Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 126 (2016) 2388–2409
stochastic processes and their applications
www.elsevier.com/locate/spa

Evolutionary games on the torus with weak selection

J. Theodore Cox[a,1], Rick Durrett[b,*,1]

[a] Department of Math., 215 Carnegie Building, Syracuse U., Syracuse NY, 13244-1150, United States
[b] Department of Math., Duke U., P.O. Box 90320, Durham NC, 27708-0320, United States

Received 7 December 2015; received in revised form 1 February 2016; accepted 3 February 2016
Available online 2 March 2016

Abstract

We study evolutionary games on the torus with N points in dimensions $d \geq 3$. The matrices have the form $\bar{G} = \mathbf{1} + wG$, where $\mathbf{1}$ is a matrix that consists of all 1's, and w is small. As in Cox Durrett and Perkins (2011) we rescale time and space and take a limit as $N \to \infty$ and $w \to 0$. If (i) $w \gg N^{-2/d}$ then the limit is a PDE on $\mathbb{R}^d$. If (ii) $N^{-2/d} \gg w \gg N^{-1}$, then the limit is an ODE. If (iii) $w \ll N^{-1}$ then the effect of selection vanishes in the limit. In regime (ii) if we introduce mutations at rate μ so that $\mu/w \to \infty$ slowly enough then we arrive at Tarnita's formula that describes how the equilibrium frequencies are shifted due to selection.

Keywords: Voter model; Voter model perturbation; PDE limit; Tarnita's formula

1. Introduction

Here we will be interested in n-strategy evolutionary games on the torus $\mathbb{T}_L = (\mathbb{Z} \bmod L)^d$. Throughout the paper we will suppose that $n \geq 2$ and $d \geq 3$. The dynamics are described by a game matrix $G_{i,j}$ that gives the payoff to a player who plays strategy i against an opponent who plays strategy j. As in [7,8], we will study games with matrices of the form $\bar{G} = \mathbf{1} + wG$, and $\mathbf{1}$ is a matrix that consists of all 1's, and $w = \epsilon^2$. We use two notations for the small parameter to make it easier to connect with the literature.

* Corresponding author.
E-mail address: rtd@math.duke.edu (R. Durrett).
[1] JTC was partially supported by NSF grant DMS 1208984, RD by DMS 1305997.

http://dx.doi.org/10.1016/j.spa.2016.02.004
0304-4149/

There are two commonly used update rules. To define them introduce

Assumption 1. Let p be a probability distribution on $\mathbb{Z}^d$ with finite range, $p(0) = 0$ and that satisfies the following symmetry assumptions.

- If π is a permutation of $\{1, 2, \ldots, d\}$ and $(\pi z)_i = z_{\pi(i)}$ then $p(\pi z) = p(z)$.
- If we let $\hat{z}^i_i = -z_i$ and $\hat{z}^i_j = z_j$ for $j \neq i$ then $p(\hat{z}^i) = p(z)$.

For example, if $p(z) = f(\|z\|_p)$ where $\|z\|_p$ is the L^p norm on $\mathbb{Z}^d$ with $1 \le p \le \infty$ then the symmetry assumptions are satisfied.

Birth–Death Dynamics. In this version of the model, a site x gives birth at a rate equal to its fitness

$$\psi(x) = \sum_y p(y - x)\bar{G}(\xi(x), \xi(y))$$

and the offspring, which uses the same strategy as the parent, replaces a "randomly chosen neighbor of x". Here, and in what follows, the phrase in quotes means z is chosen with probability $p(z - x)$. Note that we use the same transition probability to compute the fitness and do the displacement. In general they can be different.

Death–Birth Dynamics. In this case, each site x dies at rate 1 and is replaced by the offspring of a neighbor y chosen with probability proportional to $p(y - x)\psi(y)$.

Tarnita et al. [23,24] have studied the behavior of evolutionary games on more general graphs when $w = o(1/N)$ and N is the number of vertices. To describe their results, we begin with the two strategy game written as

$$G = \begin{array}{ccc} & 1 & 2 \\ 1 & \alpha & \beta \\ 2 & \gamma & \delta. \end{array} \tag{1}$$

In [23] strategy 1 is said to be favored by selection (written $1 > 2$) if the frequency of 1 in equilibrium is $>1/2$ when w is small. Assuming that

(i) the transition probabilities are differentiable at $w = 0$,

(ii) the update rule is symmetric for the two strategies, and

(iii) strategy 1 is not disfavored in the game given with $\beta = 1$ and $\alpha = \gamma = \delta = 0$

they argued that

I. $1 > 2$ *is equivalent to* $\sigma\alpha + \beta > \gamma + \sigma\delta$ *where* σ *is a constant that only depends on the spatial structure and update rule.*

In [8] it was shown that for games on $\mathbb{Z}^d$ with $d \ge 3$.

Theorem 1. *I holds for the Birth–Death updating with* $\sigma = 1$ *and for the Death–Birth updating with* $\sigma = (\kappa + 1)/(\kappa - 1)$ *where*

$$\kappa = 1 \Big/ \sum_x p(x)p(-x) \tag{2}$$

is the effective number of neighbors.

The name for κ comes from the fact that if each $p(z) \in \{0, 1/m\}$ for all z then $\kappa = m$.

In [24] strategy k is said to be favored by selection in an n strategy game if, in the presence of weak selection, i.e., w is small its frequency in equilibrium is $>1/n$. To state their result we need some notation.

$$\hat{G}_{*,*} = \frac{1}{n}\sum_{i=1}^{n} G_{i,i} \qquad \hat{G}_{k,*} = \frac{1}{n}\sum_{i=1}^{n} G_{k,i}$$

$$\hat{G}_{*,k} = \frac{1}{n}\sum_{i=1}^{n} G_{i,k} \qquad \hat{G} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} G_{i,j}$$

where $*$'s indicate values that have been summed over. To make it easier for us to prove the result and to have nicer constants, we will rewrite their condition for strategy k to be favored as

$$\alpha_1(\hat{G}_{k,*} - \hat{G}) + \alpha_2(G_{k,k} - \hat{G}_{*,*}) + \alpha_3(\hat{G}_{k,*} - \hat{G}_{*,k}) > 0 \tag{3}$$

and refer to it as *Tarnita's formula*. The parameters α_i depend on the population structure and the update rule, but they do not depend on the number of strategies or on the entries $G_{i,j}$ of the payoff matrix. In [24] they divide by α_3 so $\sigma_2 = \alpha_1/\alpha_3$ and $\sigma_1 = \alpha_2/\alpha_3$.

When $n > 2$, (3) is different from the result for almost constant sum three strategy games on $\mathbb{Z}^d$ proved in [8]. The condition (3) is linear in the entries in the game matrix while the condition (8.13) in [8] for the infinite graph is quadratic. This paper arose from an attempt to understand this discrepancy. The resolution, as we will explain, is that the two formulas apply to different weak selection regimes.

The path we take to reach this conclusion is somewhat lengthy. In Section 2, we introduce the voter model and describe its duality with coalescing random walk. Section 3 introduces the voter model perturbations studied by Cox, Durrett, and Perkins [7]. Section 4 states their result that when space and time are scaled appropriately, the limit is a partial differential equation. The limit PDE is then computed for birth–death and death–birth updating. They are reaction diffusion equations with a reaction term that is a cubic polynomial. Section 5 uses the PDE limit to analyze 2×2 games. Section 6 introduces a duality for voter model perturbations, which is the key to their analysis. Section 7 gives our results for regimes (i) and (ii) in the abstract and our version of Tarnita's formula for games with $n \geq 3$ strategies. Sections 8–11 are devoted to proofs.

2. Voter model

Our results for evolutionary games are derived from results for a more general class of processes called *voter model perturbations*. To introduce those we must first describe the *voter model*. The state of the voter model at time t is $\xi_t : \mathbb{Z}^d \to S$ where S is a finite set of states, and $\xi_t(x)$ gives the state of the individual at x at time t. To formulate this class of models, let $p(z)$ be a probability distribution on $\mathbb{Z}^d$ satisfying the conditions in Assumption 1. In the voter model, the rate at which the voter at x changes its opinion from i to j is

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

where $f_j(x,\xi) = \sum_y p(y-x)1(\xi(y) = j)$ is the probability that a neighbor of x chosen at random is in state i. In words at times of a rate 1 Poisson process the voter at x wakes up and with probability $p(y-x)$ imitates the opinion of the individual at y.

To analyze the voter model it is convenient to construct the process on a *graphical representation* introduced by Harris [14] and further developed by Griffeath [13]. For each

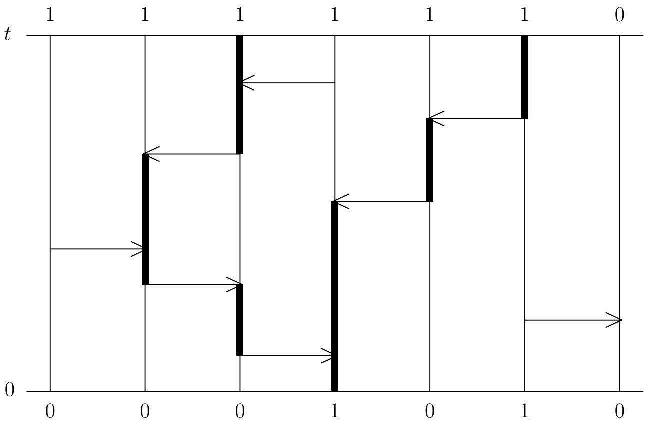


Fig. 1. Voter model graphical representation and duality.

$x \in \mathbb{Z}^d$ and y with $p(y-x) > 0$ let $T_n^{x,y}, n \geq 1$ be the arrival times of a Poisson process with rate $p(y-x)$. 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, we draw an arrow from $(x, T_n^{x,y})$ to $(y, T_n^{x,y})$. To calculate the state of the voter model on a finite set, we start at the bottom and work our way up determining what should happen at each arrow. 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.

To define the *dual process* we start with $\zeta_0^{x,t} = x$ and work down the graphical representation. The process stays at x until the first time $t - r$ that it encounters the tail of an arrow x. At this time, the particle jumps to the site y at the head of the arrow, i.e., $\zeta_r^{x,t} = y$. The particle stays at y until the next time the tail of an arrow is encountered and then jumps to the head of the arrow etc. Intuitively $\zeta_s^{x,t}$ gives the source at time $t - s$ of the opinion at x at time t so for all $s \in [0, t]$

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

The example in Fig. 1 should help explain the definitions. The family of particles $\zeta_s^{x,t}$ are coalescing random walk. Each particle at rate 1 makes jumps according to p. If a particle $\zeta_s^{x,t}$ lands on the site occupied by $\zeta_s^{y,t}$ they coalesce to 1, and we know that $\xi_t(x) = \xi_t(y)$. The dark lines indicate the locations of the two dual particles that coalesce.

Using duality it is easy to analyze the asymptotic behavior of the voter model. The results we are about to quote were proved by Holley and Liggett [16], and can also be found in Liggett's book [17]. In dimensions 1 and 2, random walks satisfying our assumptions are recurrent, so the voter model clusters, i.e.,

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

In $d \geq 3$ random walks are transient so differences in opinion persist as $t \to \infty$. Consider, for simplicity, the case of two opinions, 0 and 1. Let ξ_t^u be the voter model starting from product measure with density u, i.e., the initial voter opinions are independent and $= 1$ with probability u. As $t \to \infty$, ξ_t^u converges to a limit distribution ν_u.

A consequence of this duality relation is that if we let $p(0|x)$ be the probability that two continuous time random walks, one starting at the origin 0, and one starting at x never hit then

$$\nu_u(\xi(0) = 1, \xi(x) = 0) = p(0|x)u(1-u) \tag{4}$$

since in order for the two opinions to be different at time t, the corresponding random walks cannot hit, and they must land on sites with the indicated opinions, an event of probability $u(1-u)$.

To extend this reasoning to three sites, consider random walks starting at 0, x, and y. Let $p(0|x|y)$ be the probability that the three random walks never hit and let $p(0|x,y)$ be the probability that the walks starting from x and y coalesce, but they do not hit the one starting at 0. Considering the possibilities that the walks starting from x and y may or may not coalesce:

$$\nu_u(\xi(0)=1,\xi(x)=0,\xi(y)=0) = p(0|x|y)u(1-u)^2 + p(0|x,y)u(1-u). \tag{5}$$

All the finite dimensional distributions of ν_u can be computed in this way.

3. Voter model perturbations

The processes that we consider have flip rates

$$c^v_{i,j}(x,\xi) + \epsilon^2 h^\epsilon_{i,j}(x,\xi). \tag{6}$$

The perturbation functions h^ϵ_{ij}, $j \neq i$, may be negative but in order for the analysis in [7] to work, there must be a law q of $(Y^1,\dots,Y^M) \in (\mathbb{Z}^d)^M$ and functions $g^\epsilon_{i,j} \ge 0$, which converge to limits $g_{i,j}$ as $\epsilon \to 0$, so that for some $\gamma < \infty$, we have for $\epsilon \le \epsilon_0$

$$h^\epsilon_{i,j}(x,\xi) = -\gamma f_i(x,\xi) + E_Y[g^\epsilon_{i,j}(\xi(x+Y^1),\dots\xi(x+Y^M))]. \tag{7}$$

In words, we can make the perturbation positive by adding a positive multiple of the voter flip rates. This is needed so that [7] can use $g^\epsilon_{i,j}$ to define jump rates of a Markov process.

For simplicity we will assume that both p and q are finite range. Applying Proposition 1.1 of [7] now implies the existence of suitable $g^\epsilon_{i,j}$ and that all our calculations can be done using the original perturbation. However, to use Theorems 1.4 and 1.5 in [7] we need to suppose that

$$h_{i,j} = \lim_{\epsilon\to 0} h^\epsilon_{i,j} \tag{8}$$

has $|h_{i,j}(\xi) - h^\epsilon_{i,j}(\xi)| \le C\epsilon^r$ for some $r > 0$, see (1.41) in [7].

Birth–Death Dynamics. If we let $r_{i,j}(0,\xi)$ be the rate at which the state of 0 flips from i to j,

$$\begin{aligned} r_{i,j}(0,\xi) &= \sum_x p(x)1(\xi(x)=j)\sum_y p(y-x)\bar{G}(j,\xi(y)) \\ &= \sum_x p(x)1(\xi(x)=j)\left(1+\epsilon^2\sum_k f_k(x,\xi)G_{j,k}\right) \\ &= f_j(0,\xi) + \epsilon^2\sum_k f^{(2)}_{j,k}(0,\xi)G_{j,k}, \end{aligned} \tag{9}$$

where $f^{(2)}_{j,k}(0,\xi) = \sum_x\sum_y p(x)p(y-x)1(\xi(x)=j,\xi(y)=k)$. Thus the perturbation, which does not depend on ϵ is

$$h_{i,j}(0,\xi) = \sum_k f^{(2)}_{j,k}(0,\xi)G_{j,k}. \tag{10}$$

If p is uniform on the nearest neighbors of 0, then q is nonrandom and $Y^1,\dots,Y^m$ is a listing of the nearest and next nearest neighbors of 0.

Death–Birth Dynamics. Using the notation in (9) the rate at which $\xi(0) = i$ jumps to state j is

$$\bar{r}_{i,j}(0,\xi) = \frac{r_{i,j}(0,\xi)}{\sum_k r_{i,k}(0,\xi)} = \frac{f_j(0,\xi) + \epsilon^2 h_{i,j}(0,\xi)}{1 + \epsilon^2 \sum_k h_{i,k}(0,\xi)}$$
$$= f_j(0,\xi) + \epsilon^2 h_{i,j}(0,\xi) - \epsilon^2 f_j(0,\xi) \sum_k h_{i,k}(0,\xi) + O(\epsilon^4). \tag{11}$$

The new perturbation, which depends on ϵ, is

$$\bar{h}^{\epsilon}_{i,j}(0,\xi) = h_{i,j}(0,\xi) - f_j(0,\xi) \sum_k h_{i,k}(0,\xi) + O(\epsilon^2). \tag{12}$$

As noted above the technical condition (7) holds because p has finite range. (8) holds with $r = 2$.

4. PDE limit

Let ξ^ϵ_t be the process with flip rates given in (6). The next result is the key to the analysis of voter model perturbations on $\mathbb{Z}^d$. Intuitively, it says that if we rescale space to $\epsilon\mathbb{Z}^d$ and speed up time by ϵ^{-2} the process converges to the solution of a partial differential equation. The first thing we have to do is to define the mode of convergence. Given $r \in (0, 1)$, let $a_\epsilon = \lceil \epsilon^{r-1} \rceil \epsilon$, $Q_\epsilon = [0, a_\epsilon)^d$, and $|Q_\epsilon|$ the number of points in Q_ϵ. For $x \in a_\epsilon\mathbb{Z}^d$ and $\xi \in \Omega_\epsilon$ the space of all functions from $\epsilon\mathbb{Z}^d$ to S let

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

We endow Ω_ϵ with the σ-field $\mathcal{F}_\epsilon$ generated by the finite-dimensional distributions. Given a sequence of measures λ_ϵ on $(\Omega_\epsilon, \mathcal{F}_\epsilon)$ and continuous functions w_i, we say that λ_ϵ has asymptotic densities w_i if for all $0 < \delta$, $R < \infty$ and all $i \in S$

$$\lim_{\epsilon \to 0} \sup_{x \in a_\epsilon \mathbb{Z}^d, |x| \le R} \lambda_\epsilon(|D_i(x,\xi) - w_i(x)| > \delta) = 0.$$

Theorem 2. *Suppose $d \ge 3$. Let $w_i : \mathbb{R}^d \to [0, 1]$ be continuous with $\sum_{i \in S} w_i = 1$. Suppose the initial conditions ξ^ϵ_0 have laws λ_ϵ with local densities w_i and let*

$$u^\epsilon_i(t,x) = P(\xi^\epsilon_{t\epsilon^{-2}}(x) = i).$$

If $x_\epsilon \to x$ then $u^\epsilon_i(t, x_\epsilon) \to u_i(t, x)$ the solution of the system of partial differential equations:

$$\frac{\partial}{\partial t} u_i(t,x) = \frac{\sigma^2}{2} \Delta u_i(t,x) + \phi_i(u(t,x)) \tag{13}$$

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

$$\phi_i(u) = \sum_{j \ne i} \langle 1_{(\xi(0)=j)} h_{j,i}(0,\xi) - 1_{(\xi(0)=i)} h_{i,j}(0,\xi) \rangle_u \tag{14}$$

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

This result is Theorem 1.2 in [7]. Intuitively, on the fast time scale the voter model runs at rate ϵ^{-2} versus the perturbation at rate 1, so the process is always close to the voter equilibrium for the current density vector u. Thus, we can compute the rate of change of u_i by assuming the nearby sites are in that voter model equilibrium.

In a homogeneously mixing population the frequencies of the strategies in an evolutionary game follow the replicator equation, see e.g., Hofbauer and Sigmund's book [15]:

$$\frac{du_i}{dt} = \phi_R^i(u) \equiv u_i \left(\sum_k G_{i,k} u_k - \sum_{j,k} u_j G_{j,k} u_k \right). \tag{15}$$

We will now compute the reaction terms ϕ_i for our two examples.

Birth–Death Dynamics. On $\mathbb{Z}^d$ we let v_i be independent with $P(v_i = x) = p(x)$. Let

$$p_1 = p(0|v_1|v_1 + v_2) \quad \text{and} \quad p_2 = p(0|v_1, v_1 + v_2)$$

where the p's are defined just before (5). In this case the limiting PDE in Theorem 2 is $\partial u_i/\partial t = (1/2d)\Delta u + \phi_B^i(u)$ where

$$\phi_B^i(u) = p_1 \phi_R^i(u) + p_2 \sum_{j \neq i} u_i u_j (G_{i,i} - G_{j,i} + G_{i,j} - G_{j,j}). \tag{16}$$

See Section 12 of [8] for a proof. Formula (4.8) in [8] implies that

$$2p(0|v_1, v_1 + v_2) = p(0|v_1) - p(0|v_1|v_1 + v_2),$$

so it is enough to know the two probabilities on the right-hand side.

If coalescence is impossible then $p_1 = 1$ and $p_2 = 0$ so $\phi_B^i = \phi_R^i$. There is a second more useful connection to the replicator equation. Let

$$A_{i,j} = \frac{p_2}{p_1}(G_{i,i} + G_{i,j} - G_{j,i} - G_{j,j}).$$

The matrix A is skew symmetric. That is, $A_{i,i} = 0$ and if $i \neq j$ $A_{i,j} = -A_{j,i}$. This implies $\sum_{i,j} u_i A_{i,j} u_j = 0$ and it follows that $\phi_B^i(u)$ is p_1 times the RHS of the replicator equation for the game matrix $A + G$. This observation is due to Ohtsuki and Nowak [22] who studied the limiting ODE that arises from the pair approximation.

Death–Birth Dynamics. On $\mathbb{Z}^d$ we let v_i be independent with $P(v_i = x) = p(x)$, let

$$\bar{p}_1 = p(v_1|v_2|v_2 + v_3) \quad \text{and} \quad \bar{p}_2 = p(v_1|v_2, v_2 + v_3).$$

With Death–Birth updating the limiting PDE is $\partial u_i/\partial t = (1/2d)\Delta u + \phi_D^i(u)$ where

$$\begin{aligned} \phi_D^i(u) = {} & \bar{p}_1 \phi_R^i(u) + \bar{p}_2 \sum_{j \neq i} u_i u_j (G_{i,i} - G_{j,i} + G_{i,j} - G_{j,j}) \\ & - (1/\kappa) p(v_1|v_2) \sum_{j \neq i} u_i u_j (G_{i,j} - G_{j,i}) \end{aligned} \tag{17}$$

where $\kappa = 1/P(v_1 + v_2 = 0)$ is the "effective number of neighbors". Again see Section 12 of [8] for a proof. The first two terms are the ones in (16) with p_i replaced by $\bar{p}_i$. The similarity is not surprising since the numerators of the flip rates in (11) are the flip rates in (9). The third term comes from the denominator in (11). Formula (4.9) in [8] implies that

$$2p(v_1|v_2, v_2 + v_3) = (1 + 1/\kappa)p(v_1|v_2) - p(v_1|v_2|v_2 + v_3),$$

so again it is enough to know the two probabilities on the right-hand side.

As in the Birth–Death case, if we let

$$\bar{A}_{i,j} = \frac{\bar{p}_2}{\bar{p}_1}(G_{i,i} + G_{i,j} - G_{j,i} - G_{j,j}) - \frac{p(v_1|v_2)}{\kappa \bar{p}_1}(G_{i,j} - G_{j,i}),$$

then $\phi_i^D(u)$ is $\bar{p}_1$ times the RHS of the replicator equation for $\bar{A} + G$.

5. Two strategy games

In a homogeneously mixing population playing the game G in (1), the fraction of individuals playing the first strategy, u, evolves according to the replicator equation (15):

$$\begin{aligned}\frac{du}{dt} &= u\{\alpha u + \beta(1-u) - u[\alpha u + \beta(1-u)] - (1-u)[\gamma u + \delta(1-u)]\}\\ &= u(1-u)[\beta - \delta + \Gamma u] \equiv \phi_R(u) \qquad (18)\end{aligned}$$

where we have introduced $\Gamma = \alpha - \beta - \gamma + \delta$. Note that $\phi_R(u)$ is a cubic with roots at 0 and at 1. If there is a fixed point in (0, 1) it occurs at

$$\bar{u} = \frac{\beta - \delta}{\beta - \delta + \gamma - \alpha}. \qquad (19)$$

Using results from the previous section gives the following.

Birth–Death Dynamics. The limiting PDE is $\partial u/\partial t = (1/2d)\Delta u + \phi_B(u)$ where $\phi_B(u)$ is p_1 times the RHS of the replicator equation for the game

$$\begin{pmatrix} \alpha & \beta + \theta \\ \gamma - \theta & \delta \end{pmatrix} \qquad (20)$$

and $\theta = (p_2/p_1)(\alpha + \beta - \gamma - \delta)$.

Death–Birth Dynamics. The limiting PDE is $\partial u/\partial t = (1/2d)\Delta u + \phi_D(u)$ where $\phi_D(u)$ is $\bar{p}_1$ times the RHS of the replicator equation for the game in (20) but now

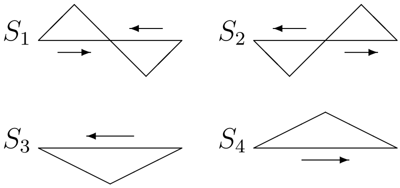
$$\theta = \frac{\bar{p}_2}{\bar{p}_1}(\alpha + \beta - \gamma - \delta) - \frac{p(v_1|v_2)}{\kappa \bar{p}_1}(\beta - \gamma).$$

5.1. *Analysis of* 2×2 *games*

Suppose that the limiting PDE is $\partial u/\partial t = (1/2d)\Delta u + \phi(u)$ where ϕ is a cubic with roots at 0 and 1. There are four possibilities

S_1	$\bar{u}$ attracting	$\phi'(0) > 0, \phi'(1) > 0$
S_2	$\bar{u}$ repelling	$\phi'(0) < 0, \phi'(1) < 0$
S_3	$\phi < 0$ *on* (0, 1)	$\phi'(0) < 0, \phi'(1) > 0$
S_4	$\phi > 0$ *on* (0, 1)	$\phi'(0) > 0, \phi'(1) < 0.$

To see this, we draw a picture. For convenience, we have drawn the cubic as a piecewise linear function.



We say that *i's take over* if for all K

$$P(\xi_s(x) = i \text{ for all } x \in [-K, K]^d \text{ and all } s \geq t) \to 1 \quad \text{as } t \to \infty.$$

Let Ω_0 be the configurations with infinitely many 1's and infinitely many 2's. We say that *coexistence occurs* if there is a stationary distribution μ for the spatial model with $\mu(\Omega_0) = 1$. The next result follows from Theorems 1.4 and 1.5 in [7]. The PDE assumptions and the other conditions can be checked as in the arguments in Section I.4 of [7] for the Lotka–Volterra system. Intuitively, the result says that the behavior of the particle system for small ϵ is the same as that of the PDE.

Theorem 3. *If $\epsilon < \epsilon_0(G)$, then: In case S_3, 2's take over. In case S_4, 1's take over. In case S_2, 1's take over if $\bar{u} < 1/2$, and 2's take over if $\bar{u} > 1/2$. In case S_1, coexistence occurs. Furthermore, if $\delta > 0$ and $\epsilon < \epsilon_0(G, \delta)$ then any stationary distribution with $\mu(\Omega_0) = 1$ has*

$$\sup_x |\mu(\xi(x) = 1) - \bar{u}| < \delta.$$

This result, after some algebra gives Tarnita's formula for two person games, Theorem 1. The key observation is

Lemma 1. $1 > 2$ *if and only if the reaction term in the PDE has* $\phi(1/2) > 0$.

Proof. Clearly $\phi(1/2) > 0$ in case S_4 but not S_3. In case S_1, $\phi(1/2) > 0$ implies $\bar{u} > 1/2$, while in case S_2 $\phi(1/2) > 0$ implies $\bar{u} < 1/2$ and hence the 1's take over. □

With this result in hand, Theorem 1 follows from the formulas for ϕ_B and ϕ_D.

6. Duality for voter model perturbations

We now return to the general case with n strategies. Our next step is to introduce a duality that generalizes the one for the voter model. Suppose now that we have a voter model perturbation of the form

$$h^\epsilon_{i,j}(x, \xi) = -\gamma f_i(x, \xi) + E_Y[g^\epsilon_{i,j}(\xi(x + Y^1), \dots \xi(x + Y^M))].$$

For each $x \in \mathbb{Z}^d$ and y with $p(y) > 0$ let $T^{x,y}_k$, $k \geq 1$ be the arrival times of a Poisson process with rate $p(y)$. At the times $T^{x,y}_k$, $k \geq 1$, x decides to change its opinion to match the one at $x + y$ where the arithmetic is done modulo L in each coordinate. We call this a voter event.

To accommodate the perturbation we let

$$\|g^\epsilon_{i,j}\| = \sup_{\eta \in S^M} g^\epsilon_{i,j}(\eta_1, \dots \eta_M)$$

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

$$g_{i,j}^\epsilon(\xi_{t-}(x+Y^1), \dots \xi_{t-}(x+Y^m)) < r_{i,j} U_k^{x,i,j} \tag{21}$$

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, suppose we start with particles at $X_1(0), \dots, X_k(0)$ at time t. We let $K(0) = k$ be the number of particles, $J(0) = \{1, 2, \dots, k\}$ be the indices of the active particles, and $T_0 = 0$. Suppose we have constructed the dual up to time T_m with $m \geq 0$. No particle moves from its position at time T_m until the first time $r > T_m$ that the tail of an arrow touches one of the active particles at time $t - r$. Call that time T_{m+1}. We extend the definitions of $K(t)$, $X_i(t)$, $i \leq K(t)$, and $J(t)$ to be constant on $[T_m, T_{m+1})$.

If the arrow is from a voter event affecting particle number i then X_i jumps to the head of the arrow at time T_{m+1}. If there is another active particle X_j on that site, the two coalesce to 1 and the higher numbered particle is removed from the active set at time T_{m+1}. If the event is a branching event, we add new particles numbered $K(T_m) + k$ at $X_i(T_m) + Y^k$ for $1 \leq k \leq M$ and set $K(T_{m+1}) = K(T_m) + M$. If there are collisions between the newly created particles and existing active particles, those newly created particles are not added to the active set. Our proof will show that in the situation covered in Theorem 6 the probability of a collision at a branching event will go to zero as $N \to \infty$.

Durrett and Neuhauser [11] called $I(s) = \{X_i(s) : i \in J(s)\}$ the influence set because

Lemma 2. *If we know the values of ξ_{t-s} on $I(s)$, the locations and types of arrows that occurred at the jump times $T_m \leq s$, and the associated uniform random variables U_m then we can compute the values of ξ_t at $X_1(0), \dots, X_k(0)$ 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 should be clear from the construction. A formal proof can be found in Section 2.6 of [7]. The computation process, as it is called in [7] is complicated, but is useful because up to time t/ϵ^2 there will only be $O(1)$ branching events. In between these events there will be many random walk steps that on the rescaled lattice will converge to Brownian motions.

7. Results for the torus

To motivate the results that we are about to state, recall that if we have a random walk on the torus $\mathbb{T}_L = (\mathbb{Z} \bmod L)^d$ that takes jumps at rate 1 with a distribution p that satisfies our assumptions in Assumption 1, then:

- One random walk needs of order L^2 steps to converge to the uniform distribution on the torus (see Proposition 1 in Section 10).
- Two independent random walks starting from randomly chosen points will need of order $N = L^d$ steps to meet for the first time. See e.g., [6].

7.1. Regime 1. $\epsilon_L^{-1} \ll L$, or $w \gg N^{-2/d}$

In this case when we rescale space by multiplying by ϵ_L then the limit of the torus is all of $\mathbb{R}^d$ and the PDE limit, Theorem 2, holds. Thus, one can apply results from Section 7 in [7] to show that the conclusions of Theorem 3 hold in cases S_2, S_3, and S_4. Indeed, since we are on the torus the linearly growing "dead zone" produced by the block construction eventually covers the entire torus and the weaker type becomes extinct at a time $O(L)$.

Case S_1 is more interesting. We cannot have a nontrivial stationary distribution since we are dealing with a Markov chain on a finite set in which $\xi(x) = i$ for all x are absorbing. However, as is the case for many other particle systems, e.g., the contact process on a finite set, we will have a quasi-stationary distribution that will persist for a long time. Using the comparison with oriented percolation described in Chapter 6 of [7] and in [9] we can show

Theorem 4. *Consider a two strategy evolutionary game with types* 0 *and* 1 *in case* S_1*, so* $\phi_1(u) = \lambda u(u - \rho)(1 - u)$*. Suppose that* $\epsilon_L^{-1} \sim CL^\alpha$ *where* $0 < \alpha < 1$ *and that for each* L *we start from a product measure in which each type has a fixed 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* $\delta > 0$ *if* L *is large and* $\log L \le t \le \exp(cL^{(1-\alpha)d})$ *then* $N_1(t)/N \in (\rho - \delta, \rho + \delta)$ *with high probability.*

The $\log L$ time needed to come close to equilibrium could be replaced by a fixed time T that depends on λ, ρ, δ, and the initial density of 1's. Our proof will show that with high probability at any time $\log L \le t \le \exp(cL^{(1-\alpha)d})$ the density is close to ρ (in the sense used to define the hydrodynamic limit) across most of the torus.

In many situations, e.g., the supercritical contact process on the d-dimensional cube [20], and power-law random graphs [21], the quasi-stationary distribution persists for time $\exp(\gamma N^d)$. However, we think that is not true in Theorem 4. For a simpler situation where we can prove this, consider the

Contact process with fast voting, studied by Durrett, Liggett, and Zhang [10]. In this voter model perturbation, there are two states, 0 and 1.

- $h_{1,0}(x, \xi) \equiv 1$: particles die at rate 1.
- $h_{0,1}(x, \xi) = \lambda f_1(x, \xi)$: a particle at x gives birth to a new one at $x + y$ at rate $\lambda p(y)$.

We only have to keep track of 1's so the reaction term

$$\begin{aligned}\phi(u) \equiv \phi_1(u) &= \langle 1_{(\xi(0)=0)} \lambda f_1(x, \xi) - 1_{(\xi(0)=1)} \rangle_u \\ &= \lambda p(0|v_1) u(1 - u) - u.\end{aligned}$$

If $\beta = \lambda p(0|v_1) > 1$ then 0 is an unstable equilibrium for $du/dt = \phi(u)$ and there is a fixed point at $\rho = (\beta - 1)/\beta$. The proof of Theorem 4 can be easily extended to show survival up to time $\exp(cL^{(1-\alpha)d})$ with high probability. In this case we can prove a partial converse.

Theorem 5. *There is an* $C < \infty$ *so that there are no occupied sites by time* $\exp(CL^{d-2\alpha} \log L)$ *with high probability.*

Note that the powers of L in the two results, $d(1 - \alpha)$ and $d - 2\alpha$, do not match. We suspect that the larger value is the correct answer. However it is not clear how to improve the proof of Theorem 4 to close the gap.

7.2. Regime 2. $L \ll \epsilon_L^{-1} \ll L^{d/2}$ *or* $N^{-2/d} \gg w \gg N^{-1}$.

In this case the time scale for the perturbation to have an effect, ϵ_L^{-2} is much larger than the time $O(L^2)$ needed for a random walk to come to equilibrium, but much smaller than the time

$O(L^d)$ it takes for two random walks to hit. Because of this, the particles in the dual will (except for times $O(L^2 \log L)$ after the initial time or a branching event) be approximately independent and uniformly distributed across the torus. Thus, if we speed up time by ϵ_L^{-2} the fraction of sites on the torus in state i will converge to an ordinary differential equation. To formulate a precise result define the empirical density by

$$U_i(t) = \frac{1}{N} \sum_{x \in \mathbb{T}_L} 1\left(\xi^\epsilon_{t\epsilon_L^{-2}}(x) = i\right).$$

Theorem 6. *Suppose that $L^2 \ll \epsilon_L^{-2} \ll L^d$. If $U_i(0) \to u_i$ then $U_i(t)$ converges uniformly on compact sets to $u_i(t)$, the solution of*

$$\frac{du_i}{dt} = \phi_i(u) \qquad u_i(0) = u_i$$

where ϕ_i is the reaction term in (14).

Thus in Regime 2, we have "mean-field" behavior, but the reaction function in the ODE is computed using the voter model equilibrium, not the product measure that is typically used in heuristic calculations. The asymptotic behavior of the particle system is now the same as that of the limiting ODE. In particular in case S_2, it will converge to 0 or 1, depending on whether the initial density $u_1 < \bar{u}$ or $u_1 > \bar{u}$.

7.3. Tarnita's formula

Suppose that in addition to the game dynamics each individual switches to a strategy chosen at random from the n possible strategies at rate μ.

Theorem 7. *Suppose that $N^{-2/d} \gg w \gg N^{-1}$. If $\mu \to 0$ and $\mu/w \to \infty$ slowly enough, then in an n-strategy game strategy k is favored by mutation if and only if*

$$\phi_k(1/n, \ldots, 1/n) > 0.$$

Note the similarity to Lemma 1. Intuitively, the change from uniformity will be due to lineages that have one branching event. We do not claim that these conditions are necessary for the conclusion to hold but they are needed for our proof to work. Our next step is to show that we recover the formula in [24] and identify the coefficients.

Birth–Death Dynamics. In this case the limiting PDE is $\partial u_k/\partial t = (1/2d)\Delta u + \phi_i^B(u)$ where

$$\phi_k^B(u) = p_1 \phi_k^R(u) + p_2 \sum_j u_k u_j (G_{k,k} - G_{j,k} + G_{k,j} - G_{j,j})$$

see (16). If we take $u_i \equiv 1/n$ then

$$\begin{aligned} p_1 \phi_k^R(1/n, \ldots, 1/n) &= \frac{p_1}{n}\left(\sum_i G_{k,i}\frac{1}{n} - \sum_{i,j} \frac{1}{n} G_{i,j} \frac{1}{n}\right) \\ &= \frac{p_1}{n}(\hat{G}_{k,*} - \hat{G}) \end{aligned}$$

while the second term in (16) is

$$\frac{p_2}{n}(G_{k,k} - \hat{G}_{*,k} + \hat{G}_{k,*} - \hat{G}_{*,*}).$$

For Birth–Death dynamics (3) holds with $\alpha_1 = p(0|e_1|e_1 + e_2)$ and $\alpha_2 = \alpha_3 = p(0|e_1, e_1 + e_2)$.

Death–Birth Dynamics. In this case the limiting PDE is $\partial u_k/\partial t = (1/2d)\Delta u + \phi_k^D(u)$ where

$$\begin{aligned}\phi_k^D(u) &= \bar{p}_1\phi_k^R(u) + \bar{p}_2 \sum_j u_k u_j (G_{k,k} - G_{j,k} + G_{k,j} - G_{j,j}) \\ &\quad - (1/\kappa)p(v_1|v_2)\sum_j u_k u_j (G_{k,j} - G_{j,k}) \end{aligned} \tag{22}$$

see (17). The computations for the first two terms are as in Birth–Death case with p_i replaced by $\bar{p}_i$. The third term is

$$-\frac{(1/\kappa)p(v_1|v_2)}{n}(\hat{G}_{k,*} - \hat{G}_{*,k}).$$

Thus for Death–Birth dynamics (3) holds with $\alpha_1 = p(v_1|v_2|v_2 + v_3)$, $\alpha_2 = p(v_1|v_2, v_2 + v_3)$, $\alpha_3 = p(v_1|v_2, v_2 + v_3) - (1/\kappa)p(v_1|v_2)$, where $1/\kappa = P(x_1 + x_2 = 0) = \sum_x p(x)p(-x)$ is the effective number of neighbors.

These calculations for Theorem 7 apply to graphs other than the torus. For example, a random r-regular graph looks locally like a tree in which each vertex has r neighbors. Of course the values of the constants for the random regular graph will be different from those on the torus.

8. Proof of Theorem 4

Proof. We will prove only the asymptotic lower bound on the number of 1's. Once this is done, the upper bound follows by interchanging the roles of 0's and 1's. The reaction term $\phi(u) = \lambda u(u - \rho)(1 - u)$ so the limiting PDE satisfies Assumption 1 in [7] with $u_* = u^* = \rho$.

There are constants $0 < v_0 < \rho < v_1 < 1$ and $w, L_i > 0$ so that

(i) if $u(0, x) \geq v_0$ when $|x| \leq L_0$ then $\liminf_{t\to\infty} \inf_{|x|\leq wt} u(t, x) \geq \rho$.

(ii) If $u(0, x) \leq v_1$ when $|x| \leq L_1$ then $\limsup_{t\to\infty} \sup_{|x|\leq wt} u(t, x) \leq \rho$.

In our case if w is chosen small enough we can take the v_i and L_i to be any positive numbers. See Aronson and Weinberger [1,2]. We will take $v_0 = \min\{\rho/2, u_1/2\}$ where u_1 is the density of 1's in the initial product measure.

As in the proof of Theorem 1.4 in [7] on the infinite lattice, we use a block construction. We let K be the largest odd integer so that we can fit K^d adjacent cubes with sides $= 2\epsilon_L^{-1}$ into the torus. Asymptotically we have $K \sim (C/2)L^{1-\alpha}$. Suppose that the origin is in the middle of one of the blocks and call that box I_0. The other blocks can be indexed by $\{-K/2, -K/2 + 1, \ldots, K/2\}^d$. There is some space leftover outside our blocks, so the block construction lattice is not a torus but a flat cube.

To achieve the PDE limit we scale space by multiplying by ϵ_L and speed up time by ϵ_L^{-2}. To define our block event we consider the initial condition for the PDE in which $u(0, x) \geq v_0$ when $|x| \leq 1/2$. Given $\delta > 0$, we choose T large enough so that $u(T, x) \geq \rho - \delta/2$ when $|x| \leq 3$.

As in the hydrodynamic limit, given $r \in (0, 1)$, let $a_\epsilon = \lceil \epsilon^{r-1} \rceil \epsilon$, $Q_\epsilon = [0, a_\epsilon)^d$, and $|Q_\epsilon|$ the number of points in Q_ϵ. For $x \in a_\epsilon \mathbb{Z}^d$ and $\xi \in \Omega_\epsilon$ the space of all functions from $\epsilon\mathbb{Z}^d$ to S let

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

We say that the configuration in the box I_k is good at time t if in each small box $xa_\epsilon + Q_\epsilon$ contained in $k + [-1/2, 1/2]^d$ the density $D_i(x, \xi_t) \geq v_0$, and it is very good if in each small box $xa_\epsilon + Q_\epsilon$ contained in $k + [-1, 1]^d$ we have $D_i(x, \xi_t) \geq \rho - \delta$. It follows from the hydrodynamic limit that

Lemma 3. *Let $\theta > 0$. Suppose the configuration in I_k is good at time mT. If L is large enough then with probability $\geq 1 - \theta/2$ all of the boxes I_ℓ with $\ell - k \in \{-1, 0, 1\}^d$ are very good at time $(m + 1)T$.*

Determining whether that all of the boxes I_k with $k \in \{-1, 0, 1\}^d$ are very good at time T can be done by running the dual processes from all of these points back to time 0. Bounding the dual by a branching random walk and then using exponential estimates it is easy to show:

Lemma 4. *Given $\theta > 0$ and T, there is a constant C so that if L is large enough with probability $\geq 1 - \theta/2$ none of the duals starting in $[-3, 3]^d$ escape from $[-3 - CT, 3 + CT]^d$ by time T.*

If the configuration in I_k is good at time mT, all of the boxes I_ℓ with $\ell - k \in \{-1, 0, 1\}^d$ are very good at time $(m+1)T$, and none of the dual processes escape from $k + [-3 - CT, 3 + CT]^d$ as we work backwards from time $(m + 1)T$ to time mT we set $\eta(k, m) = 1$. If we fail to achieve any one of our goals we set $\eta(k, m) = 0$. If the configuration in I_k is not good at time mT we define $\eta(k, m)$ by an independent Bernoulli that is 1 with probability $1 - \theta$.

These variables $\eta(k, m)$ define for us an oriented site percolation process on the graph $(\mathbb{Z} \bmod K)^d \times \{0, 1, 2, \ldots\}$ in which (k, m) is connected to $(\ell, m + 1)$ when $\ell - k \in \{-1, 0, 1\}^d$. Writing z as shorthand for (k, m) the collection of $\eta(z)$ is "M-dependent with density at least $1 - \theta$" which means that for any k,

$$P(\eta(z_i) = 1|\eta(z_j), j \neq i) \geq (1 - \theta), \quad \text{whenever } z_i, 1 \leq i \leq k \text{ satisfy } |z_i - z_j| > M \text{ for all } i \neq j. \tag{23}$$

It is typically not difficult to prove results for M-dependent percolation processes with θ small (see Chapter 4 of [9]), but here it will be useful to simplify things by applying Theorem 1.3 of Liggett, Schonmann, and Stacy [18] to reduce to the case of independent percolation. By that result, under (23), there is a constant Δ depending on d and M such that if

$$1 - \theta' = \left(1 - \frac{\theta^{1/\Delta}}{(\Delta - 1)^{(\Delta-1)/\Delta}}\right)\left(1 - (\theta(\Delta - 1))^{1/\Delta}\right), \tag{24}$$

we may couple $\eta(z)$ with a family $\zeta(z)$ of iid Bernoulli random variables with $P(\zeta(z) = 1) = 1 - \theta'$ such that $\zeta(z) \leq \eta(z)$ for all z.

With this result in hand we can prove the long time survival of our process on the torus by using

Lemma 5. *Suppose that $\theta < \theta_0$. Start the oriented percolation on $[-K/2, K/2]^d$ with all sites occupied and let τ be the first time all sites are vacant. There is a constant $c > 0$ so that as $K \to \infty$*

$$P(\tau > \exp(cK^d)) \to 1. \tag{25}$$

Mountford [20] proved for the contact process on $[1, N]^d$ that for all $\lambda > \lambda_c$

$$(\log E\tau)/N^d \to \gamma. \tag{26}$$

Earlier he showed, see [19]

$$\tau/E\tau \Rightarrow \mathcal{E} \tag{27}$$

where $\mathcal{E}$ has an exponential distribution with mean 1 and $\Rightarrow$ denotes weak convergence.

Proof of (25). We claim that our result can be proved by the method Mountford used to prove the sharp result in (26) for all $\lambda > \lambda_c$. To explain why the reader should believe this, we note that Lemma 1.1 in [20] concerns connectivity properties of an oriented percolation process in which sites are open with probability $1 - \epsilon_0$ which is then extended to the contact process with $\lambda > \lambda_c$ by using the renormalization argument of Bezuidenhout and Grimmett [4].

To be specific, Mountford, who only gives the details in $d = 2$, shows that if $\lambda > \lambda_c$ there are constants L and $c > 0$ so that for the contact process $\bar{\xi}_t^x$ in $D_n = [1, n] \times [1, 4L]$ starting with a single occupied site at x

$$\inf_{\lambda n^2 \le t \le n^8} \inf_{x,y \in D_n} P(\bar{\xi}_t^x(y) = 1) \ge c > 0.$$

See his Corollary 1.1. This result which is the key to the proof is also true for our oriented percolation process.

The last step in the proof of Theorem 4 is to note that if we start with product measure with a density u_1 of 1 then with high probability all of the I_m are good at time 0. To do this we note that the number of small boxes in the torus is a polynomial in L but the probability of an error in one is $\le \exp(-cL^{(1-r)d})$ where r is the constant used to define the sizes of the little boxes. $\square$

To justify the remark after Theorem 4 we use Proposition 1.2 of Mountford [19]. Write ξ_t^1 for the contact process in $[1, n]^2$ starting from all sites occupied. He shows that there are sequences $a(n), b(n) \to \infty$ with $b(n)/a(n) \to \infty$ so that $P^1(\tau < b(n)) \to 0$ and

$$\inf_{\xi_0} P^{\xi_0}(\xi_{a(n)}^1 = \xi_{a(n)} \text{ or } \tau < a_n) \to 1.$$

In words the process either dies out before time $a(n)$ or at time $a(n)$ agrees with the process starting from all 1's. This idea, which is due to Durrett and Schonmann [12] allows one to prove the limit is exponential by showing that it has the lack of memory property. Unfortunately, Mountford writes sup rather than inf in the conclusion. He cannot mean sup because that is attained by $\xi_0 \equiv 1$ and the probability is 0.

9. Proof of Theorem 5

Proof. Suppose for simplicity that $\epsilon_L^{-1} \sim L^\alpha$. Start a coalescing random walk $\bar{\zeta}_t^L$ with one particle at each site of the torus.

Lemma 6. *Let $\bar{N}_L(t)$ be the number of particles in the coalescing random walks at time t. There is a constant C_1 so that for all L,*

$$E\bar{N}_L(t) \le C_1 N/(1+t) \quad \textit{for all } 0 \le t \le L^{2\alpha}.$$

With high probability, i.e., one that tends to 1 as $L \to \infty$,

$$\bar{N}_L(t) \le 4C_1 N/(1+t) \quad \textit{for all } 0 \le t \le L^{2\alpha}.$$

The constant C_1 is special but all the others C's are not and will change from line to line.

Proof. Let S_t be the first coordinate of our d-dimensional random walk that makes jumps according to p at rate 1.

$$E\exp(\theta S_t) = \exp(t(\phi(\theta)-1)) \quad \text{where } \phi(\theta) = \sum_z e^{\theta z_1} p(z)$$

Since $z \to \exp(\theta z_1)$ is convex and p has mean zero and finite range

$$0 \le \phi(\theta) - 1 \sim \frac{\sigma^2\theta^2}{2} \quad \text{as } \theta \to 0.$$

Let $1/2 < \rho < 1$. If we take $\theta = t^{\rho-1}/2\sigma^2$ and t is large we have

$$P(S_t > t^\rho) \le \exp\left(-\theta t^\rho + \sigma^2\theta^2 t\right) \le \exp(-t^{2\rho-1}/4\sigma^2).$$

Let $\beta \in (\alpha, 1)$ so that $\rho = \beta/2\alpha < 1$, and let $\delta = 2\rho - 1$. Using the last result on each of the d coordinates.

(*) With probability $\ge 1 - N \cdot d\exp(-L^\delta/4\sigma^2)$ no particle that starts outside $[-2L^\beta, 2L^\beta]^d$ will enter $[-L^\beta, L^\beta]^d$ by time $L^{2\alpha}$ and no particle that starts inside $[-2L^\beta, 2L^\beta]^d$ will exit $[-3L^\beta, 3L^\beta]^d$.

Let $\bar\zeta_t$ be the coalescing random walk on $\mathbb{Z}^d$, let $p(t) = P(x \in \bar\zeta_t)$ and let $p_L(t) = P(x \in \bar\zeta_t^L)$. The last two probabilities do not depend on x by translation invariance. (*) implies that

$$|p_L(t) - p(t)| \le N \cdot d\exp(-L^\delta/4\sigma^2) \quad \text{for all } t \le L^{2\alpha}.$$

The result for the expected value now follows from a result of Bramson and Griffeath [5] that shows $p(t) \sim c/t$ as $t \to \infty$.

We begin by proving the second result for each fixed t. A result of Arratia [3], see Lemma 1 on page 913, shows

$$P(x \in \bar\zeta_t, y \in \bar\zeta_t) \le P(x \in \bar\zeta_t)P(y \in \bar\zeta_t),$$

so we have

$$\operatorname{var}(\bar N_t^L) \le N p_L(t)(1 - p_L(t)) \le E\bar N_L(t).$$

Using Chebyshev's inequality, and $\bar N_L(t) \le C_1 N/(1+t)$

$$P(|\bar N_L(t) - E\bar N_L(t)| \ge C_1 N/(1+t)) \le \frac{1+t}{C_1 N} \le \frac{C}{L^{d-2\alpha}}.$$

To complete the proof now let $M = [\log_2 L^{2\alpha}]$. Applying the last result to the $O(\log L)$ values $s_i = 2^i$ with $1 \le i \le M$ we see that

$$P(\bar N_L(s_i) \le 2C_1 N/(1+t) \text{ for } 1 \le i \le M) \ge 1 - \frac{C\log L}{L^{d-2\alpha}}.$$

Using the fact that $t \to \overline{N}_L(t)$ is decreasing and $s_{i+1}/s_i = 2$ we have

$$P(\bar N_L(t) \le 2C_1 N/(1+t/2) \text{ for } 0 \le t \le L^{2\alpha}) \ge 1 - \frac{C\log L}{L^{d-2\alpha}},$$

which proves the desired result. □

The dual for the contact process with fast voting is a branching coalescing random walk. The maximum branching rate is $\|h\|_\infty \epsilon_L^2$, so the expected number of branchings that occur on the space–time set covered by particles in the coalescing random walk is

$$\le 4C_1 \|h\|_\infty \epsilon_L^2 \int_0^{\epsilon_L^{-2}} N/(1+t)\, dt \le CL^{d-2\alpha} \log L.$$

Since the total number of branchings on the space–time set occupied by particles is Poisson, the probability that no branching occurs is

$$\ge \exp(-CL^{d-2\alpha} \log L).$$

Since there are deaths in the graphical representation for the contact process, the rest is easy. With probability $\ge \exp(-CL^{d-2\alpha})$ all of the particles at time $L^{2\alpha}/2$ will be hit by a death by time $L^{2\alpha}$. Thus even if the process is in the all 1's state at time 0, the probability it dies out by time $L^{2\alpha}$ is at least

$$\ge \exp(-C_2 L^{d-2\alpha} \log L) \equiv M.$$

If we are given $M^2 = \exp(2C_2 L^{d-2\alpha} \log L)$ trials the probability we always fail is

$$\le (1 - 1/M)^{M^2} \le \exp(-M).$$

This completes the proof. □

We could try to prove the upper bound on the survival time by arguing that with probability $\ge \exp(-CL^{d-2\alpha})$ all the particles at time $L^{2\alpha}$ land on a 0 in the configuration at time $t - L^{2\alpha}$. In the contact process we can argue this by noting that the state at time $t - L^{2\alpha}$ is dominated by the process starting from all 1 at time $t - 2L^{2\alpha}$. Since $\xi \equiv 1$ is absorbing in evolutionary games, this simple argument is not possible.

10. Proof of Theorem 6

We use the notation for the dual introduced in Section 6. Let R_n denote the increasing subsequence of jump times T_m that are branching times, and $N(t)$ be the number that occur by time t/ϵ^2. Since branching times occur at rate $C_B k \epsilon^2$ when there are k particles in the dual, and the number of particles in the dual is bounded by a branching process in which each particle gives birth to M new particles at rate $C_B \epsilon^2$, the expected number of particles at time t/ϵ^2 is $\le \exp(C_B M t)$ times the number at time 0.

Choose a time K_L so that $K_L/L^2 \to \infty$ and $\epsilon_L^2 K_L \to 0$. This is possible since $L^2 \ll \epsilon_L^{-2}$. If there are k particles at time 0, then the second condition implies

$$P(R_1 \le K_L) \to 0$$

as $L \to \infty$. From this it follows easily that

$$P(R_m - R_{m-1} \le K_L \text{ for some } m \le N(t)) \to 0 \tag{28}$$

as $L \to \infty$ and

$$P(t/\epsilon^2 - R_{N(t)} \le K_L) \to 0. \tag{29}$$

Let $S_n = R_n + L^2 + K_L$.

The next three results, which will be proved at the end of the section, concern the time intervals $[R_n, R_n + L^2]$, $[R_n + L^2, S_n]$ and $[S_n, R_{n+1}]$. In each of the three lemmas we suppose that at time 0 there are k particles in the dual and no two particles are within distance $L^{3/4}$ of each other.

Lemma 7. *Suppose the first particle encounters a birth event in the dual at time* 0*. With high probability* (i) *there is no coalescence between the newborn particles or with their parent after time L^2 and before the next birth event,* (ii) *up to time L^2 there is no coalescence between the new born particles and particles* $2, \ldots, k$*, and* (iii) *at time L^2 all particles are separated by $L^{3/4}$.*

Lemma 8. *Even if we condition on the starting locations, then the particle locations at time K_L are almost independent and uniformly distributed over the torus, i.e., the total variation distance between the random positions and independent uniformly distributed positions tends to* 0*.*

Lemma 9. *With high probability,* (i) *there is no coalescence before the first time a birth event affects a particle, and* (ii) *just before the first birth time the existing particles are separated by $L^{3/4}$.*

Proof of Theorem 6. We will now argue that the three lemmas imply the desired conclusion. (29) implies that with high probability there is no branching event in $[t/\epsilon^2 - K_L, t/\epsilon^2]$. Let $N(t)$ be the last branching time before time t/ϵ^2. By the last remark $S_{N(t)} < t/\epsilon^2$. Lemma 9 implies that no coalescence occurs in the dual in $[S_{N(t)}, t/\epsilon^2]$.

When (28) holds, Lemma 9 implies that the particles at time $R_{N(t)}$ are all separated by $L^{3/4}$ so using Lemma 7 all of the coalescences between the new born particles and their parent occur before $R_{N(t)} + L^2$ and there is no coalescence with other particles during that time interval. At time $R_{N(t)} + L^2$ all the particles are separated by $L^{3/4}$, so Lemmas 8 and 9 imply that there is no coalescence during $[R_{N(t)} + L^2, S_{N(t)}]$ and at time $S_{N(t)}$ the particles are almost uniform over the torus.

The results in the last paragraph imply that when the jump occurs at time $R_{N(t)}$ the joint distribution of the focal site and its neighbors are approximately that of the voter model with the density at time $S_{N(t)}$. Since $\epsilon_L^2 K_L \to 0$ this is almost the same as the density at time $R_{N(t)}$. Working backwards in time and using induction we see that as $L \to \infty$ the dual on the time scale $t\epsilon_L^{-2}$ converges to a limit in which branchings occur at rate $\|h\|_\infty$ and when they occur the joint distribution of the state of the focal site and its neighbors is given by the density at time t.

The last observation implies $EU_i(t)$ converges to a limit $u_i(t)$. To show that the limit satisfies the differential equation we consider $u_i(t+h) - u_i(t)$. When h is small the probability of two or more branching events in the time interval $[t, t+h]$ is $o(h^2)$. By considering the effect of a single event and letting $h \to 0$ we conclude that $du_i/dt = \phi_i(u)$. For more details in the more complicated situation of convergence to a PDE, see Section 2 in Durrett and Neuhauser [11] or Chapter 2 of [7].

The final detail is to show that $\operatorname{var} U_i(t) \to 0$ as $L \to \infty$. To do this, we note that if x and y are separated by a distance greater than $L^{3/4}$ then with high probability their dual processes never intersect before time t/ϵ^2. Writing $\eta_t(x)$ as shorthand for $\xi^\epsilon_{t\epsilon_L^{-2}}(x)$, we see that if $\delta > 0$ and L is large

$$\begin{aligned}\operatorname{var}(U_i(t)) &= \frac{1}{N^2}\sum_{x\in\mathbb{T}_L}\operatorname{cov}(\eta_t(x), \eta_t(y))\\ &\le \frac{c_d L^{3d/4}}{L^d} + \delta.\end{aligned}$$

Since $\delta > 0$ is arbitrary, letting $L \to \infty$, we conclude $\operatorname{var}(U_i(t)) \to 0$. Using Chebyshev's inequality gives the desired result.

10.1. Proofs of the three lemmas

Let $p(x)$ be a finite range, irreducible symmetric random walk kernel on $\mathbb{Z}^d$, some $d \geq 3$, with characteristic function

$$\phi(\theta) = \sum_{x \in \mathbb{Z}^d} e^{i\theta x} p(x), \quad \theta \in [-\pi, \pi]^d.$$

Then ϕ is real-valued, and by PII.5 in Spitzer there is a constant $\lambda > 0$ such that

$$1 - \phi(\theta) \geq \lambda |\theta|^2, \quad \theta \in [-\pi, \pi]^d. \tag{30}$$

Let Z_t be a rate-one continuous time random walk on $\mathbb{Z}^d$ with jump kernel p, starting at 0. Then Z_t has characteristic function

$$\phi_t(\theta) = E_0(e^{i\theta Z_t}) = \exp(-t(1 - \phi(\theta))),$$

and by (30),

$$\phi_t(\theta) \leq e^{-\lambda t |\theta|^2}, \quad \theta \in [-\pi, \pi]^d. \tag{31}$$

Let $Z_t^L = Z_t \bmod L$ be the corresponding walk on the torus $\mathbb{T}_L$.

Proposition 1. (a) *There is a constant $C > 0$ such that*

$$P(Z_t^L = x) \leq C\left(L^{-d} \vee t^{-d/2}\right), \quad t \geq 0, x \in \mathbb{T}_L. \tag{32}$$

(b) *If $s_L \to \infty$ then*

$$\sup_{t \geq L^2 s_L} \max_{x \in \mathbb{T}_L} L^d |P(Z_t^L = x) - L^{-d}| = 0 \tag{33}$$

(b) implies Lemma 8.

Proof. By a standard inversion formula,

$$P(Z_t^L = x) = L^{-d} \sum_{y \in \mathbb{T}_L} \phi_t(2\pi y/L) e^{2\pi i x y/L}.$$

Pulling out the $y = 0$ term and using (31) gives

$$|P(Z_t^L = x) - L^{-d}| \leq L^{-d} \sum_{y \in \mathbb{T}_L, y \neq 0} e^{-\lambda t |2\pi y|^2/L^2}. \tag{34}$$

After bounding the sum by an integral and then changing variables, there is a constant $C' > 0$ such that

$$\begin{aligned} \sum_{y \in \mathbb{T}_L, y \neq 0} e^{-\lambda t |2\pi y|^2/L^2} &\leq C' \int_1^L e^{-\lambda t (2\pi r/L)^2} r^{d-1} dr \\ &\leq C' \left(\frac{L}{2\pi\sqrt{\lambda t}}\right)^d \int_0^\infty u^{d-1} e^{-u^2} du. \end{aligned} \tag{35}$$

Since $C_d = \int_0^\infty r^{d-1}e^{-r}dr < \infty$, (34) and (35) imply

$$P(Z_t^L = x) \le L^{-d} + C_d C'(2\pi\lambda)^{-d} t^{-d/2}$$

which proves part (a).

For part (b), plug any $t \ge L^2 s_L$ into (34) and (35) to get

$$L^d |P(Z_t^x = x) - L^{-d}| \le C_d C'(2\pi\sqrt{\lambda})^{-d} s_L^{-d/2}.$$

Since $s_L \to \infty$ this proves part (b). $\square$

Proposition 2. (a) *If* $1 \ll r_L \le L/2$ *and* $t_L \ll L^d$, *then*

$$\max_{x\in\mathbb{T}_L, |x|\ge r_L} P(Z_s^{L,x} = 0 \text{ for some } s \le t_L) \to 0.$$

(b) *If* $r_L \ll L$ *then*

$$\sup_{t\ge L^2} P(|Z_t^{L,0}| \le r_L) \to 0$$

(b) with $r_L = L^{3/4}$ gives (iii) of Lemma 7 and (ii) of Lemma 9. (a) gives (i) of Lemma 9 and (ii) of Lemma 7. To prove (i) of Lemma 7, combine (iii) of Lemma 7 with (i) of Lemma 9.

Proof. Let $\tau = \inf\{t \ge 0 : Z_t^{L,x} = 0\}$, let $|x| \ge r_L$, and choose $\{t_L'\}$ such that $t_L' \to \infty$ and $t_L' \ll t_L \wedge r_L$. Using a standard martingale argument,

$$P(\tau \in [0, t_L']) \le P(\sup_{0\le t\le t_L'} |Z_t^{L,0}| \ge r_L) \le 2d P(|Z_{t_L'}^0| \ge r_L) \le 2d E(|Z_{t_L'}^0|)/r_L,$$

which tends to 0 because $t_L' \ll r_L$. (Alternatively, we could use the assumption p has finite range and (32) for some $t_L' \to \infty$, which is all that is needed for below.)

Next, by the Markov property and the fact that $Z_t^{L,x}$ is a rate one walk, (32) implies

$$P(\tau \in [t_L', t_L]) \le e\int_{t_L'}^{t_L+1} P(Z_t^x = 0)\,dt \le Ce\int_{t_L'}^{L^2\wedge t_L'} t^{-d/2}dt + Ce\int_{t_L'\wedge L^2}^{t_L'+1} L^{-d}dt.$$

This tends to 0 because $t_L' \to \infty$ and $t_L'/L^d \to 0$, proving (a).

For (b), the bound (32) implies $P(Z_{L^2}^{L,0} = 0) \le CL^{-d}$, so we have

$$P(|Z_{L^2}^{L,0}| \le r_L) \le CL^{-d}(|r_L|+1)^d \to 0$$

which is the desired result. $\square$

11. Proof of Theorem 7

For this result we need to augment the construction with Poisson processes T_n^x, $n \ge 1$ that have rate μ, and random variables V_n^x that are uniform over the strategy set. At time T_n^x the value at x is set equal to V_n^x. Since mutations tell us the value at a site, when we work backward in the dual, we kill a particle when it encounters a mutation. When all of the particles have been killed then we can compute the value of the process at all the sites used to initialize the dual.

Suppose first that there are no mutations. Since w satisfies the conditions for Regime 2, it follows from Theorem 6 that if we run time at rate $1/w$ then in the limit as $L \to \infty$ the density of type k satisfies

$$\frac{du_k}{dt} = \phi_k(u). \tag{36}$$

If we now return to the case with mutations and assume that $\mu/w \to c$ then the limiting equations become

$$\frac{du_k}{dt} = \phi_k(u) + \frac{\mu}{w}(1/n - u_k) \tag{37}$$

so equilibria are solutions of

$$u_k = \frac{1}{n} + \frac{w}{\mu}\phi_k(u).$$

Doing some algebra gives

$$u_k - \frac{1}{n} = \frac{w}{\mu}\phi_k(1/n, \dots, 1/n) + \frac{w}{\mu}(\phi_k(u) - \phi_k(1/n, \dots, 1/n))$$

and hence

$$\left| u_k - \frac{1}{n} - \frac{w}{\mu}\phi_k(1/n, \dots, 1/n) \right| \le \frac{w}{\mu} \left| \phi_k(u) - \phi_k(1/n, \dots, 1/n) \right|.$$

Using the fact that ϕ_k is Lipschitz continuous we conclude

$$\left| u_k - \frac{1}{n} \right| \le C_1 w/\mu,$$

$$\left| u_k - \frac{1}{n} - \frac{w}{\mu}\phi_k(1/n, \dots, 1/n) \right| \le C_2(w/\mu)^2.$$

If $\mu/w \to \infty$ slowly enough then we can use the last result to conclude

$$u_k > 1/n,$$

when w is small giving the desired formula.

References

[1] D.G. Aronson, H.F. Weinberger, Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation, in: Partial Differential Equations and Related Topics, in: Springer Lecture Notes in Math, vol. 446, Springer, New York, 1975, pp. 5–49.

[2] D.G. Aronson, J.F. Weinberger, Multidimensional nonlinear diffusion arising in population genetics, Adv. Math. 30 (1978) 33–76.

[3] R. Arratia, Limiting point processes for rescalings of coalescing and annhilating random walks on $\mathbb{Z}^d$, Ann. Probab. 9 (1981) 909–936.

[4] C. Bezuidenhout, G. Grimmett, The critical contact process dies out, Ann. Probab. 18 (1990) 1462–1482.

[5] M. Bramson, D. Griffeath, Asymptotics for interacting particle systems on $\mathbb{Z}^d$, Z. Wahr. verw. Gebiete. 53 (1980) 183–196.

[6] J.T. Cox, Coalescing random walks and voter model consensus times on the torus $\mathbb{Z}^d$, Ann. Probab. 17 (1989) 1333–1366.

[7] T. Cox, R. Durrett, E. Perkins, Voter model perturbations and reaction diffusion equations, Astérisque 349 (2011). arXiv:1103.1676.

[8] R. Durrett, Spatial evolutionary games with small selection coefficients, Electron. J. Probab. 19 (2014) paper 121.
[9] R. Durrett, Ten lectures on particle systems, in: École d'Été de Probabilités de Saint Flour, in: Lecture Notes Math., vol. 1608, Springer-Verlag, Berlin, 1995, pp. 97–201.
[10] R. Durrett, T.M. Liggett, Y. Zhang, The contact process with fast voting, Electron. J. Probab. 18 (2014) paper 28.
[11] R. Durrett, C. Neuhauser, Particle systems and reaction diffusion equations, Ann. Probab. 22 (1994) 289–333.
[12] R. Durrett, R.H. Schonmann, The contact process on a finite set, ii, Ann. Probab. 16 (1988) 1570–1583.
[13] D.S. Griffeath, Additive and cancellative interacting particle systems, in: Springer Lecture Notes in Math., 1978.
[14] T.E. Harris, On a class of set-valued markov processes, Ann. Probab. 4 (1976) 175–194.
[15] J. Hofbauer, K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge U. Press, 1998.
[16] R. Holley, T. Liggett, Ergodic theorems for weakly interacting systems and the voter model, Ann. Probab. 4 (1975) 195–228.
[17] T.M. Liggett, Stochastic Interacting Systems: Contact, Voter and Exclusion Processes, Springer, New York, 1999.
[18] T.M. Liggett, R.H. Schonmann, A.M. Stacey, Domination by product measures, Ann. Probab. 25 (1997) 71–95.
[19] T.S. Mountford, A metastable result for the finite multidimensional contact process, Canad. Math. Bull. 36 (1993) 216–226.
[20] T.S. Mountford, Existence of a finite constant for finite system extinction, J. Stat. Phys. 96 (1999) 1331–1341.
[21] T. Mountford, J.-C. Mourrat, D. Daniel Valesin, Q. Yao, Exponential extinction time of the contact process on finite graphs. arXiv:1203.2972.
[22] H. Ohtsuki, M.A. Nowak, Evolutionary games on graphs, J. Theoret. Biol. 243 (2006) 86–97.
[23] C.E. Tarnita, H. Ohtsuki, T. Antal, F. Feng, M.A. Nowak, Strategy selection in structured populations, J. Theoret. Biol. 259 (2009) 570–581.
[24] C.E. Tarnita, N. Wage, M. Nowak, Multiple strategies in structured populations, Proc. Natl. Acad. Sci. 108 (2011) 2334–2337.