.1 Voter model perturbations

The linear voter model has a rich theory due to its duality with coalescing random walk. This duality exists because the process can be constructed from a graphical representation. See Section 2.1 for details. However, the inherent asymmetry between 1’s and 0’s in the graphical representation makes it impossible to construct nonlinear voter models where the flip rates depend only on  $f_x$ . See Section 2.2 for a proof.

To get around this difficulty, we will suppose  $q$  is close to 1 and view the  $q$ -voter model as a voter model perturbation in the sense of Cox, Durrett, and Perkins [10]. On  $\mathbb{Z}^d$ , this theory requires  $d \geq 3$  so that the voter model has a one parameter family of stationary distributions  $\nu_u$ ,  $0 \leq u \leq 1$ . That is, for an initial voter configuration generated from the product measure with marginals  $u$ , the distribution  $\nu_u$  is the limiting invariant measure of the voter model. For this and other elementary facts about the voter model that we use, see Liggett’s 1999 book [23].

In general, for a voter at  $x$ , the rate of flipping from state  $i$  to  $j \neq i$  in a voter perturbation has the form

$$c_{i,j}^\delta(x, \xi) = f_j(x, \xi) + \delta^2 h_{i,j}(x, \xi)$$

where  $f_j(x, \xi)$  is the fraction of neighbors of the voter at  $x$  in state  $j$ , and  $h_{i,j}(x, \xi)$  is the perturbation to the rate of flipping from  $i$  to  $j$ . Usually the perturbation variable is  $\epsilon$ , but here it will be convenient to let  $\epsilon = \delta^2$ . To simplify formulas we will assume  $h_{i,j}(x, \xi) = 0$  when  $\xi(x) \neq i$ . Here we will consider the special case in which the neighborhood has size  $k$  and the flip rate only depends on the number of neighbors  $n(x)$  in state  $j$ :

$$c_{i,j}^\delta(x, \xi) = f_j(x, \xi) + \delta^2 r_{n(x)}^k \quad \text{for } 1 \leq n(x) \leq k,$$

where  $r_{n(x)}^k$  is the perturbation to the rate of flipping from  $i$  to  $j$  and depends only on the number of neighbors of  $x$  in state  $j$ . The  $r_{n(x)}^k$  do not have to be nonnegative, see (1.7) in [10], but we will suppose  $r_0^k = 0$  so that  $\equiv 0$  and  $\equiv 1$  are absorbing states. For simplicity, we will restrict our attention to three dimensions. In that context, we will consider neighborhoods  $x + \mathcal{N}$  with  $0 \notin \mathcal{N}$  and  $|\mathcal{N}| \geq 3$  chosen so that the group generated by  $\mathcal{N}$  is  $\mathbb{Z}^3$

**q-voter model.** The rate at which a site  $x$  flips to 0 in the  $q$ -voter model is  $f_x^q$ , where  $f_x$  is the fraction of neighbors with the opposite opinion. Suppose for the moment that  $q < 1$ . In this case, if we write

$$f_x^q = f_x + (f_x^q - f_x),$$

then the term in parentheses is  $\geq 0$ . Let  $q = 1 - \delta^2$  and write  $u$  instead of  $f_x$ . Then,

$$u^q - u = u \left( u^{-\delta^2} - 1 \right) = u \left( \exp(\delta^2 \log(1/u)) - 1 \right) \approx \delta^2 u \log(1/u).$$

From this we see that if  $q < 1$  and  $m = n(x)$ , then the perturbation is

$$r_m^k = (m/k) \log(k/m). \quad (1.1)$$

which vanishes when  $n(x) = 0$  or  $k$ .

If we let  $q = 1 + \delta^2$  and again write  $u$  instead of  $f_x$ , then

$$u^q - u = u \left( u^{\delta^2} - 1 \right) = u \left( \exp(\delta^2 \log(u)) - 1 \right) \approx \delta^2 u \log(1/u).$$

Hence when  $q > 1$ , for  $m = n(x)$ , the perturbation is

$$r_m^k = -(m/k) \log(k/m). \quad (1.2)$$

Intuitively, Theorem 1.1 holds due to a separation of time scales. The voter model runs at a fast rate, so when the density is  $u$  on the torus, the system has distribution  $\approx \nu_u$ . The rate of change of the density can then be computed by looking at the expected rate of change when the state is  $\nu_u$ . Writing  $\langle \cdot \rangle_u$  for expected value with respect to  $\nu_u$ , the right hand side of the ODE is

$$\phi(u) = \langle h_{0,1} - h_{1,0} \rangle_u = \sum_{m=1}^{k-1} r_m^k (\rho_m^0(u) - \rho_m^1(u)). \quad (1.3)$$

This result will be proved by constructing the process on a graphical representation and then defining a dual that is a coalescing branching random walk. The voter part of the process leads to a coalescing random walk. When a perturbation event occurs at a point  $x$ , the dual branches to include all of the points in  $x + \mathcal{N}$ . This will be described in detail

in Section 2.3. The proof of Theorem 1.1 is almost identical to the proof of Theorem 6 in Cox and Durrett [8] so we will only outline the proof, referring to [8] for details. When  $\epsilon_n \gg n^{-2/3}$  the particles in the dual have time to wrap around the torus and come to equilibrium in between branching events. It is known that on the torus if we start two random walks from independent randomly chosen locations, then the time to coalesce is of order  $n$ . Thus the assumption  $\epsilon_n \ll n^{-1}$  is needed for the perturbation to have an effect.

## 1.2 ODE limit

Following the approach of Cox and Durrett [8], who used the voter perturbation machinery to study evolutionary games on the torus in dimension  $d \geq 3$ , we will consider the  $q$ -voter model in what they called the weak-selection regime. (For results in the strong selection regime see Section 1.4.) Let  $\mathbb{T}_n$  be the three dimensional torus with  $n$  points and hence side length  $L = n^{1/3}$ . Let  $\epsilon_n > 0$  and define  $\delta_n$  such that  $\epsilon_n = \delta_n^2$ . The first thing to do is to prove convergence of the density of 1's in a scaled time regime,

$$U_n(t) = \frac{1}{n} \sum_{x \in \mathbb{T}_n} \xi_{t/\epsilon_n}(x),$$

to the solution of an ODE. Let  $\rho_m^i$  denote the probability that in  $\nu_u$  the origin is in state  $i$  while exactly  $m$  of the neighbors are in state  $1 - i$ . We write  $a_n \ll b_n$  for positive quantities  $a_n$  and  $b_n$  to indicate  $a_n/b_n \rightarrow 0$  as  $n \rightarrow \infty$ .

**Theorem 1.1.** *Suppose  $q = 1 - \epsilon_n$  with  $n^{-1} \ll \epsilon_n \ll n^{-2/3}$ . If  $U_n(0) \rightarrow u_0$ , then  $U_n(t)$  converges uniformly on compact sets to the solution of the ODE*

$$\frac{du}{dt} = \sum_{m=1}^{k-1} r_m^k (\rho_m^0(u) - \rho_m^1(u)), \quad u(0) = u_0, \quad (1.4)$$

where  $r_m^k$  is as defined in (1.1).

Intuitively, Theorem 1.1 holds due to a separation of time scales. The voter model runs at a fast rate, so when the density is  $u$  on the torus, the system has distribution  $\approx \nu_u$ . The rate of change of the density can then be computed by looking at the expected rate of change when the state is  $\nu_u$ . Writing  $\langle \cdot \rangle_u$  for expected value with respect to  $\nu_u$ , the right hand side of the ODE is

$$\phi(u) = \langle h_{0,1} - h_{1,0} \rangle_u = \sum_{m=1}^{k-1} r_m^k (\rho_m^0(u) - \rho_m^1(u)).$$

This result will be proved by constructing the process on a graphical representation and then defining a dual that is a coalescing branching random walk. The voter part of the process leads to a coalescing random walk. When a perturbation event occurs at a point  $x$ , the dual branches to include all of the points in  $x + \mathcal{N}$ . This will be described in detail in Section 2.3. The proof of Theorem 1.1 is almost identical to the proof of Theorem 6 in Cox and Durrett [8] so we will only outline the proof, referring to [8] for details. When  $\epsilon_n \gg n^{-2/3}$  the particles in the dual have time to wrap around the torus and come to equilibrium in between branching events. It is known that on the torus if we start two random walks from independent randomly chosen locations, then the time to coalesce is of order  $n$ . Thus the assumption  $\epsilon_n \ll n^{-1}$  is needed for the perturbation to have an effect.

Computing the  $r_m^k$ , see Section 5, leads to the following ODE

**Theorem 1.2.** *In three dimensions, when the neighborhood has size  $k$ , the limiting ODE is*

$$\frac{du}{dt} = \pm c_k u(1-u)(1-2u)f_k(u)$$

where  $c_k > 0$  and  $f_k$  is a polynomial that is positive on  $[0, 1]$  and  $f(0) = f(1) = 1$ . We have  $+$  for  $q < 1$  and  $-$  for  $q > 1$ .

When  $q < 1$ , the fixed point at  $1/2$  is attracting and we have

**Theorem 1.3.** *Suppose  $q = 1 - \epsilon_n$  and  $\epsilon_n \sim Cn^{-a}$  for some  $a \in (2/3, 1)$ . There is a  $T_0$  that only depends on  $u_0$ , so that for any  $\gamma > 0$  and  $m < \infty$ , if  $n$  is large then with high probability*

$$|U_n(t) - 1/2| \leq \gamma \text{ for all } t \in [T_0, n^m].$$

Here and in what follows “with high probability” means with probability  $\rightarrow 1$  as  $n \rightarrow \infty$ .

To prove Theorem 1.3, we will follow the approach of Huo and Durrett [20] who proved a similar result for the latent voter model on a random graph generated by the configuration model. Although the random graph has a more complicated geometry than the torus, the proof in that setting is simpler than the one given here, since on the graph random walks mix in time  $O(\log n)$  rather than in time  $O(n^{2/3})$ .

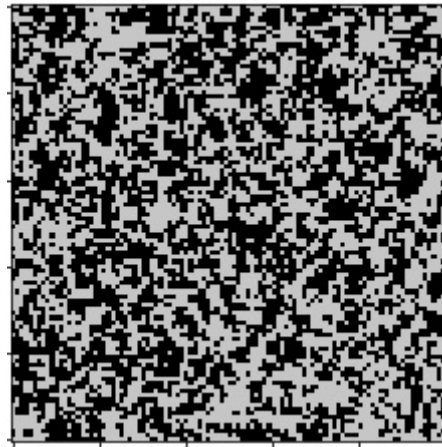


Figure 1: Cross-section from a simulation of  $q = 0.9$  on a  $100 \times 100 \times 100$  grid with periodic boundary conditions.

### Outline of the proof of Theorem 1.3.

- Section 3.1 introduces a general result for proving convergence of stochastic processes to limiting ODEs, due to Darling and Norris [11], which is the key to the proofs of the persistence results for our model (and for the latent voter model). The main difficulty is to bound the difference between the drift in the density  $U_n$  of the particle system and the drift in the ODE. In particular, one must prove that the drift in the density of  $U_n$ , which is a function of the configuration, is almost a function of the overall density.
- In Section 3.2 we take the first step in the proof, which is to show that if  $2/3 < b < a$  then we can ignore the perturbation on  $[t/\epsilon_n - n^b, t/\epsilon_n]$ , i.e., the process will evolve like the voter model. This has the consequence that if there are  $n \cdot u$  1's at time  $t/\epsilon_n - n^b$ , then at time  $t/\epsilon_n$  the process is close to the voter equilibrium  $\nu_u$ . The argument here is an improvement over the one in Section 3.1 of [20]. We use Azuma's inequality to get error estimates that are stretched exponentially small, i.e.,  $\leq C \exp(cn^{-\alpha})$  with  $\alpha > 0$  rather than polynomial, i.e.,  $\leq Ct^{-p}$ .

- In Section 3.3 we introduce a result about “renormalizing” the voter model, that comes from work of Bramson and Griffeath [4] in  $d = 3$  and Zähle [30] in  $d \geq 3$ . They show that if we consider the number of 1’s in the voter model equilibrium with density  $\lambda$ ,  $\xi^\lambda$ , in a cube  $Q(r)$  of side length  $r$ , then

$$\hat{S}_r = (\lambda(1 - \lambda))^{-1/2} r^{-5/2} \left( \sum_{x \in Q(r)} \xi^\lambda(x) - \lambda \right) \Rightarrow \text{Normal}(0, C) \quad (1.5)$$

We use this to obtain information about a similar normalized sum  $T_r$  of the number of ones in a cube of side  $r$  on the torus at time  $t/\epsilon_n$  when the number of 1’s at time  $t/\epsilon_n - n^b$  is  $\lambda n$ . To be specific, we let  $\tilde{S}_n$  be the normalized sum of  $\xi_{\sigma(n)}^\lambda(x)$  in the process that starts at time 0 from product measure with density  $\lambda$  and is run for time  $\sigma(n) = n^{0.6}$ . We show that  $\tilde{S}_r \leq T'_r \leq \hat{S}_r$ , where  $T'_r$  is a small modification of  $T_r$ .

- In Section 3.4 we bound the difference between  $T'_r$  and  $T_r$ . This in turn gives us a bound on the largest coalescing random walk cluster in  $T_r$  in  $Q(r)$ , see (3.14), and a bound on the fluctuations of the density in the cubes, which is important for completing the next step.
- In Section 3.5 we bound the difference between the drifts in the particle system and the ODE. To do this, we have to show that the empirical finite distributions on the torus  $\mathbb{T}_n$  (see Section 3.5 for a precise definition) are close to the values that come from voter model stationary distributions  $\nu_u$ . In doing this we rely on the result about the density in cubes proved in Section 3.3 to divide space at time  $t/\epsilon_n + s_n$  into cubes with  $n^{b(3)}$  sites, where  $r = n^{b(2)/3}$  and  $b(3) > b(2)$ . Here  $s_n = n^{(2+\alpha)b(2)/3}$  with  $\alpha$  small, so that the empirical f.d.d.’s in cubes of volume  $n^{b(3)}$  that do not touch are almost independent. This leads to errors of size  $C \exp(-n^{1-b(3)-2\alpha})$ .
- In Section 3.6 we put the pieces together to prove the result. As in Section 3.5 of [20] we do this by showing that if the density  $U_n(t)$  reaches  $|U_n(t) - 1/2| = 4\epsilon$ , then with very high probability (i.e., for any  $k$  with probability  $\geq 1 - n^{-k}$  for large  $n$ ) it will return to  $|U_n(t) - 1/2| \leq \epsilon$  before we have  $|U_n(t) - 1/2| > 5\epsilon$ . Taking  $\delta = 5\epsilon$  gives the desired result

In all of our estimates except those in Sections 3.3 and 3.4, the errors are stretched exponentially small, so we have the following:

**Conjecture.** When  $q < 1$  the process persists for time  $\exp(n^\beta)$  for some  $\beta > 0$ .

This could be proved with a rather small value of  $\beta$  if the errors in (3.12) and (3.14) could be improved to be stretched exponentially small. Readers familiar with long time survival results for the contact process, see e.g., Section 3 in part I of Liggett [23], might expect the conjecture to say survival occurs for time  $\exp(\gamma n)$  with  $\gamma > 0$ . However, the conjecture above cannot hold for  $\beta > 1/3$ . If we run time backwards from  $t/\epsilon_n$  to  $t/\epsilon_n - n^{2/3}$  then the  $n$  initial particles in the CRW will have coalesced to  $n^{1/3}$  particles. If all of these happen to land on sites in state 0 at time  $t/\epsilon_n - n^{2/3}$ , then the process will go extinct at time  $t/\epsilon_n$ .

### 1.3 Rapid Extinction when $q > 1$

When  $q > 1$ , the fixed point at  $1/2$  is unstable while the ones at 0 and 1 are locally attracting. To get rid of the constant  $c_k$  in the ODE limit in Theorem 1.2 we consider

$$U_n(t) = \frac{1}{n} \sum_{x \in \mathbb{T}_n} \xi_{t/\epsilon_n c_k}(x)$$

**Theorem 1.4.** Suppose  $q = 1 + \epsilon_n$  and  $\epsilon_n \sim Cn^{-a}$  for some  $a \in (2/3, 1)$ . If  $U_n(0) = u_0 < 1/2$  and  $\alpha > 1/3$  then

$$P(U_n(\alpha \log n) = 0) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

This is proved in Section 5. Much of the work for the proof of Theorem 1.4 has already been done in the proof of Theorem 1.3. Those results imply that the density in the particle system stays close to the solution of the ODE. To be precise, we can show that with high probability

$$|U_n(t) - u(t)| \leq \epsilon u(t) \quad \text{until } \tau = \inf\{t : x_t \leq n^{-(1-b(0))}\}$$

where  $2/3 < b(0) < \min\{b, 1 - \alpha\}$ . Since the ODE is  $u'(t) = -f(u)$  with  $f(u)/u \rightarrow 1$  as  $u \rightarrow 0$ , the limiting ODE has  $u(\alpha \log n) \approx n^{-\alpha}$ . Our proof shows that when the density gets to  $\leq n^{-b(0)}$  fluctuations in the voter model make the system go extinct in a time that is  $\leq Cn^b$ . See Section 4 for details. The keys to the voter extinction result are (i) the observation that the number of 1's in the voter model is a time change of continuous-time symmetric random walk, and (ii) results on the size of the boundary of the voter model in the low density regime due to Cox, Durrett, and Perkins [9].

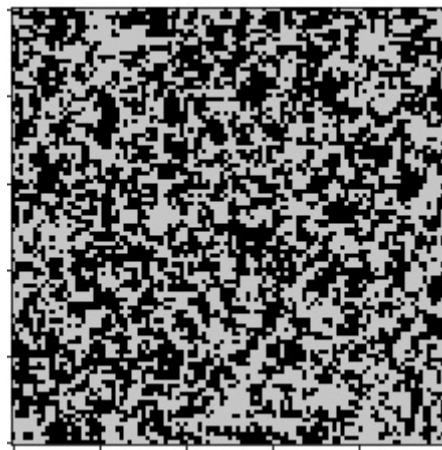


Figure 2: Cross-section from a simulation of  $q = 1.1$  on a  $100 \times 100 \times 100$  grid with periodic boundary conditions.

#### 1.4 Results for strong selection

Let  $\xi_t^\epsilon$  be a voter model perturbation on  $\mathbb{Z}^d$  with flip rates

$$c_{i,j}^{\delta_n}(x, \xi) = f_j(x, \xi) + \delta_n^2 h_{i,j}(x, \xi)$$

where  $f_j(x, \xi)$  is the fraction of neighbors in state  $j$  and the second term is the perturbation. As before we let  $\epsilon_n = \delta_n^2$ . In this section we will examine the case  $\epsilon_n \gg n^{-2/3}$ , which we call the strong selection regime.

Intuitively, the next result says that if we rescale space to  $\delta_n \mathbb{T}_n$  (recall  $\mathbb{T}_n$  is the three dimensional torus with  $n$  points) and speed up time by  $\delta_n^{-2}$ , then the process converges to the solution of a partial differential equation on  $\mathbb{R}^3$ . The torus turns into  $\mathbb{R}^3$  in the limit because  $\delta_n \ll n^{-1/3}$  while the torus has side  $n^{1/3}$ . To make a precise statement, the first thing we have to do is to define the mode of convergence. To simplify the writing we drop the subscript  $n$  on  $\delta$ . Given  $r \in (0, 1)$ , let  $a_\delta = \lceil \delta^{r-1} \rceil \delta$ ,  $Q_\delta = [0, a_\delta)^3$ , and  $|Q_\delta|$  the

number of points in  $Q_\delta$ . For  $x \in a_\delta \mathbb{Z}^3$  and  $\xi \in \Omega_\delta$ , the space of all functions from  $\delta \mathbb{Z}^3$  to  $S = 0, 1$ , let

$$D_i(x, \xi) = |\{y \in Q_\delta : \xi(x + y) = i\}| / |Q_\delta|.$$

We endow  $\Omega_\delta$  with the  $\sigma$ -field  $\mathcal{F}_\delta$  generated by the finite-dimensional distributions. Given a sequence of measures  $\lambda_\delta$  on  $(\Omega_\delta, \mathcal{F}_\delta)$  and continuous functions  $w_i : \mathbb{R}^3 \rightarrow [0, 1]$ , we say that  $\lambda_\delta$  has asymptotic densities  $w_i$  if for all  $0 < \eta, R < \infty$  and all  $i \in S$

$$\lim_{\delta \rightarrow 0} \sup_{x \in a_\delta \mathbb{Z}^3, |x| \leq R} \lambda_\delta(|D_i(x, \xi) - w_i(x)| > \eta) \rightarrow 0.$$

**Theorem 1.5.** Suppose  $d = 3$ . Let  $w_i : \mathbb{R}^d \rightarrow [0, 1]$  be continuous with  $\sum_{i \in S} w_i = 1$ . Suppose the initial conditions  $\xi_0^\delta$  have laws  $\lambda_\delta$  with asymptotic densities  $w_i$  and let

$$u_i^\delta(t, x) = P(\xi_{t\delta^{-2}}^\delta(x) = i).$$

If  $x_\delta \rightarrow x$ , then  $u_i^\delta(t, x_\delta) \rightarrow u_i(t, x)$ , where  $u_i$  is the solution of the system of partial differential equations:

$$\frac{\partial}{\partial t} u_i(t, x) = \frac{\sigma^2}{2} \Delta u_i(t, x) + \phi_i(u(t, x)) \quad (1.6)$$

with initial condition  $u_i(0, x) = w_i(x)$ . The reaction term is

$$\phi_i(u) = \sum_{j \neq i} \langle h_{j,i}(0, \xi) - h_{i,j}(0, \xi) \rangle_u, \quad (1.7)$$

where the brackets are expected value with respect to the voter model stationary distribution  $\nu_u$  in which the densities are given by the vector  $u$ .

This result is Theorem 2 in [8]. For more details see that paper.

The intuition is similar to that for the ODE limit in Theorem 1.1. On the fast time scale the voter model runs at rate  $\delta^{-2}$  versus the perturbation at rate 1, so the states of sites near  $x$  at time  $t$  is always close to the voter equilibrium  $\nu_{u(t,x)}$ . Thus, we can compute the rate of change of  $u_i(t, x)$  by assuming the nearby sites are distributed according to the voter model equilibrium  $\nu_{u(t,x)}$ .

Cox and Durrett considered evolutionary games on the torus in  $d \geq 3$  with game matrix  $1 + wG$ , where  $1$  is a matrix of 1's. Their  $w$  corresponds to our  $\epsilon_n$ . When  $w = 0$  the system reduces to the voter model. They found convergence to an ODE when  $n^{-1} \ll w \ll n^{-2/d}$  and convergence to a PDE when  $w \gg n^{-2/d}$ . Their results can be used to prove a PDE limit for our system when  $\epsilon_n \gg n^{-2/d}$ . Since there are only two opinions we only need one variable  $u_1$ , which corresponds to our  $u$ . The  $\phi$  in (1.7) is the same as the second term on the right hand side of our ODE, which should be clear from (1.3).

In the case of a  $2 \times 2$  game with a stable mixed strategy equilibrium that uses strategy 1 with probability  $\rho$  and strategy 2 with probability  $1 - \rho$ , the limiting function is  $\phi(u) = cu(\rho - u)(1 - u)$  with  $c > 0$ . Here, as in the case  $q < 1$ , the fixed point  $\rho$  is attracting. To translate Theorem 4 in [8] to our situation, we note that  $w = \epsilon_L^2$  and  $n = L^3$ .

**Theorem 1.6.** Suppose that  $w \sim Cn^{-2\alpha/3}$ , where  $0 < \alpha < 1$ , and that we start from a product measure in which each type has positive density. Let  $N_1(t)$  be the number of sites occupied by 1's at time  $t$ . There is a  $c > 0$  so that for any  $\eta > 0$  if  $n$  is large and  $\log n \leq t \leq \exp(cn^{(1-\alpha)})$ , then  $N_1(t)/N \in (\rho - \eta, \rho + \eta)$  with high probability.

The intuition behind the answer is that after space is rescaled the volume of the torus is asymptotically  $n^{(1-\alpha)}$ . Theorem 1.6 is a lower bound so it does not rule out survival for time  $\exp(cn)$ . However, Cox and Durrett proved for the contact process with fast voting introduced by Durrett, Liggett, and Zhang [13]

**Theorem 1.7.** *There is a  $C < \infty$  so that extinction in the contact process plus fast voting occurs by time  $\exp(cn^{1-2\alpha/d} \log n)$  in  $d \geq 3$ .*

Theorem 1.6 can be generalized to the  $q$ -voter with  $q < 1$  since it only relies on the hydrodynamic limit in Theorem 1.5 and a block construction. Theorem 1.7 does not extend, because  $\xi \equiv 1$  is an absorbing state, and this limits our ability to suddenly kill the process.

## 2 Graphical representation, duality

### 2.1 Voter model

We begin by describing the graphical representation and duality for the voter model in which the neighbors of  $x$  are  $x + \mathcal{N}$  and  $\mathcal{N} = \{y_1, \dots, y_k\}$ . The state of the voter model at time  $t$  is  $\xi_t : \mathbb{Z}^d \rightarrow \{0, 1\}$  where  $\xi_t(x)$  gives the opinion of the individual at  $x$  at time  $t$ . We write  $y \sim x$  to indicate that  $y$  is a neighbor of  $x$ . In the usual voter model, the rate at which the voter at  $x$  changes its opinion from  $i$  to  $j$  is

$$c_{i,j}^v(x, \xi) = 1_{(\xi(x)=i)} f_j(x, \xi),$$

where  $f_j(x, \xi) = (1/k) \sum_{i=1}^k 1_{(\xi(x + y_i) = j)}$  is the fraction of neighbors in state  $j$ .

To study the voter model, it is convenient to construct the process on a **graphical representation**, introduced by Harris [18] and further developed by Griffeath [16]. For each  $x \in \mathbb{Z}^d$  and  $y \in x + \mathcal{N}$  let  $T_m^{x,y}$ ,  $m \geq 1$ , be the arrival times of a Poisson process with rate  $1/k$ . At the times  $T_n^{x,y}$ ,  $n \geq 1$ , the voter at  $x$  decides to change its opinion to match the one at  $y$ . To indicate this, at time  $T_n^{x,y}$  we write a  $\delta$  at  $x$  and draw an arrow from  $y$  to  $x$ . To calculate the state of the voter model on a finite set, we start at the bottom and work our way up. We think of the 1's in the initial configuration as sources of fluid, the  $\delta$ 's as dams that block the fluid, while the arrows move the fluid in the direction indicated. Arrows from  $y$  to  $x$  arrive just after the  $\delta$ . A nice feature of this approach is that it simultaneously constructs the process for all initial conditions so that if  $\xi_0(x) \leq \xi'_0(x)$  for all  $x$ , then for all  $t > 0$  we have  $\xi_t(x) \leq \xi'_t(x)$  for all  $x$ .

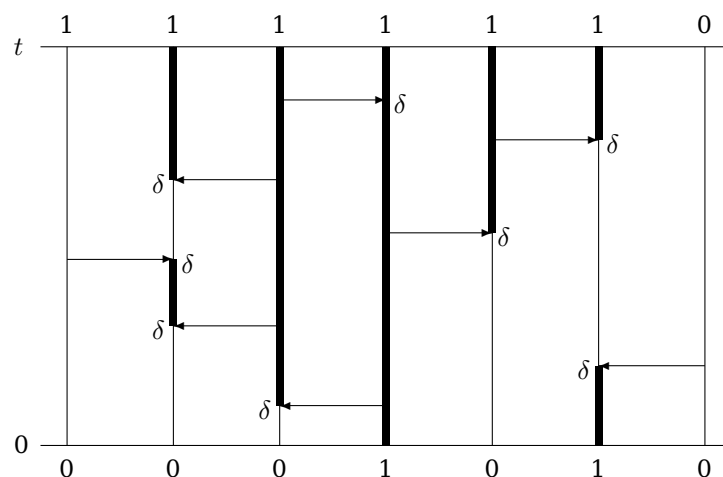


Figure 3: Voter model graphical representation

To define the **dual process** starting from  $x$  at time  $t$ , we set  $\zeta_0^{x,t} = x$  and work down the graphical representation. A particle stays at its current location until the first time that it encounters a  $\delta$ . At this point it jumps across the edge in the direction opposite its orientation. A little thought reveals that the path of a single particle in  $\zeta_s^{x,t}$ ,  $0 \leq s \leq t$ , is a random walk that at rate 1 jumps to a randomly chosen neighbor. Intuitively,  $\zeta_s^{x,t}$  gives the source at time  $t - s$  of the opinion at  $x$  at time  $t$ . That is,

$$\xi_t(x) = \xi_{t-s}(\zeta_s^{x,t}).$$

The example in Figure 3 should help explain the definitions. Here we work backwards to determine the states of the two sites marked by '?'. The dark lines indicate the locations of the two dual particles. The family of particles  $\zeta_s^{x,t}$  are coalescing random walks. That is, if a particle  $\zeta_s^{x,t}$  lands on the site occupied by  $\zeta_s^{y,t}$ , the two particles coalesce to form a single particle, and we know that  $\xi_t(x) = \xi_t(y)$ .

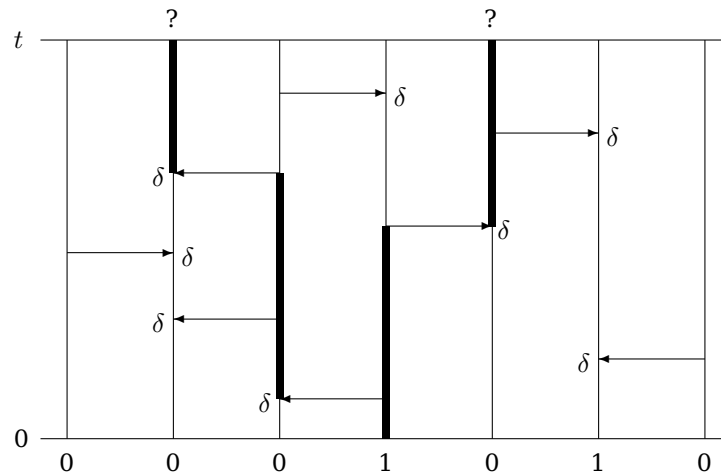


Figure 4: Dual coalescing random walk

To illustrate the power of duality, we analyze the asymptotic behavior of the voter model on  $\mathbb{Z}^d$ , proving a result of Holley and Liggett [19]. In dimensions 1 and 2, nearest neighbor random walk is recurrent, so the voter model clusters, i.e.,

$$P(\xi_t(x) \neq \xi_t(y)) \leq P(\zeta_t^{x,t} \neq \zeta_t^{y,t}) \rightarrow 0.$$

In  $d \geq 3$  random walks are transient so differences in opinion persist as  $t \rightarrow \infty$ . Let  $\xi_t^u$  be the voter model starting from product measure in which 1's have density  $u$ , i.e., the initial voter opinions are independent and = 1 with probability  $u$ . For a finite set  $B \subset \mathbb{Z}^d$ , let  $\zeta_s^{B,t} = \cup_{x \in B} \zeta_s^{x,t}$ . The distribution of  $\zeta_s^{B,t}$  does not depend on  $t$  so we drop the superscript  $t$ . Duality implies

$$P(\xi_t^u \equiv 0 \text{ on } B) = P(\xi_0^u(y) = 0 \text{ for all } x \in \zeta_t^B) = E\left((1-u)^{|\zeta_t^B|}\right)$$

As  $t \uparrow \infty$ ,  $|\zeta_t^B| \downarrow |\zeta_\infty^B|$ . From this it follows that

$$P(\xi_t^u(x) \equiv 0 \text{ on } B) \rightarrow E\left((1-u)^{|\zeta_\infty^B|}\right) \quad (2.1)$$

The probabilities on the left-hand side of (2.1) are enough to determine the distribution of the limit  $\xi_\infty^u$ . Since the limit exists, it is a stationary distribution that we denote by  $\nu_u$ .

Before moving on, we note that the duality equation can be written as

$$P(\xi_t^A \cap B \neq \emptyset) = P(A \cap \zeta_t^B \neq \emptyset) \quad (2.2)$$

where  $\xi_t^A$  is the voter model starting with 1's on  $A$  and  $\zeta_t^B$  is the coalescing random walk starting with particles on  $B$ . This holds because the left-hand side is the probability of a path from  $A \times \{0\}$  up to  $B \times \{t\}$ , while the right-hand side is the probability of a path from  $B \times \{t\}$  down to  $A \times \{0\}$ .

## 2.2 Nonlinear voter models

There are several types of duality. In general, see page 11 in [23],  $\eta_t$  and  $\zeta_t$  (with possibly different state spaces) are in duality with respect to the function  $H$  if

$$E^\eta H(\eta_t, \zeta) = E^\zeta H(\eta, \zeta_t)$$

The one we have introduced above corresponds to  $H(A, B) = 1$  if  $A \cap B \neq \emptyset$ , and 0 otherwise. Harris introduced this class of process [18] and showed that they were exactly the processes that could be constructed from a graphical representation of the type used in the previous section. since  $\xi_t^A$  is defined to be the set of sites at time  $t$  that can be reached from a path starting in  $A$ ,

$$\xi_t^{A \cup B} = \xi_t^A \cup \xi_t^B.$$

For this reason the duality is called **additive**.

**Claim.** *Using the graphical representation described in the previous section we cannot construct a voter model in which the flip rates depend only on the number of neighbors with the opposite opinion  $n_x$  and are nonlinear.*

*Proof.* For simplicity, we only prove the result when the neighborhood has size 4. Consulting Griffeath's book we see that the only gadgets that can be used in the graphical representation are combination of arrows and  $\delta$ 's. To begin, we will consider the set of processes that can be constructed by only using gadgets that have a  $\delta$  at  $x$  and a number of arrows that point to  $x$  from its neighbors. We call these objects arrow- $\delta$ s. Since the flip rates only depend on the number of sites, all arrow- $\delta$ s with  $k$  arrows have the same rate,  $a_k$ .

- When there is a 1 at  $x$  the  $\delta$  will cause the 1 to flip to a 0. However, the site will only stay a 0 if *all* neighbors connected to  $x$  by arrows are in state 0.
- When there is a 0 at  $x$  then the  $\delta$  does nothing, and the site will flip to 1 if there is *at least one* neighbor in state 1 connected to  $x$  by an arrow.

The number of  $k$ -arrow gadgets is  $\binom{k}{2}$  so the flip rates are as follows

| $n_x$ | rate $1 \rightarrow 0$     | rate $0 \rightarrow 1$     |
|-------|----------------------------|----------------------------|
| 0     | 0                          | 0                          |
| 1     | $a_1$                      | $a_1 + 3a_2 + 3a_3 + a_4$  |
| 2     | $2a_1 + a_2$               | $2a_1 + 5a_2 + 4a_3 + a_4$ |
| 3     | $3a_1 + 3a_2 + a_3$        | $3a_1 + 6a_2 + 4a_3 + a_4$ |
| 4     | $4a_1 + 6a_2 + 4a_3 + a_4$ | $4a_1 + 6a_2 + 4a_3 + a_4$ |

If we add  $\delta$ 's with no arrows then they will flip 1s even when all their neighbors are 1. If  $a_2$ ,  $a_3$ , or  $a_4$  is positive the rate of flipping  $1 \rightarrow 0$  is  $<$  the rate of flipping  $0 \rightarrow 1$  when  $n_x = 1, 2, 3$ . Adding arrows with no  $\delta$ s will only further increase the rates of flips  $0 \rightarrow 1$ .  $\square$

### 2.3 Duality for voter model perturbations

In the previous section we have shown that the  $q$ -voter does not have an additive dual. In this section we will introduce a generalization of the graphical representation used in Section 2.1 that allows us to construct voter model perturbations. This idea goes back to [10]. Calculating the state of the process is not as simple as in the additive case, but it does allow us to compute the state of the process on a finite set  $B$  at time  $t$  by working backwards from time  $t$ .

Voter model perturbations have flip rates

$$c_{i,j}^\delta(x, \xi) = f_j(x, \xi) + \delta^2 h_{i,j}(x, \xi), \quad (2.3)$$

where  $f_j(x, \xi)$  is the fraction of neighbors of  $x$  in state  $j$ . The perturbation function  $h_{i,j}$ ,  $j \neq i$ , may be negative (and this happens when  $q > 1$ ) but in order for the analysis in [10] to work, there must be a law  $q$  of  $(Y_1, \dots, Y_k) \in (\mathbb{Z}^d)^k$  and a functions  $g_{i,j} \geq 0$ , so that for some  $\gamma < \infty$ , we have

$$h_{i,j}(x, \xi) = -\gamma f_j(x, \xi) + E_Y[g_{i,j}(\xi(x + Y_1), \dots, \xi(x + Y_k))]. \quad (2.4)$$

In our situation  $Y_1, \dots, Y_k$  are  $k$  neighbors in  $\mathcal{N}$  and  $g_{i,j}$ , which does not depend on  $\epsilon$ , is the fraction of sites  $x + Y_1, \dots, x + Y_k$  in state  $j = 1 - i$  raised to the  $q$ th power.

Suppose now that we have a voter model perturbation of the form (2.3) which satisfies (2.4). We construct the voter model portion as in Section 2.1. We call the arrow- $\delta$ s **voter events**. To add the perturbation we let

$$\|g_{i,j}\| = \sup_{\eta \in \{0,1\}^k} g_{i,j}(\eta_1, \dots, \eta_k)$$

and introduce Poisson processes  $T_m^{x,i,j}$ ,  $m \geq 1$  with rate  $r_{i,j} = \epsilon \|g_{i,j}\|$ , where  $\epsilon = \delta^2$ , and independent random variables  $\{U_m^{x,i,j}\}_{m \geq 1}$ , uniform on  $(0, 1)$ . At the times  $t = T_m^{x,i,j}$  with  $m \geq 1$  we draw arrows from  $x + Y^i$  to  $x$  for  $1 \leq i \leq k$ . We call this a **branching event**. If  $\xi_{t-}(x) = i$  and

$$r_{i,j} U_k^{x,i,j} < g_{i,j}(\xi_{t-}(x + Y_1), \dots, \xi_{t-}(x + Y_k)) \quad (2.5)$$

then we set  $\xi_t(x) = j$ . The uniform random variables slow down the transition rate from the maximum possible rate  $r_{i,j}$  to the one appropriate for the current configuration.

To define the dual, we proceed as before. When a particle encounters a  $\delta$  associated with a voter event, it jumps to the other end of the arrow. When a particle encounters the head of an arrow associated with a branching event it gives birth to new particles at the other ends of all of the arrows. If either action results in two particles on the same site they coalesce to 1. Let  $I_s^{B,t}$  be the set of particles at time  $t - s$  when we start with particles on  $B$  at time  $t$ . Durrett and Neuhauser [14] called  $I_s^{B,t}$  the **influence set** because

**Lemma 2.1.** *If we know the values of  $\xi_{t-s}$  on  $I_s^{B,t}$ , then using the graphical representation (including the associated uniform random variables) we can compute the values of  $\xi_t$  in  $B$  by working our way up the graphical representation starting from time  $t - s$  and determining the changes that should be made in the configuration at each jump time.*

This fact should be clear from the construction. A formal proof can be found in Section 2.6 of [10]. The **computation process**, as it is called in [10], is complicated, but is useful because up to time  $t/\epsilon_n$  there will only be  $O(1)$  branching events affecting particles in the dual.

### 3 Prolonged persistence

In this section, we will prove Theorem 1.3. The key is to bound the difference between the density of the particle system and the ODE, using a result of Darling and Norris [11]. Section 3.1 describes this result and the work needed to apply it to finish the proof of Theorem 1.3. Sections 3.2, 3.3, 3.4, and 3.5 complete this work and Section 3.6 gives the final details.

#### 3.1 Darling-Norris theorem

To state the result from [11] result we need to introduce some notation. Let  $\xi_t$  be a continuous time Markov chain with countable state space  $S$  and jump rates  $q(\xi, \xi')$ . In our case  $\xi_t$  will be the state of the  $q$ -voter model on the torus. We are interested in proving an ODE limit for  $X_t = x(\xi_{t/\epsilon_n})$  where

$$x(\xi_{t/\epsilon_n}) = \frac{1}{n} \sum_{y \in T_n} \xi_{t/\epsilon_n}(y).$$

For each  $\xi \in S$  we define the infinitesimal drift

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

We let  $b$  be the drift of the proposed deterministic limit  $x_t$ . In our case

$$x_t = x_0 \pm \int_0^t b(x_s) ds, \quad b(y) = c_k y(1-y)(1-2y)f_k(y),$$

where  $f_k$  is a polynomial with  $f_k(0) = f_k(1) = 1$  that is positive on  $[0, 1]$  and only depends on the number of neighbors  $k$ . The sign is  $+$  for  $q = 1 - \epsilon_n$  and  $-$  for  $q = 1 + \epsilon_n$ . The crucial theorem from [11] is

**Theorem 3.1.** For each fixed  $t_0$  and  $\eta > 0$ ,

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

To make this statement meaningful we need more definitions. To measure the size of the jumps we let  $\sigma_\theta(y) = e^{\theta|y|} - 1 - \theta|y|$  and let

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

The good sets  $\Omega_i$ ,  $i = 0, 1, 2$  are given by

$$\Omega_0 = \{|X_0 - x_0| \leq \gamma\} \tag{3.1}$$

$$\Omega_1 = \left\{ \int_0^t |\beta(\xi_{s/\epsilon_n}) - b(X_s)| ds \leq \gamma \right\}, \tag{3.2}$$

$$\Omega_2 = \left\{ \int_0^t \phi(\xi_{s/\epsilon_n}, \theta) ds \leq \theta^2 At/2 \right\}. \tag{3.3}$$

The parameters in these events are coupled by the following relationships. If we let  $K$  be the Lipschitz constant of the drift  $b$  and  $\eta$  be the upper bound on the error in the approximation by the differential equation in Theorem 3.1, then

$$\gamma = \eta e^{-Kt_0}/3 \quad \text{and} \quad \theta = \gamma/(At_0), \quad \text{where } A > 0.$$

It is clear that  $b$  is Lipschitz continuous. Our assumption that  $U_n(0) \rightarrow u_0$  implies that  $\Omega_0^c = \emptyset$  for large  $n$ . To bound  $P(\Omega_2^c)$ , we will choose an  $A > 0$  that works well. We begin with a useful lemma:

**Lemma 3.2.** *If  $Z \sim \text{Poisson}(\lambda)$ , then*

$$P(Z \geq 2\lambda) \leq \exp(-\gamma(2)\lambda)$$

where  $\gamma(2)$  is a constant independent of  $\lambda$ .

*Proof.* The moment generating function of  $Z$  is

$$E \exp(\theta Z) \leq \exp(\lambda(e^\theta - 1)).$$

Taking  $\theta = \log 2$ , we have  $E \exp(Z \log 2) = \exp(\lambda)$ , so using Chebyshev's inequality we have

$$P(Z \geq 2\lambda) \leq \exp(-2 \log 2 \lambda), \exp(\lambda)$$

which proves the result with  $\gamma(2) = 2 \ln 2 - 1$ .  $\square$

The process  $X_t$  has jumps of size  $1/n$  at total rate  $n/\epsilon_n$ . As  $\theta|y| \rightarrow 0$ , we have  $\sigma_\theta(y) \sim \theta^2 y^2/2$ . So, when  $\theta|y|$  is small,  $\sigma_\theta(y) \sim \theta^2 y^2$ . Using Lemma 3.2, the probability of  $2t_0 n/\epsilon_n$  jumps during time  $[0, t_0]$  is  $\leq \exp(-\gamma(2)t_0 n/\epsilon_n)$ . When this occurs, and  $n$  is large, the integral in  $\Omega_2$  is

$$\leq \frac{\theta^2}{n^2} \cdot \frac{2t_0 n}{\epsilon_n} = \theta^2 t_0 \cdot \frac{2}{n\epsilon_n}.$$

Thus, for the event  $\Omega_2$  to hold, we need  $2/(n\epsilon_n) \ll A/2$ . Since  $\epsilon_n \sim Cn^{-a}$  with  $2/3 < a < 1$ , we have

**Lemma 3.3.** *If  $t_0$  and  $\gamma$  are fixed and  $A = n^{-(1-a)/3}$  then  $e^{-\gamma^2/(2At_0)} \rightarrow 0$  and  $P(\Omega_2^c) \rightarrow 0$  exponentially fast as  $n \rightarrow \infty$ .*

### 3.2 Ignoring branching

The remainder of Section 3 is devoted to bounding  $P(\Omega_1^c)$ . To begin to do this, we return to the original time scale. We define  $\tilde{\xi}_s$  to be the same as  $\xi_s$  at time  $s = t/\epsilon_n - n^b$ , while on the time interval  $[t/\epsilon_n - n^b, t/\epsilon_n]$ ,  $\tilde{\xi}_s$  only has voter events, ignoring the perturbation. The value  $b \in (2/3, a)$  is chosen so that lineages in the dual coalescing random walk will have time to wrap around the torus but, as we will now show, the perturbation will not have much effect. Let

$$\tilde{X}_t = \frac{1}{n} \sum_{x \in \mathbb{T}_n} \tilde{\xi}_{t/\epsilon_n}(x)$$

be the density of this new process  $\tilde{\xi}$ .

We will now show that ignoring the perturbation changes the values of more than  $\eta n$  sites with a stretched exponentially small probability.

**Step 1.** The number of perturbation events  $M$  in time  $n^b$  is bounded by a  $\text{Poisson}(\lambda)$  random variable with  $\lambda = Cn^{1+b+a}$ . Lemma 3.2 implies that

$$P(M \geq 2\lambda) \leq \exp(-\gamma(2)\lambda) \leq \exp(-C\gamma(2)n^b), \quad (3.4)$$

since  $\lambda \geq Cn^b$ .

**Step 2.** Let  $\eta_t(x) = |\xi_t(x) - \tilde{\xi}_t(x)|$ , so that  $\eta_t(x) = 1$  means there is a discrepancy between the two processes  $\xi_t$  and  $\tilde{\xi}_t$  at position  $x$ . We want to prove that  $\sum_x \eta_{t/\epsilon_n}(x)$  is more than  $\eta n$  with a stretched exponentially small probability. To do this, note that when an edge  $(x, y)$  with  $\eta_s(x) = 0$  and  $\eta_s(y) = 1$  is hit by a voter event (that is, there is an arrival in the Poisson process  $T^{x,y}$  or  $T^{y,x}$ ), then the 1 is changed to 0 with probability 1/2 (when the arrival is in  $T^{x,y}$ ) and the 0 is changed to a 1 with probability 1/2 (when the arrival

is in  $T^{y,x}$ ). Thus, the change in the number of discrepancies due to voter events is a martingale. Let  $Y_m$  denote this value after  $m$  jumps. The change is always  $\leq 1$  so if there are  $N$  jumps, then by Azuma's inequality

$$P(|Y_N - Y_0| \geq z \mid N = n_0) \leq 2 \exp(-z^2/2n_0)$$

If  $N$  is the number of changes due to voter events in the time interval  $[t/\epsilon_n - n^b, t/\epsilon_n]$ , then  $N \leq \text{Poisson}(n^{b+1})$ . By Lemma 3.2,

$$P(N \geq 2n^{1+b}) \leq \exp(-\gamma(2)n^{1+b}).$$

Note that if  $n_0 < 2n^{1+b}$ , then  $2 \exp(-z^2/2n_0) < 2 \exp(-z^2/4n^{1+b})$ . So, taking  $z = \eta n$  and  $N = 2n^{1+b}$ , we get

$$P(|Y_N - Y_0| \geq \eta n) \leq 2 \exp(-\eta^2 n^{1+b}/4). \quad (3.5)$$

### 3.3 Bounding the density

The results in the previous section show that on the interval  $[t/\epsilon_n - n^b, t/\epsilon_n]$  we can ignore the perturbation and assume that the process evolves like the voter model. To understand the distribution of 1's at time  $t/\epsilon_n$  we will use results of Bramson and Griffeath [4], and Zähle [30]. The first reference only treats  $d = 3$ . The second covers  $d \geq 3$  and is more detailed, so we will follow it.

Let  $\zeta^\lambda : \mathbb{Z}^d \rightarrow \{0, 1\}$  have the distribution of the equilibrium of a finite range voter model on  $\mathbb{Z}^d$  with density  $\nu_\lambda$ . For an explanation of this and the other basic facts about the voter model that we will use, see Liggett's book [23]. For simplicity we will do calculations for the nearest neighbor case. The results are the same in the finite range case, but are more awkward to write since, for example, the limiting normal has a general covariance matrix, we cannot use the reflection principle, etc. To formulate the limit theorem in [30], we will write the process at a fixed time as a random field

$$F_\lambda(\phi) = \sum_{i \in \mathbb{Z}^d} [\zeta^\lambda(i) - \lambda] \phi(i),$$

where  $\phi$  is a member of a suitable class of test functions. To rescale space, we let

$$F_{\lambda,r}(\phi) = F_\lambda(\phi_r) \quad \text{where} \quad \phi_r(x) = r^{-(d+2)/2} \phi(x/r).$$

Theorem 1 on pages 1265–1266 of [30] shows that in our nearest neighbor case

$$F_{\lambda,r}(\phi) \Rightarrow \text{Normal}(0, a_d \lambda(1 - \lambda) B(\phi, \phi)),$$

where  $\Rightarrow$  denotes weak convergence as  $r \rightarrow \infty$ ,  $\text{Normal}(\mu, \sigma^2)$  is a one-dimensional normal distribution with mean  $\mu$  and variance  $\sigma^2$ , and  $B$  is the bilinear function

$$B(\phi, \psi) = \iint \frac{\phi(x)\psi(y)}{|x - y|^{(d-2)/2}} dx dy.$$

Restricting our attention now to  $d = 3$ , Zähle's result implies that

$$\hat{S}_r \equiv [\lambda(1 - \lambda)]^{-1/2} r^{-5/2} \sum_{x \in [-r/2, r/2]^3} [\zeta^\lambda(x) - \lambda] \Rightarrow \text{Normal}(0, c_{3,\lambda}) \quad (3.6)$$

Bramson and Griffeath [4] prove (3.6) by the method of moments, which gives

$$E(\hat{S}_r)^{2m} \rightarrow c_{3,\lambda}^{2m} \mu_m \quad \text{where} \quad \mu_m = (2m - 1)(2m - 3) \cdots 3 \cdot 1. \quad (3.7)$$

In our situation, we need a slightly different result. In particular, these results are for the voter model on  $\mathbb{Z}^3$ , and we need a result for the voter model on the 3-d torus. Let

$$T_r \equiv [\lambda(1 - \lambda)]^{-1/2} r^{-5/2} \sum_{x \in Q(r)} [\xi_{t/\epsilon_n}(x) - \lambda]$$

where  $\lambda$  is the fraction of sites in state 1 at time  $t/\epsilon_n - n^b$ , and  $Q(r)$  is a fixed cube with side  $r = n^\beta$  with  $\beta < 1/3$ . To prove a limit result for  $T_r$  we will sandwich it between  $\bar{S}_r$  and

$$\bar{S}_r \equiv [\lambda(1 - \lambda)]^{-1/2} r^{-5/2} \sum_{x \in Q(r)} [\bar{\zeta}_{\sigma(n)}^\lambda(x) - \lambda],$$

where  $\bar{\zeta}_{\sigma(n)}^\lambda$  is the voter model on the torus starting from product measure with density  $\lambda$  and run for time  $\sigma(n) = n^{0.6}$ . To couple this with  $T_r$  we create  $\bar{S}_r$  by running coalescing random walks starting at time  $t/\epsilon_n$  from points in  $Q(r)$  backwards in time for  $\sigma(n)$ , and then use independent coin flips with probability  $\lambda$  of heads (1) and  $1 - \lambda$  of tails (0) to determine the states of the sites.

**(i) With stretched exponentially small probability, no coalescing random walk will move more than  $n^{0.33}$  in any coordinate by time  $\sigma(n) = n^{0.6}$ .**

*Proof.* We will use a special case of (7.3) on page 553 in Feller volume II [15].

**Lemma 3.4.** *Let  $w_1, w_2, \dots, w_k$  be i.i.d. with  $P(w_i = 1) = P(w_i = -1)$ . Then if  $W_k = w_1 + \dots + w_k$ ,  $\epsilon > 0$ , and  $x = o(k)$ , we have*

$$P(W_k/\sqrt{k} \geq x) \leq \exp(-(1 - \epsilon)x^2/2).$$

Taking  $k = n^{0.6}$  and  $x = 0.03$ , it follows that the probability some coalescing random walk starting inside the cube  $Q(r)$  and run for time  $\sigma(n)$  moves by more than  $n^{0.33}$  in any coordinate is

$$\leq 2 \cdot 6r^3 \exp(-(1 - \epsilon)n^{0.06}/2).$$

Here the 2 comes from using the reflection principle to relate the maximum to the value at time  $n^{0.6}$ , and 6 is 3 coordinates times 2 signs.  $\square$

The result (i) implies that with very high probability there is no difference between the coalescing starting from  $Q(r)$  with  $r = n^\beta$  for  $\beta < 1/3$ , run to time  $\sigma(n) = n^{0.6}$  on the torus or on  $\mathbb{Z}^3$ .

**(ii) There is a  $\gamma > 0$  so that at all times  $t \geq (k + 1)n^{2/3}$ , the total variation between the distribution of a nearest neighbor random walk on the torus and the uniform distribution is  $\leq (1 - \gamma)^k$ .**

*Proof.* To prove the result, we use a simple coupling. At time  $n^{2/3}$  the distribution of each particle has a density that is  $\geq \gamma/n$  at each point of the torus. At time  $n^{2/3}$  the distribution has the form  $\gamma \cdot \mu_n + (1 - \gamma)q_n$ , where  $\mu_n$  is uniform on the torus and  $q_n$  is some transition probability. Uncoupled mass at time  $(k - 1)n^{2/3}$  can be coupled to the uniform distribution with probability  $\geq \gamma$  at time  $kn^{2/3}$  and the desired result follows.  $\square$

**Definition of  $T'_n$ .** We continue the construction of  $T_r$ : from the end of the construction of  $\bar{S}_r$  at time  $\sigma(n)$ , we run the coalescing random walk particles on  $\mathbb{Z}^3$ . To assign values to the lineages at time  $n^b$  we extend the configuration on the torus at that time to be periodic on  $\mathbb{Z}^3$ . It follows from (ii) that with very high probability there is no difference between flipping coins at time  $n^{0.6}$  to determine the states of the sites in the sum  $\bar{S}_n$  or continuing to run the coalescing random walks on  $\mathbb{Z}^3$  until time  $n^b$ . Having done this, we

no longer perfectly reproduce  $T_n$ , so we call the result  $T'_n$ . The good news is that when we run the coalescing random walk on  $\mathbb{Z}^3$  starting at  $\sigma(n)$ , we will have  $T'_r \prec \hat{S}_n$ . That is, the coalescing random walk clusters in  $T'_r$  are contained in clusters in  $\hat{S}_r$ .

**To prove the result in (3.6)**, Zähle defines a *cluster* to be a set of sites that coalesce to the same limiting particle, and lets  $Z_{r,k}$ ,  $1 \leq k \leq K(r)$  be the cluster sizes and lets  $\eta_{r,k}$  be independently  $= 1$  with probability  $\lambda$  and  $= 0$  with probability  $1 - \lambda$ . As she notes in (3.6) on page 1274,

$$\hat{S}_r =_d r^{-5/2} \sum_{k=1}^{K(r)} Z_{r,k} \cdot (\eta_{r,k} - \lambda) \quad (3.8)$$

If we condition on the  $Z_{r,k}$ , then we have a sum of independent random variables. If we let  $v_n^2 = \sum_k Z_{n,k}^2$ , then using Lyapunov's theorem (see the bottom of page 1275) it follows that

$$(\hat{S}_r/v_n \mid \mathcal{Z}) \Rightarrow \chi,$$

where  $\mathcal{Z}$  is the  $\sigma$ -field generated by the  $Z_{r,k}$  and  $\chi$  is a standard normal. In Lemma 1 on page 1276 in [30] she shows that  $v_n^2$  converges in probability to a constant, so if we remove the conditioning we get the same limit. Lemma 2 computes the limit of  $Ev_n^2$  and (3.6) follows.

**The last argument can be applied to  $\bar{S}_n$**  to conclude that it converges to a normal distribution. To find the limiting variance we compute

$$\sum_{x,y \in Q(r)} E(\bar{\zeta}_{\sigma(n)}^\lambda(x) - \lambda)(\bar{\zeta}_{\sigma(n)}^\lambda(y) - \lambda)$$

When the coalescing random walks starting from  $x$  and  $y$  do not coalesce, the states at  $x$  and  $y$  are independent; otherwise, they are equal. Thus, if we let  $\tau_{x,y}$  be the time the two coalescing random walks hit, then the above sum is

$$\sum_{x,y \in Q(r)} \lambda(1 - \lambda)P(\tau_{x,y} \leq n^{0.6})$$

Using the local central limit theorem,

$$P(n^{0.6} \leq \tau_{x,y} < \infty) \approx 2\beta_d \int_{n^{0.6}}^{\infty} \frac{1}{(2\pi t)^{3/2}} \exp(-|x - y|^2/2t) dt$$

The right-hand side gives the expected amount of time the two particles spend together. When they hit they spend an exponential rate 2 amount of time together. In addition, they will hit a geometric number of times with success probability  $\beta_d$ . Changing variables  $t = |x - y|^2/2s$ ,  $dt = -|x - y|^2/(2s^2)$  the integral becomes

$$\begin{aligned} & \int_0^{|x-y|^2/n^{0.6}} \left( \frac{s}{\pi|x-y|^2} \right)^{3/2} e^{-s} \left( \frac{\pi|x-y|^2}{2s^2} \right) ds \\ &= \frac{1}{2\pi^{3/2}|x-y|} \int_0^{|x-y|^2/n^{0.6}} s^{-1/2} e^{-s} ds \leq Cn^{-0.3}. \end{aligned}$$

Consulting Lemma 4 in [30] we find

$$P(\tau_{x,y} < \infty) \sim c'_3/|x - y|$$

Using the formula for  $c'_3$  it follows that the asymptotic variance for  $\bar{S}_r$  is the same as for  $\hat{S}_r$ .

**Limit theorem for  $T'_r$ .** Let  $X_{r,k} \prec Y'_{r,k} \prec Z_{r,k}$  be the cluster sizes in  $\bar{S}_r$ ,  $T'_r$ , and  $\hat{S}_r$ . The limiting variances of the unnormalized sums are

$$\sum_k EX_{r,k}^2 \leq \sum_k E(Y'_{r,k})^2 \leq \sum_k EZ_{r,k}^2$$

Since the top and bottom sums have the same asymptotics, this gives us the Gaussian limit theorem for  $T'_n$ . Replacing 2 by  $2m$  and recalling that Bramson and Griffeath [4] proved their result for  $\hat{S}_r$  by the method of moments gives the desired results for  $T'_r$ :

$$T'_r \equiv [\lambda(1-\lambda)]^{-1/2} r^{-5/2} \sum_{k=1}^{K(r)} Y'_{r,k} (\eta_{r,k} - \lambda) \Rightarrow \mathcal{N}(0, c_{3,\lambda}) \quad (3.9)$$

$$E(T'_r)^{2m} \rightarrow c_{3,\lambda}^{2m} (2m-1)(2m-3) \cdots 3 \cdot 1 \quad (3.10)$$

The last result implies

$$r^{2m\beta} P(|T'_r| \geq r^\beta) \leq E(\bar{S}_r)^{2m} \rightarrow E\chi^{2m} \quad (3.11)$$

so if we let  $\tilde{D}'_r = [\lambda(1-\lambda)]^{1/2} r^{5/2} T'_r$ , (i.e., we remove the scaling) then

$$P(|D'_r| \geq [\lambda(1-\lambda)]^{1/2} r^{5/2+\beta}) \leq C_m r^{-2m\beta}. \quad (3.12)$$

This is the concentration result we desired for  $T'_n$ . Recall that  $T'_n$  was constructed as a slight modification of  $T_n$ , which is the true rescaled and centered density that we which to prove results about.

### 3.4 Controlling the difference between $T'_n$ and $T_n$

The goal in this section is to generalize (3.12) to  $T_r$ .

**Bounding the number of extra coalescences in  $T'_n$ .** When we went from the torus to  $\mathbb{Z}^3$  we may have eliminated some coalescence in  $T_n$  at times in  $[n^{0.6}, n^b]$ . For this to happen the difference in two particles positions must have wrapped around the torus, an event we call  $G$ , and the particles projected back to the torus must have hit, an event we call  $H$ . To bound this event we note that

$$P(G \cap H) \leq \min\{P(G), P(H)\}.$$

Let  $\alpha = 2(1-\epsilon)/3$ . Lemma 3.4 implies that the probability  $G$  happens during  $[n^{0.6}, n^\alpha]$  is  $\leq \exp(-n^\eta)$  for some  $\eta > 0$ . On  $[n^\alpha, n^b]$ , the probability that a random walk is at a fixed site is  $\leq 1/n^{1-\epsilon}$ . Thus, for a fixed pair of particles,

$$P(H) \leq Cn^b/n^{1-\epsilon}.$$

If  $r = n^{b(2)/3}$ , then  $n^{b(2)}$  is a trivial upper bound for the number of particles at time  $\sigma(n)$ , which holds with probability 1. We will now estimate the number of collisions of a fixed particle with all of the others. This number is increased if we ignore coalescence, and run the particles as independent. We do this so that

**Lemma 3.5.** *If  $m \geq 1$  and a particle belongs to a cluster of size  $2m$  or  $2m+1$  with  $m \geq 1$  formed by coalescence during  $[n^\alpha, n^b]$ , then there are at least  $m$  disjoint pairs of particles that have coalesced.*

*Proof.* Recall that on this time interval we are running the lineages on  $\mathbb{Z}^3$ . We will prove the result by induction. To be able to disentangle the graph constructed by coalescence we will number the particles. Once two particles hit the two future trajectories could

be assigned to either particle so we allow ourselves the liberty of exchanging the labels at any collision. If the cluster has size 2 or 3, this is trivial. Suppose now that  $m \geq 2$ . Locate the time  $t_0$  at which the first two particles coalesced. Call them  $x$  and  $y$  and let  $t_1$  be the first time after  $t_0$  that the coalesced particle collided with another one that we call  $z$ . Remove the  $Y$ -shaped part of the genealogy leading from  $x$  and  $y$  to the coalescence at time  $t_1$ . Label the lineage coming out  $t_1$  the same as the one coming in on  $z$ 's trajectory. We have identified one pair of coalescing particles and reduced the number of sites in the cluster by 2, so the result follows by induction.  $\square$

Given Lemma 3.5, our next task is to estimate the probability that  $m$  disjoint pairs will coalesce. Using the trivial upper bound  $n^{b(2)}$  on the number of lineages, the number of coalescing pairs is

$$N \leq \text{Binomial}(n^{2b(2)}, Cn^b/n^{1-\epsilon}).$$

Note that this bounds the number of coalescing pairs that coalesce in the system, not just those that form one cluster. The expected number is  $Cn^{b+2b(2)+\epsilon-1}$ , where  $b$  is larger than  $2/3$  and can be assumed to be  $\leq 0.7$ . If  $b(2) \leq 0.1$ , then  $-\nu = b + 2b(2) + \epsilon - 1 < 0$  when  $\epsilon < 0.5$ . In this case,

$$P(N = k) \leq \binom{n^{2b(2)}}{k} (Cn^{b+\epsilon-1})^k \leq \frac{C^k n^{-k\nu}}{k!},$$

so summing gives

$$P(N \geq k) \leq e^C n^{-k\nu}. \quad (3.13)$$

**Bounding the size of clusters in  $\hat{S}_r$ .** Formula (3.7) tells us that

$$E(\hat{S}_r)^{2m} \rightarrow c_{3,\lambda}^{2m} \mu_m$$

Using (3.8) we have  $[(1-\lambda)\lambda^{2m} + \lambda(1-\lambda)^{2m}] \sum_{k=1}^{K(r)} Z_{r,k}^{2m} \leq E(\hat{S}_r)^{2m}$ . From this we see that when  $r$  is large

$$r^{11m/2} P(\max_k Z_{r,k} \geq r^{5.5/2}) \leq C_m r^{5m}$$

so we have

$$P(\max_k Z_{r,k} \geq r^{5.5/2}) \leq C_{m,\lambda} r^{-m/2}. \quad (3.14)$$

Combining (3.13) and (3.14) we see that if  $Y_{r,k}$  are cluster sizes in  $T_n$ , then

$$P\left(\max_k Y_{r,k} \geq \frac{m}{2\nu} r^{5.5/2}\right) \leq C_{m,\lambda} r^{-m/2} \quad (3.15)$$

Combining (3.13) with  $k = m/2\nu$  and (3.15) we see that the combined size of the clusters in  $T'_n$  but not in  $T_n$  is

$$\leq \frac{m}{2\nu} n^{5.5/2} \quad \text{with probability } 1 - C_{m,\lambda} r^{-m/2}. \quad (3.16)$$

Using this with (3.12) and letting  $D_r = [\lambda(1-\lambda)]^{1/2} r^{5/2} T_n$  it follows that

$$P(|D_r| \geq [\lambda(1-\lambda)]^{1/2} r^{5/2+\beta}) \leq C_{m,\lambda} r^{-m\beta/2}. \quad (3.17)$$

Suppose  $r = n^{b(2)/3}$  where  $0 < b(2) < 1$ , then

$$P\left(|D_r| \geq [\lambda(1-\lambda)]^{1/2} n^{5b(2)/6+\beta}\right) \leq C_m n^{-m\beta b(2)/6}$$

Now, partition the torus into cubes of side  $n^{b(2)/3}$ . Letting  $N_i$  be the number of 1's in the  $i$ th cube we have

$$P\left(\left|\frac{N_{r,i}}{n^{b(2)}} - \lambda\right| \geq [\lambda(1-\lambda)]^{1/2} n^{-b(2)/6+\beta}\right) \leq C_m n^{-m\beta b(2)/6}.$$

For fixed  $\beta > 0$ , given a  $k < \infty$  we can pick  $m$  large enough then the right hand side is  $\leq n^{-(1-b(2))-k}$ . Then we have,

$$P\left(\text{for some } i \left|\frac{N_i}{n^{b(2)}} - \lambda\right| \geq [\lambda(1-\lambda)]^{1/2} n^{-b(2)/6+\beta}\right) \leq n^{-k}. \quad (3.18)$$

### 3.5 Bounding the difference in the drifts

For  $x, y_1, \dots, y_k \in \mathbb{Z}^d$  and  $v_0, v_1, \dots, v_k \in \{0, 1\}$  fixed we let

$$G_{x,y,v} = \{\xi(x) = v_0, \xi(x+y_1) = v_1, \dots, \xi(x+y_k) = v_k\}$$

be a finite dimensional event. For simplicity, we do not display the dependence on the sites  $y$  and the states  $i$ . The probabilities of these events for  $\xi_t$  are the *empirical finite dimensional distributions*.

Thus far we have been concerned with the overall density of particles on the torus. However, to successfully bound  $P(\Omega_1^c)$  we need to show that if  $u$  is the density of ones in the voter model at time  $t/\epsilon_n - n^b$ , then the empirical finite dimensional distributions on the torus are close to those of the voter model equilibrium  $\nu_u$  at time  $t/\epsilon_n + s_n$ , where

$$s_n = n^{(2+\beta)b(2)/3}. \quad (3.19)$$

The reasoning for introducing this extra time  $s_n$  is described below.

The first step is to partition the torus at time  $t/\epsilon_n$  into boxes with side  $r = n^{b(2)/3}$ . Using (3.18), we can conclude that with high probability the density in each box is close to  $u$ , the density of 1's at time  $t/\epsilon_n - n^b$ . We divide the torus at time  $t/\epsilon_n + s_n$  into cubes with side  $n^{b(3)/3}$ , where  $b(3) > b(2)$ . The  $\beta$  in the time guarantees that if we work backwards from time  $t/\epsilon_n + s_n$  to  $t/\epsilon_n$ , the probability a random walk particle will move by an amount much larger than  $n^{b(2)/3}$ , the size of the boxes at time  $t/\epsilon_n$ , is stretched exponentially small. See Lemma 3.4. As in [14] and [10] this implies the conditional distribution of the position given that the lineage ends in a specific box is almost uniform, and hence the probability it lands on a 1 will be close to  $u$ . A second consequence is that

**Lemma 3.6.** *With very high probability, the empirical finite dimensional distributions at time  $t/\epsilon_n + s_n$  will be close to  $\nu_u(G_{x,y,v})$ .*

*Proof.* To see this, note that we compute the probabilities of finite dimensional sets in the voter model equilibrium  $\nu_u$  by starting the CRW with points at  $y_0, \dots, y_m$ , and running time to  $s_n$ . The particles that coalesce are a partition of the original set. We then flip a coin with a probability  $u$  of heads (state 1) to determine the states. Here we are only running time to  $s_n$  so our partition is finer, but the final particles are roughly independent and uniform on the torus so whether they land on 1 or 0 are roughly independent coin flips.  $\square$

The last paragraph shows that probabilities of the f.d.d.'s are close to the voter model equilibrium  $\nu_u$ . This enables us to conclude that the expected value of the drift of our process when the density is  $x$  is close to  $b(x)$ . The next step is control the fluctuations about the mean. Using normal tail bounds on random walks in Lemma 3.4, it follows that

if  $B_n$  is the event that some coalescing random walk at time  $t/\epsilon_n + s_n$  moves by more than  $n^{b(3)/3}$  in time  $s_n$ , then for any  $\gamma > 0$  we have for large  $n$

$$\begin{aligned} P(B_n) &\leq n \exp(-(1-\gamma)n^{2b(3)/3}/2n^{(2+\beta)b(2)/3}) \\ &= n \exp\left(-\frac{1-\gamma}{2}n^{[2b(3)-(2+\beta)b(2)]/3}\right) \end{aligned} \quad (3.20)$$

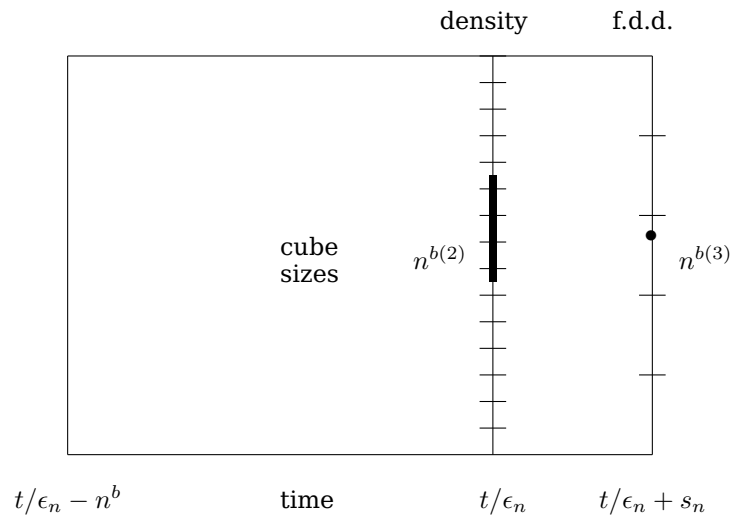


Figure 5: Picture summarizing the proof. Here  $s_n = n^{(2+\beta)b(2)/3}$ . The words at the top indicate the quantity that is “good” at each time, i.e., close to its average value on the cubes. The dark line at time  $t/\epsilon_n$  shows the interval in which we will with high probability find the lineage of the black dot when it is worked backwards in time.

For the last inequality to be useful we need to choose  $\beta$  so that  $2b(3) - (2 + \beta)b(2) > 0$ . The estimate in (3.20) implies that the states of sites in cubes in the decomposition at time  $t/\epsilon_n + s_n$  that do not touch are independent on  $B_n^c$ . We can divide our collection of cubes into 27 subcollections  $\mathcal{C}_i$  of size  $n^{1-b(3)}/27$  so that no two cubes in the subcollection touch. For  $1 \leq i \leq 27$ , let  $N_i$  be the number of times  $G_{x,y,v}$  occurs in the union of the cubes in  $\mathcal{C}_i$ , let  $N_{i,j}$  be the number of times  $G_{x,y,v}$  occurs for  $x$  in the  $j$ th cube in  $\mathcal{C}_i$ . If  $x$  is close to the edge of the cube then some of the  $x + y_i$  may be outside. However, the  $y_i$  are fixed, so for large  $n$  they will at worst be in an adjacent cube.

For fixed  $i$ , the  $N_{i,j}$  are independent on the event  $B_n^c$ , and  $0 \leq N_{i,j}/n^{b(3)} \leq 1$ . Let  $\rho_{i,j} = EN_{i,j}/n^{b(3)}$ . Let

$$X_{i,j} = \frac{N_{i,j}}{n^{b(3)}} - \rho_{i,j} \in [-\rho_{i,j}, 1 - \rho_{i,j}].$$

Finally, let  $\psi_{i,j}(\theta) = E \exp(\theta X_{i,j})$ , let  $Y_i = \sum_j X_{i,j}$ , and let  $M = n^{1-b(3)}/27$  be the number of cubes in each collection  $\mathcal{C}_i$ . If  $\theta > 0$ , then, assuming  $B_n^c$ , we have

$$e^{\theta M \eta} P(Y_i \geq M \eta) \leq \prod_j \psi_{i,j}(\theta),$$

using the independence of the  $N_{i,j}$  across  $j$ . So, we have

$$\begin{aligned} P(Y_i \geq M\eta) &\leq e^{-\theta M\eta} \prod_j \psi_{i,j}(\theta) \\ &= \exp \left( M \left[ -\theta\eta + M^{-1} \sum_j \log \psi_{i,j}(\theta) \right] \right) \end{aligned} \quad (3.21)$$

Since we do not know much about  $\psi_{i,j}(\theta)$ , we will let  $\eta_n = n^{-\alpha}$ , and later choose  $\theta_n$  so that  $\lim_{n \rightarrow \infty} \theta_n = 0$ . Expanding  $\log \psi_{i,j}$  around 0:

$$\begin{aligned} \frac{d}{d\theta} \log \psi_{i,j}(\theta) &= \frac{\psi'_{i,j}(\theta)}{\psi_{i,j}(\theta)}, \\ \frac{d^2}{d\theta^2} \log \psi_{i,j}(\theta) &= \frac{\psi''_{i,j}(\theta)}{\psi_{i,j}(\theta)} - \frac{(\psi'_{i,j}(\theta))^2}{\psi_{i,j}^2(\theta)}. \end{aligned}$$

When  $\theta = 0$ , we have  $\psi_{i,j}(0) = 1$  by definition, and also

$$\begin{aligned} \frac{d}{d\theta} \log \psi_{i,j}(0) &= EX_{i,j} = 0, \\ \frac{d^2}{d\theta^2} \log \psi_{i,j}(0) &= EX_{i,j}^2. \end{aligned}$$

So, if  $\theta_{i,n} \rightarrow 0$ , then we have the approximation

$$\log \psi_{i,j}(\theta_{i,n}) \sim \frac{\theta_{i,n}^2}{2} EX_{i,j}^2.$$

Since  $X_{i,j} \in [-\rho_{i,j}, 1 - \rho_{i,j}]$  and  $EX_{i,j} = 0$ ,

$$EX_{i,j}^2 \leq \rho_{i,j}(1 - \rho_{i,j})$$

To optimize the bound in (3.21) we  $d/d\theta$  the term in square brackets in (3.21) to get

$$0 = -\eta_n + \theta_{i,n} M^{-1} \sum_j \rho_{i,j}(1 - \rho_{i,j}), \quad (3.22)$$

which says we want to take  $\theta_n = \eta_n/\tau_i$ , where  $\tau_i = M^{-1} \sum_j \rho_{i,j}(1 - \rho_{i,j})$ . This gives the following large deviations bound

$$\begin{aligned} P(Y_i \geq M\eta_n) &\leq \exp \left( M \left[ -\frac{\eta_n^2}{\tau_i} + \frac{\eta_n^2}{2\tau_i^2} \tau_i \right] \right) \\ &= \exp \left( -\frac{M\eta_n^2}{2\tau_i} \right) \leq \exp(-M\eta_n^2), \end{aligned}$$

since  $2\tau_i \leq 1$ . The same reasoning can be used to get a bound on the other deviation. Since we have expanded the moment generating function around 0 the bound is the same, giving the final result

$$P(|Y_i| \geq M\eta_n) \leq \exp(-M\eta_n^2)$$

Define  $Y = \sum_{i=1}^{27} Y_i$ , and then use the triangle inequality to get

$$P(|Y| \geq 27M\eta_n) \leq 27 \exp(-M\eta_n^2)$$

The last task is to relate this to the difference of the drifts. To do this, we note that

$$Y = n^{-b(3)} \sum_{i,j} N_{i,j} - \sum_{i,j} \rho_{i,j}$$

so we have

$$\frac{Y}{n^{1-b(3)}} = n^{-1} \sum_x 1(G_{x,y,v}) - \frac{1}{n^{1-b(3)}} \sum_{i,j} \rho_{i,j}$$

Let  $p_{x,y,v}^n$  be the probability of  $G_{x,y,v}$  when we work backwards in the coalescing random walk starting from  $x, x + y_1, \dots, x + y_k$  then we have

$$\frac{1}{n^{1-b(3)}} \sum_{i,j} \rho_{i,j} = \frac{1}{n} \sum_x p_{x,y,v}^n$$

In the three neighbor case we only have to consider:  $y_1 = e_1$ ,  $y_2 = e_2$ , and  $y_3 = e_3$ . When there are more neighbors, we have to consider a number of other possibilities, see the calculations in Section 5. Let  $r(v) = r(v_0, v_1, v_2, v_3)$  be the jump rate of vertex  $x$  when the states are  $v_i$ . Multiplying by  $r(v)$ , summing over the relevant values of  $y, v$  we have

$$n^{-1} \sum_{x,y,v} 1(G_{x,y,v}) r(v) = \beta(\xi_{t/\epsilon_n + s_n})$$

so we have

$$P\left(\left|\beta(\xi_{t/\epsilon_n + s_n}) - \frac{1}{n} \sum_{x,y,v} p_{x,y,v}^n r(v)\right| \geq 16n^{-\alpha}\right) \leq 27 \exp\left(-n^{1-b(3)-2\alpha}/27\right) \quad (3.23)$$

The choice of  $s_n$  guarantees that as we work backwards in time the particles in the CRW move by an amount  $\gg n^{b(3)}$ . The bound in (3.18) implies that each particle in the CRW lands on a 1 with probability close to  $u$ . It follows that

$$\left|\frac{1}{n} \sum_{y,v} p_{x,y,v}^n - b(u)\right| \leq \eta/2$$

with very high probability. The bounds derived above only works for fixed  $t$ . However, it is easy to extend them so that they hold uniformly on  $[0, t_0]$  and hence are valid for the integral. To do this, we subdivide the interval into subintervals of length  $1/n^{1/2}\epsilon_n$ . Within each interval the probability there are more than  $2n^{1/2}$  flips is  $\leq \exp(-c\sqrt{n})$ . If we add this to previous error probability and multiply by the number of subinterval we still have a result that holds with very high probability.

### 3.6 Final details

To get long time survival, we will iterate. Let

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

and note that  $x_t$  is the solution of the ODE so this is not random. Theorem 3.1 implies that  $|X(T_0) - 1/2| \leq 2\eta$  with very high probability. Let

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

and note that on  $[T_0, T_1]$  we have  $|X_t - 1/2| \leq 4\eta$ . There is a constant  $t_\eta$  so that if  $x(0) = 1/2 + 4\eta$  or  $x(0) = 1/2 - 4\eta$  then  $|x(t_\eta) - 1/2| \leq \eta$ . Let  $S_1 = T_1 + t_\eta$ . Since  $T_1$  is random,  $S_1$  is a random time. However, due to the Markov process, we can translate time

to apply Theorem 3.1 again. That is, consider  $\tilde{X}_t := X_{t+T_1}$ . Then since  $|\tilde{X}_0 - 1/2| = 4\eta$ , Theorem 3.1 implies that with high probability  $|\tilde{X}_{t_\eta} - 1/2| = |X(S_1) - 1/2| \leq 2\eta$  and  $|X_t - 1/2| \leq 5\eta$  on  $[T_1, S_1]$ . For  $m \geq 2$ , let

$$T_m = \inf\{t > S_{m-1} : |X_t - 1/2| \geq 4\eta\} \quad \text{and} \quad S_m = T_m + t_\eta.$$

We can with high probability iterate the construction  $n^k$  times before it fails. Since each cycle takes at least  $t_0$  units of time, taking  $\eta = \gamma/5$  the proof of Theorem 1.3 is complete.

## 4 Rapid extinction for $q > 1$

In this section we will prove Theorem 1.4. There are two steps to the proof. First, we use the results in Section 4 to show that the fraction of 1's in the random process is close to solution of the ODE until time

$$\tau = \min\{t : x_t < n^{-(1-b(0))}\}, \quad (4.1)$$

where  $b(0)$  will be defined in the proof of Lemma 4.1. The second step is to prove that when we start with  $\leq n^{b(0)}$  ones, then fluctuations in the voter model will cause it to hit 0 in time  $\leq Cn^{b(0)}$ . This time is  $< n^b$  for large  $n$ , so by results in Section 3.2, it is legitimate to assume that the process acts like the voter model. The proof for the second step is based on a Green's function calculation and estimates for the rate of change of the number of ones in the voter model.

### 4.1 First step

**Lemma 4.1.** *Suppose  $X_0 < 1/2$  and let  $\tau$  be defined in (4.1). Then, for any  $\eta > 0$ , as  $n \rightarrow \infty$ ,*

$$\mathbb{P}\left(|X_\tau - n^{-(1-b(0))}| < \eta n^{-(1-b(0))}\right) \rightarrow 1.$$

*Proof.* We use (3.18) from Section 3.4. If  $X_0 = u$  and we divide the torus at time  $t/\epsilon_n$  into boxes of side  $r = n^{b(2)/3}$ , then taking  $m$  large in (3.18) gives

$$P\left(\text{for some } i \left| \frac{N_{r,i}}{n^{b(2)}} - u \right| \geq [u(1-u)]^{1/2} n^{-b(2)/6+\beta}\right) \leq n^{-k}, \quad (4.2)$$

for any  $\beta > 0$  and  $k < \infty$ . Since  $u^{1/2} > (u(1-u))^{1/2}$ , we can change this to

$$P\left(\text{for some } i \left| \frac{N_{r,i}}{n^{b(2)}} - u \right| \geq u^{1/2} n^{-b(2)/6+\beta}\right) \leq n^{-k}. \quad (4.3)$$

For this estimate to be useful, we need  $u \gg u^{1/2} n^{-b(2)/6+\beta}$  which is equivalent to  $u \gg n^{-b(2)/3+2\beta}$ . If  $b(2)$  is close to 1 and  $\beta$  is small, we can define  $b(0)$  by

$$1 - b(0) = b(2)/3 - 2\beta,$$

so that  $b(0) < \min\{b, 1 - \alpha\}$  where  $\alpha > 1/3$  is the quantity from Theorem 1.4. Combining these estimates and using results from the previous section we have that if  $x_0 < 1/2$  and  $\eta > 0$  then as  $n \rightarrow \infty$

$$P(|X_t - x_t| \leq \eta x_t \text{ for all } t \leq \tau) \rightarrow 1.$$

Lemma 4.1 follows.  $\square$

This result shows that the number of 1's gets driven to  $\leq (1 + \epsilon)n^{b(0)}$  at the deterministic time  $\tau$ . To complete the process of extinction we will rely on fluctuations in the voter model.

## 4.2 Green's function calculation

To motivate the calculation in the next lemma we note that the voter model is a time change of simple random walk.

**Lemma 4.2.** *Let  $S_t$  be continuous-time simple random walk on  $\{0, \dots, n\}$  with jump-rate  $r(j)$  at position  $j$ . Let  $0 < x < z \leq n$  be integers, and  $T_{0,z}$  the first time that  $S_t$  hits 0 or  $z$ . Then,*

$$E_x T_{0,z} = \sum_{y=1}^x \frac{2y}{r(y)} + \sum_{y=x+1}^z \frac{2x}{r(y)} - \sum_{y=1}^z \frac{2xy}{zr(y)}. \quad (4.4)$$

Since  $P_x(T_z < T_0) = x/z$ , this is enough to bound the extinction time if  $x/z \rightarrow 0$ .

*Proof.* First consider the embedded discrete-time chain of  $S_t$ . For  $0 \leq y \leq z$ , let  $N_x(y)$  be the number of times the random walk visits  $y$  before hitting 0 or  $z$ , starting from position  $x$ . Consider the Green's function

$$G_0(x, y) = \mathbb{E}[N_x(y)].$$

Fix  $y$  and write  $g(x) = G_0(x, y)$ . Then we have that  $g$  satisfies

$$\begin{cases} g(0) = 0 \\ g(x) = \frac{1}{2} (g(x+1) + g(x-1)), & x \neq 0, y, z \\ g(y) = 1 + \frac{1}{2} (g(y+1) + g(y-1)) \\ g(z) = 0 \end{cases}.$$

From this it is clear that  $g$  should be linear and increasing on  $[0, y]$  and linear and decreasing on  $[y, z]$ . That is,

$$\begin{cases} g(x) = c_1 x & 0 \leq x \leq y \\ g(x) = c_2(z-x) & y \leq x \leq z. \end{cases}.$$

To satisfy the conditions for  $g(x)$  and  $g(y)$ , the constants must be

$$c_1 = \frac{2(z-y)}{z}, \quad c_2 = \frac{2y}{z}.$$

The walk will spend an average of  $1/r(y)$  units of time at position  $y$  before jumping. Thus, if  $G(x, y)$  is defined to be the expected amount of time the continuous time walk spends at  $y$ , started from  $x$ , before hitting 0 or  $z$ , we have:

$$G(x, y) = \frac{1}{r(y)} \cdot G_0(x, y) = \frac{1}{r(y)} \cdot \begin{cases} 2x(z-y)/z & x \leq y \\ 2(z-x)y/z & x \geq y \end{cases}$$

Thus, the expected total time before being absorbed, started from  $x$ , is

$$\begin{aligned} E_x[T_{0,z}] &= \sum_{y=1}^z G(x, y) = \sum_{y=1}^x \frac{2y}{z} \cdot (z-x) \cdot \frac{1}{r(y)} + \sum_{y=x+1}^z \frac{2(z-y)}{z} \cdot x \cdot \frac{1}{r(y)} \\ &= \sum_{y=1}^x \frac{2y}{r(y)} + \sum_{y=x+1}^z \frac{2x}{r(y)} - \sum_{y=0}^z \frac{2xy}{zr(y)}, \end{aligned}$$

which establishes (4.4) □

### 4.3 Boundary size calculations

To use (4.4) to bound the extinction time, we need to understand the size of the boundary of the voter model:  $\partial\xi = \{\{x, y\} : x \sim y, \xi(x) \neq \xi(y)\}$ . Here  $x \sim y$  means that  $x$  and  $y$  are neighbors and  $\{x, y\}$  is the un-oriented edge that connects them. For a voter model configuration  $\xi$ , let  $|\xi| = \sum_x \xi(x)$  be the number of 1s. The next result gives trivial upper and lower bounds on  $|\partial\xi|$  when  $|\xi| = k$ :

$$C_d k^{1/d} \leq |\partial\xi| \leq 2dk. \quad (4.5)$$

Using (4.4), we see that if  $x = n^p$  and  $z = n^q$  for some  $0 < p < q < 1$ , then for  $r(y) = y$ ,

$$E_x T_{0,z} \leq \sum_{y=1}^x \frac{2y}{y} + \sum_{y=x+1}^z \frac{2x}{y} \leq Cx + 2x[\log(z) - \log(x)] \leq C'x \log(z) \quad (4.6)$$

If  $p = b(0)$  and  $q > p$ , this gives us what we want, an extinction time  $\ll n^b$ .

On the other hand, if we use the lower bound and plug in  $r(y) = y^{1/3}$ , then

$$E_x T_{0,z} \leq \sum_{y=1}^x \frac{2y}{y^{1/3}} + \sum_{y=x+1}^z \frac{2x}{y^{1/3}} \leq C(x^{5/3} + x^{2/3}z) \leq C'x^{2/3}z \quad (4.7)$$

If we take  $x = n^{b(0)}$  and  $z = n^c$  then this is  $\leq Cn^{5b(0)/3}$ , which is much longer than the interval of length  $n^b$  over which the process behaves like the voter model. Combining (4.5) and (4.7) gives

**Lemma 4.3.** *If  $x = n^p$  with  $p < 3b/5$  and  $z = n^q$  with  $q > p$  and  $2p/3 + q < b$  then*

$$P_x(T_{0,z} \leq n^b) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

This will let us show that the time spent at small values of  $|\partial\xi_t|$  can be ignored. For larger values, we need a more precise statement about the size of the boundary. This has been done by Cox, Durrett, and Perkins [9], in order to show that in  $d \geq 2$  the rescaled voter model converged in distribution to super-Brownian motion. This was later used by Bramson, Cox, and LeGall [3] to prove a result for the voter model in  $d \geq 3$  started at 0. See Theorem 4 on page 1012 in [3].

To prepare for stating our lemma we describe the result from [9]. They use a general probability kernel  $p(z)$ . In our case  $p(z) = 1/6$  for the nearest neighbors of 0. If  $\xi_t(x) = 1$  we let

$$V_t(x) = \sum_y p(y-x) 1_{(\xi_t(y)=1)}$$

If  $\xi_t(x) = 0$  we set  $V_t(x) = 0$ . This part of the definition is not really needed in the statement since  $X_s^N$  is supported by points on the rescale lattice in state 1. On page 202 of their result you find the following result.

(I1) There is a finite  $\gamma > 0$  so that for all  $\phi \in C_0^\infty(\mathbb{R}^d)$  and  $T > 0$

$$E \left[ \left( \int_0^T X_s^N ([V_{N,s} - \gamma]\phi^2) ds \right)^2 \right] \rightarrow 0$$

Here  $X_t^N$  is the voter model with space scaled by  $\sqrt{N}$  and time scaled by  $N$  and turned into a measure by assigning mass  $1/N$  to states in state 1, see (1.4), and  $V_{N,s}(x)$  is a suitably rescaled version of  $V_t(x)$ . The formula on page 202 has  $V'$  because they want to write the formula so that it is valid for  $d = 2$  and  $d \geq 3$ .

In our situation  $\gamma = 2d\beta_d$ . However, in this proof we need control on the size of the error. The reader should think of  $s$  as a point in the time interval  $[t/\epsilon_n - n^b/2, t/\epsilon_n]$  over which our process behaves like the voter model.

**Lemma 4.4.** *If  $k$  is large and the density of 1's is small then*

$$P\left(\frac{|\partial \xi_s|}{|\xi_s|} \notin [(1-\epsilon)2d\beta_d k, (1+\epsilon)2d\beta_d k] \mid |\xi_s| = k\right) \leq k^{-2/3}$$

*Proof.* Pick a site  $x$  at time  $s$  with  $\xi_s(x) = 1$ . When this holds the coalescing random walk starting at  $x$  at time  $s$  lands on a site in state 1 at time  $t/\epsilon_n - n^b$ . Let  $r = k^\alpha$  where  $\alpha$  is small and follow the CRW path backwards in time for  $r$  units of time. If we let  $h(s, x)$  be the probability the CRW starting at time  $s$  lands on 1 at time  $t/\epsilon_n - n^b$ , then an elementary conditional probability shows that the probability our conditioned CRW particle at  $x$  at time  $s$  is at  $y$  at time  $s - r$  is

$$\bar{p}_{s,r}(x, y) = p_r(x, y) \frac{h(r, y)}{h(s, x)}$$

This result is often known as Doob's  $h$ -transform. Since the lineage will wrap around the torus in the remaining  $\geq n^b/2$  units of time, the ratio is close to 1 and can be ignored.

For each neighbor  $y$  of an  $x$  with  $\xi_t(x) = 1$ , let  $V_{x,y} = 1$  if it does not coalesce with  $x$  by time  $r$  and 0 otherwise. For any  $\alpha > 0$ , if  $k$  is large and the density of 1's is  $u$  which is small then

$$\left| \frac{P(V_{x,y} = 1)}{2d\beta_d k} - 1 \right| < \eta/2.$$

Here we are using the hydrodynamic limit Lemma 3.6 to conclude that the distribution of the process is close to  $\nu_u$  at time  $r$ .

Let  $W_x = \sum_{y \sim x} V_{x,y}$ ,  $\mu(x) = \sum_{y \sim x} EV_{x,y}$ , and

$$S_k = \sum_x^* \bar{W}_x \quad \text{where} \quad W_x - \mu(x)$$

where  $\Sigma_x^*$  is short for  $\sum_{x: \xi_t^0(x)=1}$ . Arguments in Section 3.5 imply that if  $|x - x'| > s$  then the correlation between  $\bar{W}_x$  and  $\bar{W}_{x'}$  is small enough to be ignored so

$$E(S_k^2) = \sum_x^* \sum_y^* E[\bar{W}_x \bar{W}_y] \leq 36k \cdot Cr^3$$

since  $|\bar{W}_x| \leq 6$  and for a given  $x$  there are at most  $Cr^3$  values of  $y$  with  $|x - y| \leq r$ . If we use Chebyshev's inequality

$$P(|S_k| \geq \epsilon k) \leq \frac{36k \cdot Cs^3}{k^2} \leq Ck^{-1+3\delta}$$

If  $\alpha < 1/10$  this gives the desired result.  $\square$

#### 4.4 Extinction time

The results about the boundary of the voter model can now be applied to the Green's function calculation to get the result

**Lemma 4.5.** *Consider the voter model started with configuration  $|\xi_0| = x$  and let  $T_{0,z}$  be the first time the configuration hits 0 or  $z$ . If  $x = n^{b(0)}$  and  $z = n^c$  with  $c > b(0)$  then*

$$E_x[T_{0,z}] \leq Cn^{b(0)}$$

*Proof.* We can divide the sum in (4.4) into the pieces where Lemma 4.3 can be applied. That is, define  $x' = n^p < x$  so that  $p < 3b(0)/5$  and  $2p/3 + c < b(0)$ . Then,

$$\mathbb{E}_x[T_{0,z}] \leq \mathbb{E}_{x'}[T_{0,z}] + \mathbb{E}_x[T_{x',z}].$$

The first term is less than a constant times  $n^{b(0)}$  by Lemma 4.3. To bound the second hitting time, we use (4.6) and Lemma 4.4 to conclude that the expected amount of time when  $|\partial\xi_s|/|\xi_s|$  is not within  $\epsilon$  of  $2d\beta_d$  is

$$\leq \sum_{y=1}^x \frac{2y}{y^{1/3}} y^{-2/3} + \sum_{y=x+1}^z \frac{2x}{y^{1/3}} y^{-2/3} \leq \sum_{y=1}^x \frac{2y}{y^{1/3}} + \sum_{y=x+1}^z \frac{2x}{y^{1/3}} \leq Cn^b$$

which finally completes the proof.  $\square$

Theorem 1.4 now immediately follows: apply Lemma 4.1 to get that  $U_n(\alpha \log n) < n^{-(1-b(0))}$  with high probability. Next, use Section 3.2 so that with high probability we can assume the  $q$ -voter model only experiences voter branching events for the remainder of the time. Lemma 4.5 then proves that with high probability the unscaled voter model started with  $n^{b(0)}$  occupied sites will hit 0 or  $n^c$  in an additional time of  $Cn^{b(0)}$ . The probability that the process hits 0 first is simply  $(n^c - n^{b(0)})/n^c \rightarrow 1$ . Since  $b(0) > 2/3$ , this additional time is  $o(1)$  for the time-scaled process  $U_n(t)$ . Thus,

$$P(U_n(\alpha \log n) = 0) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

## 5 Computing the perturbation

In this section, Theorem 1.2 is proved. Recall Theorem 1.1 state that the limiting ODE for the model with a  $k$ -sized neighborhood is

$$\frac{du}{dt} = \sum_{m=1}^{k-1} r_i^k(\rho_m^0(u) - \rho_m^1(u)),$$

where  $\rho_m^i(u)$  is the probability under the voter model equilibrium  $\nu_u$  that the origin is in state  $i$  and a exactly  $m$  of the neighbors are in state  $1 - i$ . In this section, we analyze these quantities. Before stating the proof for a general  $k$ , we first show an explicit proof for a neighborhood of size 3 to give a flavor of how the individual terms are computed, while introducing some necessary notations in an organic manner.

### 5.1 k=3

To compute  $\rho_i^0$  we have to compute the coalescence fate of 0,  $e_1$ ,  $e_2$ ,  $e_3$ . There are 7 possibilities

|       |          |         |       |      |
|-------|----------|---------|-------|------|
| one   | 0 ; 3    | 1: 2    | 2 ; 1 | 3; 0 |
| two   | 0; 2, 1  | 1: 1, 1 |       |      |
| three | 0; 1,1,1 |         |       |      |

The first number in each string gives the number of neighbors that coalesce with 0. The others give the size of the limiting coalescing clusters formed by the remaining neighbors. The word at the beginning of the row is the number of numbers after the semi-colon. We can ignore 3;0 because in that case all the neighbors have the same state as 0.

Let  $\rho_i^0$  be the probability that in the voter equilibrium  $\nu_u$  the origin is 0 while exactly  $i$  of the neighbors are 1. Factoring out the probability the origin is we have  $\rho_i^0 = (1-u)q_i(u)$ . To compute the  $q_i(u)$  we use the following table.

- The coefficients of  $u$  come from the “one” terms.
- The coefficients of  $u^2$  and  $u(1-u)$  come from the “two” terms. There is no  $(1-u)^k$  since all the neighbors would be 0.  $p(1; 1, 1)$  appears three times since only 0,0 is impossible.  $p(0; 2, 1)$  only appears twice since 0,0 and 1,1 are impossible.

- The coefficients of  $u^2(1-u)$  and  $u(1-u)^2$  come from the “three” terms. There is no  $u^3$  or  $(1-u)^3$  since all neighbors would be 0 or 1. For this reason  $p_{0;1,1,1}$  appears  $2^3 - 2 = 6$  times

The meaning of the first column will become clear when the reader reaches (5.2)

| $\Delta_i(u)$ | <i>term</i> | $q_1(u)$                 | $q_2(u)$       |
|---------------|-------------|--------------------------|----------------|
| 0             | $u$         | $p_{2;1}$                | $p_{1;2}$      |
| -1            | $u^2$       |                          | $p_{1;1,1}$    |
| 1             | $u(1-u)$    | $p_{0;2,1} + 2p_{1;1,1}$ | $p_{0;2,1}$    |
| 1             | $u(1-u)^2$  | $3p_{0;1,1,1}$           |                |
| 0             | $u^2(1-u)$  |                          | $3p_{0;1,1,1}$ |

(5.1)

so reading down the columns we have

$$q_1(u) = p_{2,1}u + [2p_{1,1,1} + p_{0,2,1}]u(1-u) + 3p_{0,1,1,1}u(1-u)^2$$

$$q_2(u) = p_{1,2}u + p_{1,1,1}u^2 + p_{0,2,1}u(1-u) + 3p_{0,1,1,1}u^2(1-u)$$

Let  $\rho_i^1$  be the probability that in the voter equilibrium  $\nu_u$  the origin is 1 while exactly  $i$  of the neighbors are 0. From the previous calculation we see that  $\rho_i^1 = uq_i(1-u)$  so we have

$$\langle h_{0,1} - h_{1,0} \rangle_u = \sum_{i=1}^2 r_i(\rho_i^0 - \rho_i^1)$$

The quantity in parentheses is  $\Delta_i(u) \equiv (1-u)q_i(u) - uq_i(1-u)$ . Taking difference we have (the first column indicates the term in  $q_i(u)$ )

|            |                                                              |  |
|------------|--------------------------------------------------------------|--|
| $u$        | $u(1-u) - (1-u)u = 0$                                        |  |
| $u^2$      | $u^2(1-u) - (1-u)u^2 = u(1-u)(2u-1)$                         |  |
| $u(1-u)$   | $u(1-u)^2 - u^2(1-u) = u(1-u)(1-2u)$                         |  |
| $u(1-u)^2$ | $u(1-u)^3 - u^3(1-u) = u(1-u)[(1-u)^2 - u^2] = u(1-u)(1-2u)$ |  |
| $u^2(1-u)$ | $u^2(1-u)^2 - (1-u)^2u^2 = 0$                                |  |

(5.2)

so consulting (5.1) we have

$$\Delta_1(u) = [2p_{1,1,1} + p_{0,2,1} + 3p_{0,1,1,1}]u(1-u)(1-2u)$$

$$\Delta_2(u) = [-p_{1,1,1} + p_{0,2,1}]u(1-u)(1-2u)$$

and the reaction term is

$$\frac{\phi(u)}{u(1-u)(1-2u)} = r_1[2p_{1,1,1} + p_{0,2,1} + 3p_{0,1,1,1}]$$

$$+ r_2[-p_{1,1,1} + p_{0,2,1}]$$

If  $2r_1^3 > r_2^3$  so the right-hand side is positive. Using (1.1) we see that in the q-voter model with  $q < 1$

$$2r_1^3 = 2/3 \log(3) > 2/3 \log(3/2) = r_2^3$$

so the reaction term is  $c_3u(1-u)(1-2u)$  with  $c_3 > 0$ . When  $q > 1$  the reaction term is  $-c_3u(1-u)(1-2u)$ .

## 5.2 General k

In this case we have to compute the coalescence fate of 0 with  $k$  neighbors. Again  $\rho_i^0 = (1-u)q_i(u)$ , where the functions  $q_i(u)$ ,  $i \leq k-1$  defined as before are polynomials with terms of the type  $u^a(1-u)^b$ . First let us look at the difference  $\Delta_{a,b}(u)$  of these terms, where  $\Delta_{a,b}(u) = \rho_i^0 - \rho_i^1 = u^a(1-u)^{b+1} - u^{b+1}(1-u)^a$ . Note that  $\Delta_{a,b}(u) = 0$  if  $a = b+1$ .

In the case  $a \leq b$  we have

$$\begin{aligned}\Delta_{a,b}(u) &= u^a(1-u)^{b+1} - u^{b+1}(1-u)^a \\ &= u^a(1-u)^a[(1-u)^{b-a+1} - u^{b-a+1}] \\ &= u^a(1-u)^a(1-2u) \left[ \sum_{j=0}^{b-a} u^j(1-u)^{b-a-j} \right].\end{aligned}$$

To see the last step write  $1-2u = (1-u) - u$  and telescope the sum. In the case  $a > b+1$

$$\begin{aligned}\Delta_{a,b}(u) &= u^a(1-u)^{b+1} - u^{b+1}(1-u)^a \\ &= u^{b+1}(1-u)^{b+1}[u^{a-b-1} - (1-u)^{a-b-1}] \\ &= -u^{b+1}(1-u)^{b+1}(1-2u) \left[ \sum_{j=0}^{a-b-2} u^j(1-u)^{a-b-2-j} \right]\end{aligned}$$

Since  $\sum_{j=0}^n u^j(1-u)^{n-j} > 0$  on  $[0,1]$  we have that 0, 1 and  $1/2$  are the only roots of  $\Delta_{a,b}(u)$ . Also note that  $\Delta_{a,b}(u) = -\Delta_{b+1,a-1}(u)$ . We claim

$$\frac{\phi(u)}{u(1-u)(1-2u)} = f(u),$$

where  $f(u)$  is a positive polynomial in  $u$  with no real roots. To prove this, given a coalescence fate  $s_0; s_1, s_2, s_3, \dots, s_j$  where  $\sum_j s_j = k$  we look at number of ways to obtain  $a$  clusters with opinion 1 (which gives the coefficients of the terms  $u^a(1-u)^b$ ,  $a > b+1$ ) and compare it with the number of ways to obtain  $b+1$  clusters with opinion 1 (which gives the coefficients of the terms  $u^{b+1}(1-u)^{a-1}$ ).

First, suppose  $b = 0$  and  $a \geq 2$ . Let  $s_0$  be the number of neighbors that have coalesced with 0, and  $s_1, s_2, \dots, s_a$  be the sizes of the limiting coalescing clusters formed by the rest of the neighbors, where we assume that the sizes are arranged in an increasing order, i.e.,  $s_1 \leq s_2 \leq \dots \leq s_a$ . The coefficient of  $\Delta_{a,0}(u)$  in  $\phi(u)$  is given by  $r_{s_1+\dots+s_a} p_{s_0; s_1, \dots, s_a}$  (Since all the clusters have opinion 1, there is only one way to choose). Similarly the coefficient of  $\Delta_{1,a-1}(u)$  in  $\phi(u)$  is given by  $(r_{s_1} + \dots + r_{s_a}) p_{s_0; s_1, \dots, s_a}$  (Since exactly one of the clusters has opinion 1, there are  $a$  different choices, the coefficient of each of the clusters needs to be added individually).

Since  $s_i$ 's are increasing in  $i$ , so

$$\log(k/s_a) \leq \log(k/s_j) \quad \forall j \in \{1, 2, \dots, a-1\}.$$

So by the definition  $r_i^k = \frac{i}{k} \log(k/i)$ , and using the inequality above we have

$$\begin{aligned} r_{s_1+\dots+s_a} &= \frac{s_1+\dots+s_a}{k} \log(k/(s_1+\dots+s_a)) \\ &\leq \frac{s_1+\dots+s_a}{k} \log(k/s_a) \\ &= \frac{s_1}{k} \log(k/s_a) + \frac{s_2}{k} \log(k/s_a) + \dots + \frac{s_a}{k} \log(k/s_a) \\ &\leq \frac{s_1}{k} \log(k/s_1) + \frac{s_2}{k} \log(k/s_2) + \dots + \frac{s_a}{k} \log(k/s_a) \\ &= r_{s_1} + r_{s_2} + \dots + r_{s_a}. \end{aligned}$$

Since  $\Delta_{a,0}(u) = -\Delta_{1,a-1}(u)$ , if we only look at terms of the type  $\Delta_{1,a-1}(u)p_{s_0;s_1,\dots,s_a}$  (which is non-negative) in  $\phi(u)$ , we get a non-negative polynomial in  $u$  with no roots other than 0, 1 and  $1/2$ .

Now suppose  $b \neq 0$  and  $a \geq b+2$ . As explained in the previous case, let  $s_0$  be the number of neighbors that coalesce with 0, and  $s_1, s_2, \dots, s_{a+b}$  be the sizes of the limiting coalescing clusters formed by the rest of the neighbors, where we assume that the sizes are arranged in an increasing order, i.e.,  $s_1 \leq s_2 \leq \dots \leq s_{a+b}$ . There are  $\binom{a+b}{a}$  ways of choosing  $a$  clusters out of the  $a+b$  clusters. Denote the total size of each of these clusters by  $x_i$ , where  $1 \leq i \leq \binom{a+b}{a}$ , where wlog we assume that the sizes are arranged in an ascending order. The coefficient of  $\Delta_{a,b}(u)$  in  $\phi(u)$  is given by  $p_{s_0;s_1,s_2,\dots,s_{a+b}} \sum_{i=1}^{\binom{a+b}{a}} r_{x_i}$ . Given  $1 \leq i \leq a+b$ , the number of clusters in which cluster  $s_i$  has opinion 1 is given by  $\binom{a+b-1}{a-1}$ . Hence the total size of all the clusters, where  $a$  of them have opinion 1, is given by

$$\sum_{i=1}^{\binom{a+b}{a}} x_i = \binom{a+b-1}{a-1} (s_1 + s_2 + \dots + s_{a+b}).$$

Using a similar argument there are  $\binom{a+b}{b+1}$  ways of choosing  $b+1$  clusters out of the  $a+b$  clusters. Denote the total size of each of these clusters by  $y_i$ , where  $1 \leq i \leq \binom{a+b}{b+1}$ , where wlog we assume that the sizes are arranged in an ascending order. The coefficient of  $\Delta_{b+1,a-1}(u)$  in  $\phi(u)$  is given by  $p_{s_0;s_1,s_2,\dots,s_{a+b}} \sum_{i=1}^{\binom{a+b}{b+1}} r_{y_i}$ . Given  $1 \leq i \leq a+b$ , the number of clusters in which cluster  $s_i$  has opinion 1 is given by  $\binom{a+b-1}{b} = \binom{a+b-1}{a-1}$ . Hence the total size of all the clusters, where  $b+1$  of them have opinion 1, is given by

$$\sum_{i=1}^{\binom{a+b}{b+1}} y_i = \binom{a+b-1}{b} (s_1 + s_2 + \dots + s_{a+b}).$$

For ease of notation, let us denote  $\binom{a+b}{a}$  by  $n$  and  $\binom{a+b}{b+1}$  by  $m$ . Then  $m > n$  since

$$\begin{aligned} \binom{a+b}{b+1} - \binom{a+b}{a} &= \binom{a+b}{a} \left( \frac{a}{b+1} - 1 \right) \\ &= \binom{a+b}{a} \left( \frac{a-b-1}{b+1} \right) > 0. \end{aligned}$$

Since  $\sum_{i=1}^n x_i = \sum_{i=1}^m y_i$ , and the  $x_i$ s as well as the  $y_i$ s are arranged in ascending order, we have  $x_i > y_i + m - n$ , for  $1 \leq i \leq n$ .

$$\begin{aligned} &x_1 \log x_1 + x_2 \log x_2 + \dots + x_n \log x_n - y_1 \log y_1 - y_2 \log y_2 - \dots - y_m \log y_m \\ &> x_1 \log x_1 + x_2 \log x_2 + \dots + x_n \log x_n - y_1 \log y_1 - y_2 \log y_2 - \dots - y_m \log x_n \\ &= x_1 \log x_1 + x_2 \log x_2 + \dots + x_n \log x_n - y_1 \log y_1 - y_2 \log y_2 - \dots - y_{j-1} \log y_{j-1} - c \log y_j, \end{aligned}$$

where  $y_n + y_{n+1} + \dots + y_j - c = x_n$ . Now we have  $\sum_{i=1}^{n-1} x_i = c + \sum_{i=1}^{j-1} y_i$ . Repeating the same process as explained above  $n - 1$  times, we have

$$\sum_{i=1}^n x_i \log x_i > \sum_{i=1}^m y_i \log y_i.$$

Now using the definition of  $r_m^k$

$$\begin{aligned} \sum_{i=1}^n r_{x_i}^k &= \sum_{i=1}^n \frac{x_i}{k} \log(k/x_i) = \sum_{i=1}^n \frac{x_i}{k} [\log(k) - \log(x_i)] \\ &= \sum_{i=1}^m \frac{y_i}{k} \log(k) - \sum_{i=1}^n \frac{x_i}{k} \log(x_i) < \sum_{i=1}^m \frac{y_i}{k} \log(k) - \sum_{i=1}^m \frac{y_i}{k} \log(y_i) \\ &= \sum_{i=1}^m \frac{y_i}{k} \log(k/y_i) = \sum_{i=1}^m r_{y_i}^k. \end{aligned}$$

Now using the above inequality along with the fact that  $\Delta_{a,b} = -\Delta_{b+1,a-1}$ , if we only look at terms of the type  $\Delta_{b+1,a-1}(u)p_{s_0;s_1,\dots,s_{a+b}}$  (which is non-negative) in  $\phi(u)$ , we get a non-negative polynomial in  $u$  with no roots other than 0, 1 and  $1/2$ . This proves Theorem 1.2 for  $q < 1$ .

**Corollary 5.1.** Fix  $q > 1$ . For a  $q$ -voter model with  $k$ -neighbors, the reaction function defined in (1.3) simplifies to

$$\phi(u) = -c_k u(1-u)(1-2u)f_k(u), \quad (5.3)$$

where  $c_k > 0$  and  $f_k(u)$  is a strictly positive polynomial in  $u$ .

*Proof.* Recalling the perturbation from (1.1) and (1.2), note that the perturbation when  $q > 1$  has the same value as the perturbation when  $q < 1$  but with the opposite sign. This along with the above work proves the corollary.  $\square$

## References

- [1] Abraams, D.M., and Strogatz, S.H. (2003) Modelling the dynamics of language death. *Nature*. 424, 900
- [2] Bailey, N.T.J. (1990) The Elements of Stochastic Processes. Wiley Classic Edition. MR-1038907
- [3] Bramson, M., Cox, J.T. and LeGall, J.-F. (2001). Super-Brownian limits of voter model clusters. *Ann. Probab.* 29, 1001–1032 MR-1872733
- [4] Bramson, M., and Griffeath, D. (1979) Renormalizing the 3-dimensional voter model. *Ann. Probab.* 7, 418–432 MR-0528320
- [5] Bramson, M., and Griffeath, D. (1980) Asymptotics for interacting particle systems on  $\mathbb{Z}^d$ . *Prob. Theory Rel. Fields.* 53, 183–196 MR-0580912
- [6] Castellano, C., Muñoz, M., and Pastor-Satoros, R. (2009) The nonlinear  $q$ -voter. *Physical Review E*. 80, paper 041129
- [7] Centola, D. (2018) *How behavior spreads: The Science of Complex Contagions*. Princeton University Press
- [8] Cox, J.T., and Durrett, R. (2016) Evolutionary games on the torus with weak selection. *Stoch. Proc. Appl.* 126, 2388–2409 MR-3505231
- [9] Cox, J.T., Durrett, R., and Perkins, E.A. (2000) Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* 28, 185–224 MR-1756003
- [10] Cox, J.T., Durrett, R., and Perkins, E.A. (2013) Voter model perturbations and reaction diffusion equations. *Astérisque*. Volume 349. arXiv:1103.1676 MR-3075759

- [11] Darling, R.W.R., and Norris, J.R. (2008) Differential equation approximation for Markov chains. *Probability Surveys*. 5, 37–79 MR-2395153
- [12] Durrett, R., and Levin, S. (1994) The importance of being discrete (and spatial). *Theoret. Pop. Biol.* 46, 363–394
- [13] Durrett, R., Liggett, T.M., and Zhang, Y. (2014) The contact process with fast voting. *Electronic Journal of Probability*. 18 paper 28 MR-3174840
- [14] Durrett, R. and Neuhauser, C. (1994) Particle systems and reaction-diffusion equations. *Ann. Probab.* 22, 2890333 MR-1258879
- [15] Feller, W.F. (1970) *An Introduction to Probability Theory and Its Applications, Volume II*. John Wiley and Sons, New York MR-0270403
- [16] Griffeath, D.S. (1978) *Additive and Cancellative Interacting Particle Systems*. Springer Lecture Notes in Math. MR-0538077
- [17] Hammal, O.A., Chaté, H., Dornic, I., and Munoz, M.A. (2005) Langevin description of critical phenomenon with two symmetric absorbing states. *Physical Review Letters*. 94, paper 230601
- [18] Harris, T.E. (1976) On a class of set-valued Markov processes. *Ann. Probab.* 4, 175–194 MR-0400468
- [19] Holley, R., and Liggett, T. (1975) Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* 4, 195–228 MR-0402985
- [20] Huo, R. and Durrett, R. Latent voter model on locally tree-like random graphs. *Stoch. Proc. Appl.* 128, 1590–1614 MR-3780690
- [21] Jędrzejewski, A. (2017) Pair approximation for the  $q$ -voter model with independence on complex networks. *Physical Review E*. 98, paper 0123907
- [22] Lambiotte R., Saramaki, J., and Blondel, V.D. (2009) Dynamics of latent voters. *Physical Review E*. 79, paper 046107 MR-2551223
- [23] Liggett, T.M. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, New York. MR-1717346
- [24] Min, B., and San Miguel, M. (2017) Fragmentation transitions in a coevolving nonlinear voter model. *Scientific Reports*. 7, paper 12864
- [25] Moretti, P., Liu, S., Castellano, C., and Pastor-Satorras, S. (2013) Mean-field analysis of the  $q$ -voter model on networks. *Journal Statistical Physics*. 151, 113–130 MR-3045809
- [26] Romero, D.M., Meeder, B., and Kleinberg, J. (2011) Differences in the mechanisms of information diffusion across topics: Idioms, political hashtags, and complex contagion on Twitter. *Proceedings of WWW2011, Hyderabad, India* Association for Computing Machinery
- [27] Vazquez, F., Castelló, X., and San Miguel, M. (2010) Agent based models of language competition: macroscopic description and order-disorder transitions. *Journal of Statistical Mechanics*. paper P04007
- [28] Vazquez, F., and Lopez, C. (2008) Systems with two absorbing states: Relating the macroscopic dynamics with macroscopic behavior. *Physical Review E* 78, paper 061127
- [29] Vasconcelos, V.V., Levin, S.A., and Pinheiro, F.L. (2019) Consensus and polarization in competing complex contagion processes. *Journal of the Royal Society Interface*. June 2019 arXiv:1811.08525
- [30] Zähle, I. (2001) Renormalization of the voter model in equilibrium. *Ann. Probab.* 29, 1262–1302 MR-1872743

**Acknowledgments.** This work was begun during the 2019 AMS Math Research Communities meeting on Stochastic Spatial Models, June 9–15, 2019. We would like to thank Hwai-Ray Tung, a graduate student at Duke for producing the figures.

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>