A simple evolutionary game arising from the study of the role of IGF-II in pancreatic cancer

Ruibo Ma and Rick Durrett

Dept of Math, Duke U., P.O. Box 90320, Durham NC 27708-0320

March 13, 2017

Abstract

We study an evolutionary game in which a producer at x gives birth at rate 1 to an offspring sent to a randomly chosen point in $x + \mathcal{N}_c$, while a cheater at x gives birth at rate $\lambda > 1$ times the fraction of producers in $x + \mathcal{N}_d$ and sends its offspring to a randomly chosen point in $x + \mathcal{N}_c$. We first study this game on the d-dimensional torus $(\mathbb{Z} \bmod L)^d$ with $\mathcal{N}_d = (\mathbb{Z} \bmod L)^d$ and $\mathcal{N}_c =$ the $2d$ nearest neighbors. If we let $L \to \infty$ then $t \to \infty$ the fraction of producers converges to $1/\lambda$. In $d \geq 3$ the limiting finite dimensional distributions converge as $t \to \infty$ to the voter model equilibrium with density $1/\lambda$. We next reformulate the system as an evolutionary game with "birth-death" updating and take $\mathcal{N}_c = \mathcal{N}_d = \mathcal{N}$. Using results for voter model perturbations we show that in $d = 3$ with $\mathcal{N} =$ the six nearest neighbors, the density of producers converges to $(2/\lambda) - 0.5$ for $4/3 < \lambda < 4$. Producers take over the system when $\lambda < 4/3$ and die out when $\lambda > 4$. In $d = 2$ with $\mathcal{N} = [-c\sqrt{\log N}, c\sqrt{\log N}]^2$ there are similar phase transitions, with coexistence occurring when $(1 + 2\theta)/(1 + \theta) < \lambda < (1 + 2\theta)/\theta$ where $\theta = (e^{3/(\pi c^2)} - 1)/2$.

1 Introduction

Archetti, Ferraro, and Christofori [1] have recently analyzed the dynamics of the production of insulin-like growth factor II in tumor cell lines from mouse insulinomas. In this system some (cooperator) cells produce IGF-II while other mutant (defector) cells that have lost both copies of this gene "free-ride" on the growth factors produced by other cells. Thus this system is yet another example of studying the interaction of cooperators and defectors in a spatial system. For a classic example in which cooperators pay a price c to give a benefit b to each of their neighbors see Ohtsuki et al. [19], and Section 1.6 in Cox, Durrett, and Perkins [4].

In our system, space is represented by the d-dimensional lattice or torus $(\mathbb{Z} \bmod L)^d$. We use very simple dynamics that are a variant of the biased voter model. Let $\mathcal{N}_c$ be the competition neighborhood. Producers (2's) give birth at rate 1 and if the birth comes from x it replaces a randomly chosen member of $x + \mathcal{N}_c$. In $d \geq 3$, $\mathcal{N}_c$ will typically be the $2d$ nearest neighbors.

Cheaters (1's) have a diffusion neighborhood $\mathcal{N}_d$ that they use to compute the local density of 2's:

$$u(t, x) = \frac{1}{|\mathcal{N}_d|} |\{y \in \mathcal{N}_d : \xi_t(x + y) = 2\}|,$$

where $|S|$ is the number of elements in a set S. A 1 at x at time t gives birth at rate $\lambda u(t, x)$ and sends the offspring to replace a randomly chosen member of $x + \mathcal{N}_c$. Since we are thinking about diffusion, it would be more natural to replace the simple average by a weighted average using a $p(y)$ that looks like a truncated normal distribution but here we will choose simplicity over realism.

To analyze this system, it is useful to observe that it can be reformulated as an evolutionary game with matrix

$$G = \begin{matrix} & \mathbf{1} & \mathbf{2} \\ \mathbf{1} & 0 & \lambda \\ \mathbf{2} & 1 & 1 \end{matrix} \tag{1}$$

and "birth-death" dynamics. Let $\xi_t(z)$ be the strategy being used by the individual at z at time t. The individual at x has fitness

$$\phi(x) = \frac{1}{|\mathcal{N}_d|} \sum_{y \in \mathcal{N}_d} G(\xi_t(x), \xi_t(y))$$

gives birth at rate $\phi(x)$, and the offspring replaces an individual at a site randomly chosen from $x + \mathcal{N}_c$.

1.1 Homogeneously mixing dynamics

If $\mathcal{N}_d = \mathcal{N}_c = \{1, 2, \ldots N\}$ then in the limit as $N \to \infty$ the frequency of players using strategy i, u_i, follows the replicator equation:

$$\frac{du_i}{dt} = u_i(F_i - \bar{F}) \tag{2}$$

where $F_i = \sum_j G_{i,j} u_j$ is the fitness of strategy i and $\bar{F} = \sum_i u_i F_i$ is the average fitness. See e.g., Hofbauer and Sigmund's book [15]. Note that if we add a constant c_k to column k we add $u_k c_k$ to each F_i and hence also to $\bar{F}$ so the behavior of the replicator equation is not changed. The replicator equation for (1) is

$$\frac{du_2}{dt} = u_2(1 - [u_1 \lambda u_2 + u_2 \cdot 1]) = u_1 u_2(1 - \lambda u_2)$$

As $t \to \infty$, $u_2(t) \to 1/\lambda$, $u_1(t) \to 1 - 1/\lambda$, which is a mixed strategy equilibrium for the game.

1.2 Global diffusion, local competition on the torus

Consider now the system on the torus with $\mathcal{N}_d = (\mathbb{Z} \bmod L)^d$, $\mathcal{N}_c$ the nearest neighbors. For simplicity we suppose the initial state is a product measure with density u_0. Let $N = L^d$,

write $y \sim x$ if y is a nearest neighbor of x, and let

$$U_L(t) = \frac{1}{N} \sum_x 1_{\{\xi_t(x)=2\}},$$

$$V_L(t) = \frac{1}{2dN} \sum_{x,y \sim x} 1_{\{\xi_t(x)=2,\xi_t(y)=1\}}.$$

The system on the torus is difficult to study because its statistics are random, and eventually it will reach one of the absorbing states $\equiv 1$ and $\equiv 0$. To avoid this we will let $L \to \infty$ before taking $t \to \infty$.

Theorem 1. *As $L \to \infty$, $U_L \to u(t)$ and $V_L(t) \to v(t)$ with*

$$\frac{du}{dt} = v(t)(1 - \lambda u(t)) \qquad \text{where } u(0) = u_0 \text{ and } v(0) = u_0(1 - u_0).$$

In the limiting equation $u(t) \to 1/\lambda$ as $t \to \infty$.

To see why this is true, note that the initial density of 2's $u(0) > 1/\lambda$ then the density of 2's will decrease until $u(t) \approx 1/\lambda$. The difficulty in proving this is the usual one in interacting particle systems; to bound the decrease of the one-dimensional distribution $u(t)$, we need information about the two dimensional distribution $v(t)$. A new difficulty is that the dynamics of the dual coalescing branching random walk depend on the density $u(t)$

When the density $u(t) \approx 1/\lambda$ the system behaves like the voter model. Based on Theorem 1 the following should not be surprising. Let ν_p be the limit of the voter model starting with product measure with density p.

Theorem 2. *If $d \geq 3$ then as $L \to \infty$, all of the empirical finite dimensional distributions at time t converge to those of a translation invariant distribution μ_t on $\{1,2\}^{\mathbb{Z}^d}$. As $t \to \infty$, $\mu_t \Rightarrow \nu_{1/\lambda}$.*

Note that Theorem 1 holds in $d \leq 2$. The particle system should cluster in $d \leq 2$, but the first step of our proof which is to show that the convergence $u(t) \to 1/\lambda$ in Theorem 1 occurs exponentially fast uses $d \geq 3$.

1.3 Weak selection $d \geq 3$

To be able to use machinery we have developed previously, [4, 7], we replace the game by $\bar{G} = 1 + wG$ where 1 is a matrix of all 1's and w is small:

$$
\begin{array}{c c c}
 & \mathbf{1} & \mathbf{2} \\
\bar{G} = \mathbf{1} & 1 & 1 + \lambda w \\
\mathbf{2} & 1 + w & 1 + w
\end{array}
$$

Note that by a remark after (2), the replicator equation for $\bar{G}$ is the same as that for G.

If the game matrix is $\mathbf{1}$, then dynamics are those of the voter model, so when w is small this is a voter model perturbation as defined in [4]. To make it easier to compare with [4] and [7] we will let $w = \varepsilon^2$. To be able to just quote the previous results, we will assume

3

that $\mathcal{N}_c = \mathcal{N}_d = \mathcal{N}$. One can prove results when $\mathcal{N}_c \neq \mathcal{N}_d$ but there are a number of small changes in the proof and some of the symmetry that is useful in simplifying formulas is lost.

The key to the study of voter model perturbations is a result that says when the system is suitably rescaled in space and time it converges to the solutions of a reaction diffusion equation. We run time at rate ε^{-2} so that the perturbation will have an effect, and scale space by ε. That is, we look at

$$\bar{\xi}_t^\varepsilon(x) = \xi_{t\varepsilon^{-2}}(x/\varepsilon) \qquad \text{for } x \in \varepsilon\mathbb{Z}^d.$$

The last detail before we can state the result is to define the mode of convergence. Pick a small $r > 0$ and divide $\varepsilon\mathbb{Z}^d$ into boxes with side ε^r. Given an $x \in \mathbb{R}^d$ let $B^\varepsilon(x)$ be the box that contains x, and let $\bar{u}_i^\varepsilon(t,x)$ be the fraction of sites in $B^\varepsilon(x)$ in state i at time $t\varepsilon^{-2}$. We say that the rescaled spatial model $\bar{\xi}_t^\varepsilon$ converges to $u(t,x)$ if for any L

$$\sup_{x \in [-L,L]^d} |\bar{u}_i^\varepsilon(t,x) - u(t,x)| \to 0 \quad \text{as } \varepsilon \to 0.$$

Theorem 3. *Suppose $d \geq 3$. Let $v_i : \mathbb{R}^d \to [0,1]$ be continuous with $\sum_{i \in S} v_i = 1$. If the initial condition $\bar{\xi}_0^\varepsilon$ converges to v_i in the sense described above then $\bar{\xi}_t^\varepsilon$ converges to $u(t,x)$ the solution of the system of partial differential equations:*

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

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

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

where $h_{i,j}(0,\xi)$ is the rate 0 flips from i to j in the evolutionary game when the configuration is ξ. The brackets are expected value with respect to the stationary distribution ν_u for the d-dimensional nearest neighbor voter model in which the densities are given by the vector u.

To give a formula for the reaction term in the case of a k-strategy evolutionary games with weak selection, we use results in Section 12 of [7]. Let v_1 and v_2 be independent and uniform over $\mathcal{N}$. Let

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

The reaction term is p_1 times the right-hand side of the replicator equation for $H = G + A$ where

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

Note that if we add c_k to column k the perturbation matrix A is not changed, so if we subtract 1 from the second column of G the reaction term in our situation is p_1 times the right-hand side of the replicator equation for

$$
\begin{array}{ccc}
 & \mathbf{1} & \mathbf{2} \\
H = \mathbf{1} & 0 & (1+\theta)(\lambda-1) - \theta \\
\mathbf{2} & 1 + \theta - \theta(\lambda-1) & 0
\end{array}
$$

4

If we suppose that $\mathcal{N}$ = the nearest neighbors of 0 then $\theta \approx 0.5$, see page 13 of [7], and the game becomes

$$
\begin{array}{ccc}
 & \mathbf{1} & \mathbf{2} \\
\mathbf{1} & 0 & (3/2)\lambda - 2 \\
\mathbf{2} & 2 - \lambda/2 & 0
\end{array}
$$

From this and results in Section 6 of [7], we see that when w is small

- If $\lambda < 4/3$ then $(3/2)\lambda - 2 < 0$ so strategy 2 dominates strategy 1 and the 1's die out.

- If $\lambda > 4$ then $2 - \lambda/2 < 0$ so strategy 1 dominates strategy 2 and the 2's die out.

- If $4/3 < \lambda < 4$, the replicator equation converges to the mixed strategy equilibrium

$$
(\rho, 1 - \rho) = (1.5 - 2/\lambda, 2/\lambda - 0.5).
$$

  It follows that there is coexistence in the spatial game and in all stationary distributions that assign probability 1 to configurations with infinitely many 1's and infinitely many 2's, the probability that x is in state 1 is close to ρ.

The next set of simulations, done by Mridu Nanda, a student at the North Carolina School for Science and Math, shows that the theory accurately describes the behavior of the spatial game when $w = 1/10$ and works reasonably well even when $w = 1/2$. The numbers in the next table give the equilibrium frequencies of strategy 1 for the indicated values of w and λ and compare them with the theoretical predictions about the limit $w \to 0$

λ	4/3	3/2	3	3.5	4
Original game	0.11	0.25	0.75	0.83	0.89
$w = 1/2$	0.01	0.19	0.79	0.88	0.96
$w = 1/10$	0.00	0.16	0.82	0.92	0.98
$w \to 0$ limit	0	0.17	0.83	0.93	1

1.4　Two dimensions

One of the drawbacks of the voter perturbation machinery is that it requires the existence of a stationary distribution for each vector of densities and hence cannot be used in two dimensions, where the only stationary distributions for the voter model concentrate on absorbing states $\xi(x) \equiv i$. To overcome this problem, one can note that for two dimensional nearest neighbor random walk, the probability a random walk does not hit 0 by time t is asymptotically $c/\log t$ and then run time at $\varepsilon^{-2} \log(1/\varepsilon)$ to have particles created at rate $O(1)$ that don't coalesce before a fixed finite time T. However, the fact that most particles do not escape coalescence brings a number of technical difficulties and there are only a few systems that have been rigorously analyzed, see [9, 11, 5]. A second unfortunate fact is that the probability that after a branching event that produces three particles, the probability none of them coalesce by time t is $O(1/\log^3 t)$ compared to $O(1/\log t)$ for a pair of particles, see [5], so the limiting PDE can only have quadratic reaction terms.

Here will follow in the footsteps of Ted Cox [2] and suppose that when $w = 1/N$,

$$\mathcal{N}_d = \mathcal{N}_c = [-c\sqrt{\log N}, c\sqrt{\log N}]^2.$$

As we will show in Section 4 this is enough to make our voter random walks transient in the sense that a random walk starting at 0 has positive probability of not returning to 0 by time $N \log N$. In a sense we are making a large range assumption but $\sqrt{\log N}$ grows slowly so our results should be relevant for processes with fixed finite range. For example if $N = 10^6$, $\sqrt{\log N} = 3.717$.

One should be able to prove results for general "long-range" voter model perturbations in $d = 2$, but for simplicity and concreteness, we will only consider k-strategy evolutionary games with weak selection.

Theorem 4. *Suppose $d = 2$ and the interaction neighborhood is $Q_N = [-c\sqrt{\log N}, c\sqrt{\log N}]^2$. If we scale space by dividing by $\sqrt{N \log N}$ and run time at rate N the spatial game converges (in the sense described above) to the solution of the system of partial differential equations:*

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

Let $b = 3/(2\pi c^2)$, $p_1 = e^{-3b}$ and $p_2 = (e^{-b} - e^{-3b})/2$. The reaction term is p_1 times the right-hand side of the replicator equation for $H = G + A$ where

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

If we consider the special case in (1) and, as before, subtract 1 from the second column (which does not effect A or the behavior of the replicator equation for H) we see that the reaction term is p_1 times the right-hand side of the replicator equation for

$$
\begin{array}{ccc}
 & \mathbf{1} & \mathbf{2} \\
H = \mathbf{1} & 0 & (1+\theta)\lambda - (1+2\theta) \\
\mathbf{2} & (1+2\theta) - \lambda\theta & 0
\end{array}
$$

but now $\theta = p_2/p_1 = (e^{2b} - 1)/2$. The qualitative behavior is same as in $d \geq 3$ but the locations of the phase transitions have changed. The mixed strategy equilibrium has

$$\rho_1 = (1+\theta) - (1+2\theta)/\lambda, \qquad \rho_2 = (1+2\theta)/\lambda - \theta$$

so there is coexistence when

$$\frac{1+2\theta}{1+\theta} < \lambda < \frac{1+2\theta}{\theta}$$

When $\theta = 1/2$ this reduces to the previous answer.

In the analysis above, the long range assumption is needed for the proof but we do not think it is necessary for the conclusion. See the simulation in Figure 1. The fact that the interfaces between the regions occupied by the two strategies break down suggests that there is coexistence in the spatial model. For an explanation of the heuristic see the analysis of the nonlinear voter model done by Molofsky et al [18]
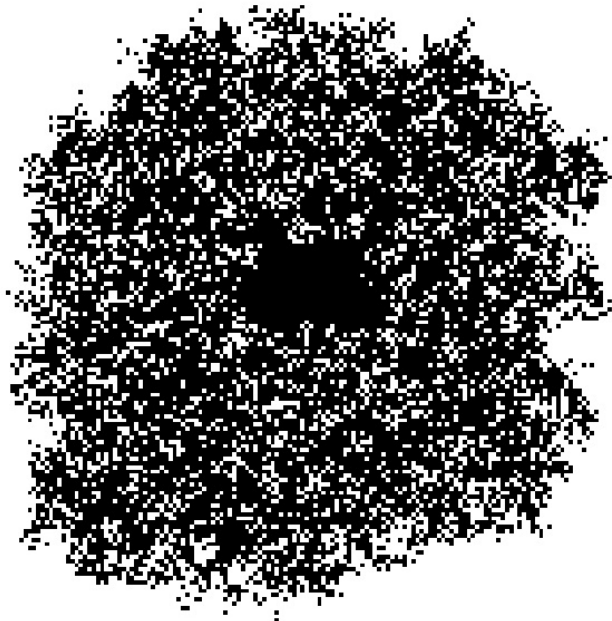.

Figure 1: In this simulation 1's are white and 2's are black. The initial state was a black square in a sea of white.

2   Proof of Theorem 1

The first step is to construct the process from a graphical representation. To make our process look more like the biased voter model, we will change the notation for producers from 2 to 0. In the new notation

$$U_L(t) = \frac{1}{N} \sum_x 1_{\{\xi_t(x)=0\}}$$

Suppose that $U_L(0) > 1/\lambda$. We will use a biased voter model type construction that only works up to time $T_0(L) = \inf\{t : U_L(t) < 1/\lambda\}$, but Lemma 2.2 will show that $T_0(L) \to \infty$ as $L \to \infty$.

For each $x \in (\mathbb{Z} \bmod L)^d$ and nearest neighbor y we have a Poisson process $T_n^{x,y}$, $n \geq 1$ with rate $1/2d$. At time $T_n^{x,y}$ we draw an arrow from x to y and write a δ at y to indicate that y will take on the "opinion" at x at that time.

For each $x \in (\mathbb{Z} \bmod L)^d$ and nearest neighbor y we have a Poisson process $S_n^{x,y}$, $n \geq 1$ with rate $(\lambda - 1)/2d$ and a collection of independent random variables $R_n^{x,y}$ that are uniform on $[0, \lambda - 1]$. At time $S_n^{x,y}$ we draw an arrow from x to y if $\lambda U_L(S_n^{x,y}) - 1 \geq R_n^{x,y}$. These arrows will create a 1 at y if there is a 1 at x.

Given an initial condition $\xi_0(x)$ we view the $\{x : \xi_0(x) = 1\}$ are sources of fluid. The fluid moves up the graphical representation, being blocked by δ's and moving across arrows in the direction of their orientation. In an arrow-δ the δ occurs just before the arrival of the

arrow, otherwise the arrow would do nothing. It is easy to see that this approach, which goes back to Harris [14] and Griffeath [13], constructs the process and the density $U_L(s)$, $0 \le s \le t$.

As with the ordinary biased voter model we can for each $x \in (\mathbb{Z} \bmod L)^d$ define a set-valued dual processes $\zeta_s^{x,t}$ for $0 \le s \le t$. If one of the particles in $\zeta_s^{x,t}$ is at y and there is an arrow-δ from x to y at time $t-s$ then the particle jumps to x. If instead it encounters an arrow from x to y then the particle at y gives birth to a new particle at x. If the jumping particle or the new born lands on an occupied site the two coalesce to 1. From the definition of the dual we see that

$$\{\xi_t(x) = 0\} = \{\xi_0(y) = 0 \text{ for all } y \in \zeta_s^{x,t}\}$$

If the initial state is a product measure with density u_0 then

$$P(\xi_t(x) = 0) = E\left[u_0^{|\zeta_t^{x,t}|}\right]$$

Let $N = L^d$. On $(\mathbb{Z} \bmod L)^d$, we have the following differential equation for $u_L(t) = EU_L(t)$:

$$
\begin{aligned}
\frac{du_L(t)}{dt} &= \frac{d}{dt} \frac{1}{N} \sum_x P(\xi_t(x) = 0) \\
&= E\frac{1}{N}\left(-\sum_{y \sim x} 1_{\{\xi_t(x)=0, \xi_t(y)=1\}} \frac{\lambda U_L}{2d} + \sum_{y \sim x} 1_{\{\xi_t(x)=1, \xi_t(y)=0\}} \frac{1}{2d}\right) \quad (3) \\
&= E[V_L(t)(1 - \lambda U_L(t))]
\end{aligned}
$$

where $V_L(t) = \frac{1}{2Nd} \sum_{y \sim x} 1_{\{\xi_t(x)=0, \xi_t(y)=1\}} \le 1$. Let $v_L(t) = EV_L(t)$. Note that the capital letters are random variables while the lower case letters are their expected values.

Our first goal is to show that $U_L(t) \to u(t)$ and $V_l(t) \to v(t)$. To prove this the following lemma will be useful.

Lemma 2.1. *Given $u(s)$ with $u(0) = u_0$ and $u(s) \ge 1/\lambda$ for all $s \ge 0$ define for each t a coalescing branching random walk (CBRW) ζ_s^t on $\mathbb{Z}^d$ in which branching at time s occurs at rate $\lambda u(t-s) - 1$. There is a unique u with $u(0) = u_0$ so that for all t*

$$u(t) = E\left[u_0^{|\zeta_s^t|}\right].$$

Proof. Suppose u_1 and u_2 are two solutions with $u_1(0) = u_2(0) = u_0$. Fix t and let ζ_s^1, ζ_s^2 be the corresponding CBRWs defined for $0 \le s \le t$. Since $x \to u_0^x$ is Lipschitz continuous on $[1, \infty)$ with Lipschitz constant 1

$$|u_1(t) - u_2(t)| = \left|E\left[u_0^{|\zeta_s^1|}\right] - E\left[u_0^{|\zeta_s^2|}\right]\right| \le E||\zeta_t^1| - |\zeta_t^2||$$

Let $v(t) = \min\{u_1(t), u_2(t)\}$ and $w(t) = \max\{u_1(t), u_2(t)\}$. Let ζ_t^v and ζ_t^w be the corresponding CBRWs. Clearly, these processes can be constructed on the same space so that:

$$|\zeta_s^v| \le |\zeta_s^1|, |\zeta_s^2| \le |\zeta_s^w|$$

8

and hence
$$\left||\zeta_s^1| - |\zeta_s^2|\right| \leq |\zeta_s^w| - |\zeta_s^v|$$

Let z_t^w and z_t^v be the corresponding BRWs with birth rates $\lambda w - 1$ and $\lambda v - 1$. If we couple the births and random walk steps in the natural way then

$$|\zeta_s^w| - |\zeta_s^v| \leq |z_s^w| - |z_s^v|$$

The branching processes have

$$\frac{d}{dt}(E|z_t^w| - E|z_t^v|) = \lambda(w(s) - v(s))$$

Integrating and the fact that $|u_1(t) - u_2(t)| = w(t) - v(t)$

$$E|z_s^w| - E|z_s^v| \leq \lambda \int_0^t |u_1(s) - u_2(s)| \, ds$$

Combining our calculations we see that

$$|u_1(t) - u_2(t)| \leq \lambda \int_0^t |u_1(s) - u_2(s)| \, ds$$

Gronwall's inequality then implies $u_1(t) = u_2(t)$. $\qquad\square$

Lemma 2.2. *If $u_L(0) = u_0$ for all L then as $L \to \infty$, $u_L(t) \to u(t)$, the function from Lemma 2.1.*

Proof. Since $V_L \leq 1$, it follows from (3) that $du_L/dt \geq 1 - \lambda u_L(t)$, so $u_L(t) > 1/\lambda$ for all t, Using (3) again we can see that $u_L(t)$ is Lipschitz continuous with Lipschitz constant 1. This implies that $u_L(t)$ is tight as a sequence of continuous functions in $C[0, \infty)$. It is easy to see that any subsequential limit will satisfy the conditions of Lemma 2.1. Since the solution is unique, the sequence converges. $\qquad\square$

Our next step is to show that the variances of $U_L(t)$ and $V_L(t)$ tend to 0. To start to do this we will prove a random walk estimate. For $x = (x_1, ..., x_d) \in (\mathbb{Z} \bmod L)^d$, define the distance from 0 by
$$|x| = \max_k \min\left(|x_k|, L - |x_k|\right).$$

Lemma 2.3. *Suppose S_t is a continuous time random walk on $(\mathbb{Z} \bmod L)^d$ with exponential rate 2 starting from 0. For all $t, \delta > 0$*

$$\lim_{L \to \infty} P\left(|S_t| \geq L^\delta - 2\right) = 0$$

Proof. In order to have $|S_t| \geq L^\delta - 2$ we must have $|S_t^i| \geq L^\delta - 2$ for some i, so it suffices to prove the result for a one dimensional random walk. Let $\varphi(\theta) = (e^\theta + e^{-\theta})/2$ be the moment generating function for one step. Steps in the direction of the ith component happen at rate $2/d$. Using Chebyshev's inequality, we see that if $\theta > 0$

$$e^{\theta(L^\delta - 2)} P\left(S_t \geq L^\delta - 2\right) \leq Ee^{\theta S_t} = \sum_{k=0}^\infty e^{-2t/d} \frac{(2t/d)^k}{k!} \varphi(\theta)^k = e^{2t(\varphi(\theta) - 1)/d}.$$

9

Taking $\theta = 1$ and rearranging

$$P(S_t \geq L^\delta - 2) \leq \exp((2t/d)(\varphi(1) - 1) - L^\delta + 2) \to 0, ,$$

as $L \to \infty$, which proves the desired result. $\qquad\square$

Lemma 2.4. *As $L \to \infty$, $var\,(U_L(t)) \to 0$ and $var\,(V_L(t)) \to 0$.*

Proof. Let $A = \{\xi_t(x) = 0\}$, $B = \{\xi_t(y) = 0\}$. If $|x - y| \leq 2L^\delta$ we use $|\text{cov}\,(1_A, 1_B)| \leq 1$. To bound the covariance when $|x - y| > 2L^\delta$, we use an old trick due to David Griffeath [13]. We construct the dual process $\zeta_s^{x,t}$ on graphical representation #1, and the dual process $\zeta_s^{y,t}$ on an independent graphical representation #2. To have the correct joint distribution we adopt the priority rule that if a particle z in $\zeta_s^{y,t}$ occupies the same site as a particle in $\zeta_s^{x,t}$, the graphical representation #1 is used for moves and births from z in $\zeta_s^{y,t}$. Let C be the event that the duals starting from x and y do not intersect. Let 1_B^* be the indicator of the event that B occurs when only graphical representation #2 is used. Since 1_A and 1_B^* are independent

$$|\text{cov}\,(1_A, 1_B)| = |E[1_A 1_B - 1_A 1_B^*]| \leq P(C^c)$$

To bound $P(C^c)$ let $\pi_{i,j}^t$ be the probability that by time t the dual starting from x branches i times and the one starting from y branches j times. Breaking things down according to the values of i and j

$$P\,(C^c) \leq \sum_{i,j \geq 0} \pi_{i,j}^t (i+1)(j+1) P\,(|S_t| \geq L^\delta - 2)$$

$$= P\,(|S_t| \geq L^\delta - 2) \sum_{i,j \geq 0} \pi_{i,j}^t (i+1)(j+1) \to 0,$$

as $L \to \infty$, because comparison with a branching process shows $\sum_{i,j \geq 0} \pi_{i,j}(i+1)(j+1) < \infty$. To bound the variance now we note that

$$\text{var}\,(U_L(t)) \leq \frac{1}{N^2}[N \cdot (2L)^{d\delta} + N^2 P(C^c)] \to 0$$

as $N \to \infty$, The argument for $V_L(t)$ is almost the same except that now four dual processes are involved. $\qquad\square$

Lemma 2.5. *There is a $v(t)$ so that $V_L(t) \to v(t)$.*

Proof. Lemmas 2.2 and 2.4 imply that $U_L(t) \to u(t)$. This implies that the dual process has a limit. It follows from this that $EV_L(t)$ converges to a limit that we call $v(t)$ and the result follows from another application of Lemma 2.4 . $\qquad\square$

To prove Lemmas 2.1–2.5 when $u(0) < 1/\lambda$, we use a different graphical representation. For each $x \in (\mathbb{Z} \bmod L)^d$ and nearest neighbor y we have a Poisson process $T_n^{x,y}$, $n \geq 1$ with rate $1/2d$, and a collection of independent random variables $R_n^{x,y}$ that are uniform on $[0,1]$. At time $T_n^{x,y}$, if $\lambda U_L(T_n^{x,y}) \geq R_n^{x,y}$, then we draw an arrow from x to y and write a δ at y to indicate that y will take on the "opinion" at x at that time. Otherwise, we draw an arrow from x to y. These arrows will create a 0 at y if there is a 0 at x.

With this new dual, we have $1 - u(t) = E(1 - u(0))^{N(t)}$, where $N(t)$ is the number of particles at time t. The proof of Lemma 2.1 goes through with minor changes. Lemma 2.2 follows as before. Lemma 2.3 is a random walk estimate. The proof of the covariance estimate Lemma 2.4 needs only minor changes and then Lemma 2.5 follows as before.

2.1 Convergence of $u(t)$ to $1/\lambda$

Again we begin with the case $u(0) > 1/\lambda$. Using (3) with Lemmas 2.2 and 2.5 we have

$$\frac{du}{dt} = v(t)\left(1 - \lambda u(t)\right),$$

Lemma 2.6. *Let $u(t)$ be the global density of producer cells at time t, and suppose the system starts from the product measure with $u(0) > 1/\lambda$. Then $u(t) \to 1/\lambda$ as $t \to \infty$.*

Proof. Take an $\varepsilon > 0$. We want to find a constant, C_ε such that if $u \geq (1/\lambda) + \varepsilon$ on $[0, C_\varepsilon)$, then $u(C_\varepsilon) < 1/\lambda$. First pick m, s.t. $u(0)^m < 1/(2\lambda)$. The branching rate of the dual process is bounded from below by $\varepsilon\lambda$. Hence there is a $K > 0$, which does not depend on ε, such that $C_\varepsilon = K/(\varepsilon\lambda)$ satisfies $P(N\left(C_\varepsilon\right) < m) < 1/(2\lambda)$. Then

$$u(C_\varepsilon) = Eu(0)^{N(C_\varepsilon)} \leq u(0)^m + P(N(C_\varepsilon) < m) < 1/\lambda.$$

Thus, the density $u(t)$ cannot stay away from $1/\lambda$. Since $1/\lambda$ is an equilibrium for the ODE of $u(t)$, the convergence is established by contradiction. □

Up to this point all of our calculations are valid in any dimension. We will now show that if $d \geq 3$, the convergence mentioned in the lemma occurs exponentially fast. The next lemma controls the covariance between neighbors. Let $p_t(x|y)$ be the probability that walks starting from x and y do not collide by time t.

Lemma 2.7. *Suppose that the initial distribution is product measure with $u(0) > 1/\lambda$ and $\int_0^t (\lambda u(s) - 1)\, ds \leq 1$. If $\delta_0 = [u(0) - u(0)^2]e^{-2}p(0|e_1)$ then*

$$\sup_{x \neq y} P(\xi_t(x) = 0, \xi_t(y) = 0) - P(\xi_t(x) = 0) \leq -\delta_0$$

Proof. Let $\zeta_t^{A,t}$ be the dual coalescing random walk starting from A occupied at time t. Let $N_x = |\zeta_t^{x,t}|$ and $N_{x,y} = |\zeta_t^{\{x,y\},t}|$. We have

$$P(\xi_t(x) = 0) = Eu(0)^{N_x} \qquad P(\xi_t(x) = 0, \xi_t(y) = 0) = Eu(0)^{N_{x,y}}$$

Notice that $N_{x,y} \geq N_y$ so $u(0)^{N_{x,y}} - u(0)^{N_x} \leq 0$. Let G be the event that the duals starting from x and y do not branch and the random walks starting from x and y do not hit. Since $\int_0^t (\lambda u(s) - 1)\, ds \leq 1$

$$P(G) \geq e^{-2}p_t(x|y) \geq e^{-2}p(0|e_1)$$

Combining our estimates

$$P(\xi_t(x) = 0, \xi_t(y) = 0) - P(\xi_t(x) = 0) \leq E[u(0)^{N_{x,y}} - u(0)^{N_x}; G]$$
$$\leq [u(0)^2 - u(0)]e^{-2}p(0|e_1) = -\delta_0$$

which completes the proof. □

We will now combine the last two lemmas to prove exponential convergence. Let $T_m = \inf\{t : u(t) \leq 1/\lambda + 2^{-m}\}$. By Lemma 2.6, $T_m < \infty$.

Lemma 2.8. *Let* $t_0 = 2/[\lambda p(0|e_1)\delta_0]$. *If* $\int_0^{T_m+t_0}(\lambda u(s) - 1)\,ds \le 1$ *then* $T_{m+1} - T_m \le t_0$.

Proof. Suppose $T_{m+1} - T_m \ge t_0$. Since $u(s) \ge 1/\lambda + 2^{-(m+1)}$ on $[T_m, T_m + t_0]$ the probability that the dual has a branching event and the two particles do not coalesce is $\ge (1 - e^{-\lambda 2^{-(m+1)}})p(0|e_1)$ so using Lemma 2.7

$$u(T_m + t_0) - u(T_m) \le (1 - e^{-\lambda 2^{-(m+1)}t})p(0|e_1)(-\delta_0)$$
$$\le -\lambda 2^{-(m+2)}p(0|e_1)\delta_0 t_0 = -2^{-(m+1)}$$

so we have $T_{m+1} \le T_m + t_0$ and the proof is complete. $\qquad\square$

Proof of Theorem 1 when $u(0) > 1/\lambda$. Lemma 2.8 implies that

$$\int_{T_m}^{T_{m+1}}(\lambda u(s) - 1)\,ds \le \lambda 2^{-m}t_0 = 2^{1-m}/[p(0|e_1)\delta_0]$$

From this we see that if $u(0) = 1/\lambda + 2^{-M}$ and M is large enough then

$$\int_0^t (\lambda u(s) - 1)\,ds \le 1$$

for all t and we have $T_{M+k} \le kt_0$ for all k which proves exponential convergence. $\qquad\square$

When $u(0) < 1/\lambda$ the branching rate is different and we need to look at $1 - \lambda u(t)$, but otherwise the proofs go through as before. This completes the proof of Theorem 1.

3   Proof of Theorem 2

The first step is to show that the the empirical finite dimensional distributions converge to those of a translation invariant measure μ_t. The convergence of their means follows from the proof of Lemma 2.5. Their variance can be shown to go to 0 using the proof of Lemma 2.4.

In this section we will show that if t is large, then the finite dimensional distribution of μ_t are close to those of the voter model stationary distribution $\nu_{1/\lambda}$. To begin we assume $u(0) > 1/\lambda$. The next lemma bounds the covariance of well separated sites.

Lemma 3.1. *Fix* $T, \delta > 0$. *There is an* $R_2 > 0$, *so that when* $|x - y| > R_2$,

$$|P(\xi_T(x) = 0, \xi_T(y) = 0) - u(T)^2| \le \delta.$$

Proof. In the dual process, the branching rate is bounded by $\lambda - 1$. Let $Z_t(x)$ be the number of particles at x at time t in a branching random walk that starts with one particle at 0, jumps at rate 1 and branches at rate $\lambda - 1$. Let $m_t(x) = EZ_t(x)$. By considering the rates at which things happen we see that $m_t(x)$ satisfies

$$\frac{dm_t(x)}{dt} = -m_t(x) + \lambda \sum_y m_t(y)p(y,x). \qquad (4)$$

Note the the second term accounts for jumps at rate 1 and branching at rate $\lambda - 1$. Let S_t be a continuous-time random walk starting from the origin with jumps at rate 1. We will show

$$m_t(x) = e^{(\lambda-1)t} P(S_{\lambda t} = x). \tag{5}$$

Both sides agree at time 0. Thus it suffices to show the RHS satisfies (4).

$$\frac{d}{dt} e^{(\lambda-1)t} P(S_{\lambda t} = x) = (\lambda - 1)e^{(\lambda-1)t} P(S_{\lambda t} = x) + e^{(\lambda-1)t} \frac{d}{dt} P(S_{\lambda t} = x)$$

$$= (\lambda - 1)e^{(\lambda-1)t} P(S_{\lambda t} = x)$$

$$+ e^{(\lambda-1)t} \left(-\lambda P(S_{\lambda t} = x) + \lambda \sum_{y \sim x} P(S_{\lambda t} = y) p(y, x) \right)$$

$$= -e^{(\lambda-1)t} P(S_{\lambda t} = x) + \lambda \sum_{y \sim x} e^{(\lambda-1)t} P(S_{\lambda t} = y) p(y, x)$$

Thus we have shown (5) satifies (4).

We can bound the decay of $m_t(x)$ by using the argument in Lemma 2.3. Again it suffices to consider $d = 1$.

$$e^x P(S_{\lambda t} = x) \leq E e^{S_{\lambda t}} = \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \varphi(1)^k = e^{\lambda t(\varphi(1)-1)}.$$

which implies $m_t(x) \leq e^{t(\lambda\varphi(1)-1)-x}$, i.e., $m_t(x)$ decays exponentially in x. Note that branching in the actual dual always has a lower rate and the branching random walk ignores coalescence, so $m_t(x)$ gives an upper bound of the probability that the dual has a particle at x at time t.

Using the reflection principle

$$P(\zeta_t^{0,T} \cap [x, \infty) \times \mathbb{Z}^{d-1} \neq \emptyset \text{ for some } t \leq T) \leq 2 \sum_{y=x}^{\infty} e^{T(\lambda\varphi(1)-1)-y} \leq C_T e^{-x}$$

Considering all of the coordinates, we see that if R_2 is large and $D = \{x : \|x\|_\infty \leq R_2/2\}$

$$P(\zeta_t^{0,T} \cap D^c \text{ for some } t \leq T) \leq \delta/4$$

This implies that if $|x - y| \geq R_2$ then

$$P(\zeta_t^{x,T} \cap \zeta_t^{y,T} = \emptyset \text{ for all } t \leq T) \geq 1 - \delta/2$$

When this occurs we say that the duals starting from x and y do not collide. We denote the event by B.

Let $A = \{\xi_T(x) = 0, \xi_T(y) = 0\}$.

$$|P(A) - u(T)^2| = |P(A \cap B) + P(A \cap B^c) - u(T)^2|$$
$$= |P(B)u(T)^2 + P(A \cap B^c) - u(T)^2|$$
$$\leq u(T)^2 P(B^c) + P(B^c) \leq 2P(B^c) \leq \delta,$$

which proves the desired result. □

Our next step is to generalize Lemma 3.1 to m sites.

Lemma 3.2. *Fix $T, \delta > 0$. Suppose we have sites x_k, $1 \le k \le m$ with $m \ge 3$. There is an $R_m > 0$, so that if $|x_i - x_j| > R_m$, for $1 \le i < j \le m$, we have*

$$|P(\xi_T(x_k) = 0 \text{ for } 1 \le k \le m) - u(T)^m| \le \delta.$$

Proof. Let $A_m = \{\xi_T(x_k) = 0 \text{ for } 1 \le k \le m\}$, and B_m be the event that the duals starting from $x_1, \ldots, x_m$ do not collide. By Lemma 3.1, there is an R_m so that if $|x - y| > R_m$, $P(B_2^c) \le \delta/m^2$. From this it follows that $P(B_m^c) \le \delta/2$. Computing as in the previous lemma gives

$$
\begin{aligned}
|P(A_m) - u(T)^m| &= |P(A_m \cap B) + P(A_m \cap B^c) - u(T)^m| \\
&= |P(B)u(T)^m + P(A_m \cap B^c) - u(T)^m| \\
&\le u(T)^m P(B^c) + P(B^c) \le 2P(B^c) \le \delta
\end{aligned}
$$

For the next proof it is convenient to note that if we take R_m to be as small as possible $m \to R_m$ is increasing. $\qquad\square$

To show the convergence of finite dimensional distributions, it is enough to consider probabilities that a set of sites are all in state 0, because all the finite dimensional distributions can be computed from these values.

Lemma 3.3. *Let $\varepsilon > 0$. Suppose we have sites x_k, $1 \le k \le m$. If $\varepsilon > 0$, then there is a $t_1 > 0$, for all time $t \ge t_1$,*

$$\left| P(\xi_t(x_k) = 0 \text{ for } 1 \le k \le m) - \nu_{1/\lambda}(\eta(x_k) = 0 \text{ for } 1 \le k \le m) \right| < \varepsilon.$$

Proof. Since we have proved exponential convergence of $u(t) \to 1/\lambda$, can choose T, so that

$$|u(T)^k - (1/\lambda)^k| \le \frac{\varepsilon}{10} \quad \text{for all } 1 \le k \le m \tag{6}$$

$$\exp\left(-\int_T^\infty \lambda u(t) - 1 \, dt\right) > 1 - \frac{\varepsilon}{10}. \tag{7}$$

Let η_t^T be a voter model starting at time T from a product measure with probability $1/\lambda$ for a site to be in state 0. Since $\eta_t^T \Rightarrow \nu_{1/\lambda}$ as $t \to \infty$ it suffices to show that if $t \ge t_2$ then

$$\left| P(\xi_t(x_k) = 0 \text{ for } 1 \le k \le m) - P(\eta_t^T(x_k) = 0 \text{ for } 1 \le k \le m) \right| < \varepsilon/2.$$

Let $\tilde\eta$ be the dual of η starting at time $t > T$ with particles at $x_1, \ldots, x_m$. Let $\tilde\xi$ be the dual of ξ starting at time t with particles at $x_1, \ldots, x_m$. Let R_m be the value from Lemma 3.2 for $\delta = \varepsilon/8m$. Let S_t^x and S_t^y be independent random walks starting at x and y. By the local central limit theorem,

$$\lim_{s \to \infty} \sup_{x,y} P(|S_s^x - S_s^y| \le R_m) = 0$$

We pick t_1 so large that for $t \geq t_1$,

$$\sup_{x,y} P(|S^x_{t-T} - S^y_{T-t}| \leq R) \leq \varepsilon/(5m^2).$$

It follows that the probability that there is a non-coalesced pair in $\tilde{\xi}_t$ within distance R_m of each other at time T is $< \varepsilon/10$.

Let A be the event that there is no branching in the dual from t to T, and any pair that has not coalesced is at least R_m away from each other. Combining the computations above, $P(A^c) \leq 2\varepsilon/10$. Let

$$B = \{\xi_t(x_k) = 0 \text{ for } 1 \leq k \leq m\}$$
$$C = \{\eta^T_t(x_k) = 0 \text{ for } 1 \leq k \leq m\}$$
$$D_k = A \cap \{\tilde{\xi} \text{ has } k \text{ particles time } T\}$$
$$E_k = \{\tilde{\eta} \text{ has } k \text{ particles time } T\}$$

On the no branching event, the random walks in the two duals can be coupled, so (7) implies

$$\sum_{k=1}^m |P(D_k) - P(E_k)| \leq \frac{\varepsilon}{10} \tag{8}$$

Breaking thing down according to the number of particles in the dual at T

$$|P(B) - P(C)| = \left| \sum_{k=1}^m P(B \cap D_k) + P(B \cap A^c) - \sum_{\ell=k}^m P(C \cap E_k) \right|$$
$$\leq \frac{2\varepsilon}{10} + \sum_{k=1}^m |P(B \cap D_k) - P(C \cap E_k)|$$

Lemma 3.2 and the choice of R_m implies

$$|P(B \cap D_k) - u(T)^k P(D_k)| \leq \frac{\varepsilon}{10} \cdot P(D_k) \tag{9}$$

Since η^T starts from product measure with density $1/\lambda$ at time T

$$P(C \cap E_k) = (1/\lambda)^k P(E_k) \tag{10}$$

Using the triangle inequality and (10)

$$|P(B \cap D_k) - P(C \cap E_k)| \leq |P(B \cap D_k) - u(T)^\ell P(D_k)|$$
$$+ P(D_k)|u(T)^k - (1/\lambda)^k| + |P(D_k) - P(E_k)|(1/\lambda)^k$$

Summing $k = 1$ to m and using (9), (6), and (8) we have

$$\sum_{k=1}^m |P(B \cap D_k) - P(C \cap E_k)| \leq \frac{\varepsilon}{10} + \frac{\varepsilon}{10} + \frac{2\varepsilon}{10} < \varepsilon/2$$

completing the proof. $\qquad\square$

As in the previous section, only minor changes are needed to treat the case $u(0) < 1/\lambda$. The formula for $m_t(x)$ changes but the rest of the proof of Lemma 3.1 stays the same. The proofs of Lemmas 3.2 and Lemma 3.3 only use the exponential convergence and the estimate in Lemma 3.1, so they go through as before and the proof of Theorem 2 is complete.

4 Two dimensions

We begin by describing the construction of the process and duality on $\mathbb{Z}^d$ for general $\mathcal{N}_c$ and $\mathcal{N}_d$. The details are different from Section 2. There we used a percolation style dual which only works for "additive processes." Here we use the approach taken in [10] and [4] which works for any process.

Recall that we are considering birth-death dynamics for the evolutionary game

$$\bar{G} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \end{array} \begin{array}{cc} \mathbf{1} & \mathbf{2} \\ 1 & 1 + \lambda\varepsilon^2 \\ 1 + \varepsilon^2 & 1 + \varepsilon^2 \end{array}$$

Construction. Our process has voter events at rate 1 per site. If $\xi_t(y) = 2$ then at rate ε^2, the 2 gives birth onto a randomly chosen $x \in y + \mathcal{N}_c$. If $\xi_t(y) = 1$ then at rate $\lambda\varepsilon^2$ y chooses a neighbor z from $y + \mathcal{N}_d$ at random. If $\xi_t(z) = 2$ then y gives birth onto a randomly chosen $x \in y + \mathcal{N}_c$. Here we have replaced the computation of the fitness by averaging over the neighborhood (as was done in [4] and [7]) by the equivalent act of making a random choice from $x + \mathcal{N}_d$ to simplify the perturbation. In two dimensions this drastically reduces the size of the dual.

To construct the process we use a large number of Poisson processes. For each ordered pair (x, y) with $x \in y + \mathcal{N}_c$ we have a Poisson process $\{T_n^{x,y}, n \geq 1\}$ with rate $1/|\mathcal{N}_c|$ and $\{S_n^{x,y}, n \geq 1\}$ with rate $\varepsilon^2/|\mathcal{N}_c|$ At times $T_n^{x,y}$, x imitates the opinion at y, so we draw an arrow from x to y. At times $S_n^{x,y}$, x imitates the opinion at y if it is a 2, so we draw an arrow from x to y and write a 2 above it. Finally, for each triple with $x \in y + \mathcal{N}_c$, $z \in y + \mathcal{N}_d$ we have Poisson processes $R_n^{x,y,z}$ with rate $\varepsilon^2/|\mathcal{N}_c||\mathcal{N}_d|$. A times $R_n^{x,y,z}$ y will give birth onto x if y is in state 1 and z is in state 2, so we draw an arrow from y to x and write a 1 above it. We then draw an unnumbered line segment with no arrows from y to z.

Duality. We have used an explicit construction so that we can define a set valued dual process $\zeta_s^{x,t}$ by working backwards starting with $\zeta_0^{x,t} = \{x\}$. Here we are working on the original time scale. If a particle is at x and $t - s = T_n^{x,y}$ then it jumps to y at time s. If a particle is at x and $t - s = S_n^{x,y}$ then it gives birth to a particle at y at time s. If a particle is at x and $t - s = R_n^{x,y,z}$, then it gives birth to particles at y and z at time s. If the jumps or births cause two particle to be on the same site they coalesce to 1. $\zeta_s^{x,t}$ is called the influence set because if we know the states of all the sites in $\zeta_s^{x,t}$ at time $t - s$ then we can compute the state of x at time t.

Two dimensions. Let $N = \varepsilon^{-2}$ and take $\mathcal{N}_c = \mathcal{N}_d = Q_N$ where

$$Q_N = [-c\sqrt{\log N}, c\sqrt{\log N}]^2 \cap \mathbb{Z}^2$$

and c is a fixed constant. To carry out our proofs we will need a local central limit theorem that is uniform in N. Let $X_1, X_2, \ldots$ be uniform on Q_N and let $S_n = X_1 + \cdots + X_n$. This and the next few things we define should have superscript N's but we suppress this to avoid clutter. The uniform distribution on Q_N has variance $\sim (c^2/3)\log N$. Let $\sigma^2 = c^2/3$ and let

$$p_n(x) = P(S_n/\sqrt{n \log N} = x) \quad \text{for } x \in \mathcal{L}_n = \{z/\sqrt{n \log N} : z \in \mathbf{Z}\}$$
$$n(x) = (2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2) \quad \text{for } x \in (-\infty, \infty)$$

Theorem 5. *If $n, N \to \infty$, and $h = 1/\sqrt{\log N}$ then*

$$\sup_{x \in \mathcal{L}_n} \left| \frac{n^{1/2}}{h} p_n(x) - n(x) \right| \to 0$$

The proof is a small modification of the proof of Theorem 3.5.2 in [8]. To encourage the reader to skip it, we put the proof in Section 5.

Convergence of the dual to branching Brownian motion. Our next goal is to show that when space is scaled by dividing by $(N \log N)^{1/2}$ and time is run at rate N the dual converges to a branching Brownian motion. To make this possible we do not add newly born particles to the dual until time

$$t_N = N/(\log^{1/3} N)$$

has elapsed since the branching occurred. In the next lemma and in what follows "with high probability" means that the probability tends to 1 as $N \to \infty$. To make it easier to say things we call the parent and its children a family.

Lemma 4.1. *With high probability, at time t_N after a birth event, all noncoalesced family members are separated by $L_N = N^{1/2} \log^{1/4} N$.*

Proof. Pick two family members, assign to them independent random walks and let S_t^x be the difference in their x coordinates t units of time after the birth event.

$$\mathrm{var}\,(S_{t_N}^x) \sim t_N(c^2/3) \log N = (c^2/3)N \log^{2/3} N$$

which corresponds to a standard deviation of $O(N^{1/2} \log^{1/3} N)$ so by the local central limit theorem, $P(|S_{t_N}^x| > L_N) \to 1$ as $N \to \infty$. The last conclusion also holds for the y coordinate, which gives the result. $\square$

Lemma 4.2. *Suppose at time 0, two particles are separated by L_N. The probability that they hit by time $N \log N$ goes to 0 as $N \to \infty$.*

Proof. From the previous proof we see that $\mathrm{var}\,(S_t^x) \le Ct \log N$. When $t = N/\log^{2/3} N$ this is $CN \log^{1/3} N$, which corresponds to a standard deviation of $CN^{1/2} \log^{1/6} N$. so using the L^2 maximal inequality on the martingale S_t^x we see that with high probability that the two particles do not hit before time $N/\log^{2/3} N$. Let $S_t \in \mathbb{R}^2$ be the difference in the two particles locations when they use independent random walks and V_N be the amount of time that $S_t = (0,0)$ in $[N/\log^{2/3} N, N \log N]$. By the local central limit theorem, if N is large

$$EV_N \le C \int_{N/\log^{2/3} N}^{N \log N} \frac{1}{t \log N}\, dt$$
$$= \frac{C}{\log N} (\log(N \log N) - \log(N \log^{-2/3} N) = (5C/3) \frac{\log \log N}{\log N}$$

which converges to 0 as $N \to \infty$. Since the random walks jump at rate 1, if they hit they will spend an exponential amount of time with mean $1/2$ together which gives the result. $\square$

Combining Lemmas 4.1 and 4.2 we see that if all particles in the dual are separated by L_N just before at time of the kth branching event (i.e., at T_k-) then all of the noncoalesced family members will be sepearated by L_N at time $T_k + t_N$. There will be no more coalescence within the family before time T_{k+1}, and there will be no coalescences between the family and other particles during $[T_k, T_{k+1}]$. Since the time of the next birth has an exponential distribution and is $O(N)$, another use of the local central limit theorem shows that at time $T_{k+1}-$ all of the existing particles are separated by distance L_N. Convergence of the rescaled dual to branching Brownian motion follows easily from this. More details than you want to read can be found in Chapter 2 of [4]. A more succinct proof with a structure that parallels the one used here can be found in Section 10 of [3].

Computation of the reaction term. Let v_i be independent and uniform on Q_N. Let $p_{N,M}(0|v_1)$ be the probability that random walks starting from 0 and v_1 do not hit by time M. Let $p_{N,M}(0|v_1|v_1 + v_2)$ be the probability that random walks starting from 0, v_1 and $v_1 + v_2$ do not hit by time M. Let $p_{N,M}(0|v_1, v_1 + v_2)$ be the probability that at time M the random walks starting from v_1 and $v_1 + v_2$ have coalesced but have not hit the one starting from 0. Ultimately we will show that if $N \to \infty$ and $N/\log N \le M(N) \le N \log N$ then the $p_{N,M}$ hitting probabilities converge to limits $p(0|v_1)$, $p(0|v_1|v_1 + v_2)$ and $p(0|v_1, v_1 + v_2)$.

Once this is done we can use result in Section 12 of [7] to compute the reaction term. To state the result we begin by recalling that the faction of individuals playing strategy i in a homogeneously mixing system satisfies the replicator equation (2), which can be written as:

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

Here $\equiv$ indicates we are defining $\phi_i^R(u)$. Formula (12.4) from [7] then implies that the reaction term for our birth-death updating is

$$\phi_B^i(u) = p(0|v_1|v_1 + v_2) \cdot \phi_R^i(u)$$
$$+ p(0|v_1, v_1 + v_2) \cdot \sum_{j \ne i} u_i u_j (G_{i,i} - G_{j,i} + G_{i,j} - G_{j,j})$$

Coalescence probabilities. We begin by computing $E p_{N,M}(0|v_1)$. It is easier to do the calculation for a discrete time random walk S_n with jumps uniform on Q_N that starts at v_1. In order for S_n to be at 0 at time $n \ge 1$, S_{n-1} has to be close enough to 0, and the jump X_n has to be exactly the right size so

$$P(S_n = 0) \le 1/|Q_N| \sim 1/4c^2 \log N \tag{11}$$

This bound implies that as $N \to \infty$

$$P(S_n = 0 \text{ for some } n \le \sqrt{\log N}\,) \to 0$$

For $n \ge \sqrt{\log N}$ we can use the local limit theorem to conclude

$$P(S_n = 0) \sim \frac{1}{(2\pi c^2/3)n \log N} \equiv \frac{b}{n \log N}$$

18

where we have set $b = 3/(2\pi c^2)$. In time M there are $\sim 2M$ jumps and

$$\sum_{n=0}^{2M} P(S_n = 0) = \frac{b \log(2M)}{\log N}. \tag{12}$$

Note that if $M(N) = N^s$ this converges to bs.

Lemma 4.3. *Let $R_N(s)$ be the number of returns of S_n to 0 by time N^s. As $N \to \infty$, $R_N(s) \Rightarrow R(s)$ is a Poisson process with rate b. This implies that if $N/\log N \le M(N) \le N \log N$*

$$E p_{N,M}(0|v_1) \to e^{-b}$$

Proof. It suffices to show

(i) If $0 \le a_1 < b_1 \le a_2 < b_2 \ldots a_n < b_n$ then $R(b_i) - R(a_i)$ are independent.

(ii) $E[R(t) - R(s)] = b(t - s)$

(iii) $P(R(t + h) - R(t) > 1) = o(h)$

To see this is sufficient note that if we subdivide $[s, t]$ into n intervals and let $X_{n,i}$ be the number of arrivals in the ith interval then (i) implies the $X_{n,i}$ are independent, while (ii) and (iii) imply $nP(X_{n,i} > 1) \to 0$ and $nP(X_{n,i} = 1) \to b$ so using a standard Poisson convergence result, see e.g., Theorem 3.6.1 in [8] that $N(t) - N(s)$ is Poisson.

To check (i), it suffices to prove that this holds when $b_i < a_{i+1}$ for $1 \le i \le n - 1$ for then a limiting argument gives the general case. To prove this weaker result we use induction. Condition on the path of the random walk up to time $N^{b(n-1)}$. With high probability $|S(N^{b(n-1)})| \le N^{b(n-1)/2} \log N$. When this is true, $|S(N^{a(n)}) - S(N^{b(n-1)})| \gg |S(N^{b(n-1)})|$, so the conditional probability of a return to 0 in $[N^{a(n)}, N^{b(n)}]$ is in the limit, independent of the value of $S(N^{b(n-1)})$.

Condition (ii) follows from the derivation of (12). To check (iii) we start with the observation that (11) implies that after a return to 0 there will not be one for the next $\sqrt{\log N}$ units of time. Using the proof of (12) again we see that the conditional probability of another return to 0 by time N^{t+h} is $\le Ch$. $\qquad \square$

Lemma 4.4. *Let S_t^0, S_t^1, S_t^2 be independent continuous time random walks that take jumps uniform on Q_N and start at 0, v_1, and $v_1 + v_2$. Let $R_N^3(s)$ be the number of collisions between these random walks up to time N^s. Then as $N \to \infty$, $R_N^3(s) \Rightarrow R(s)$ a Poisson process with rate $3b$. It follows that if $N/\log N \le M(N) \le N \log N$ then*

$$E p_{N,M}(0|v_1|v_1 + v_2) \to e^{-3b}$$

Proof. The main difficulty is to control the correlation between hits of the different pairs. Define a six dimensional random walk by $V_t = (S_t^1 - S_t^0, S_t^2 - S_t^1, S_t^0 - S_t^2)$. Since the sum of the three differences is 0 this walk lies in a four dimensional subspace. The possible values of V_t are a four dimensional lattice, so the random walk is "genuinely four dimensional" and hence transient. In 1951 Dvoretsky and Erdös [12] proved a rate of escape for simple random walk W_n. Here we have used 2 instead of 1 in the integral test to avoid the fact that $\log(1) = 0$. One can of course use any fixed value K.

Lemma 4.5. *Suppose that $\psi : [2, \infty) \to (0, \infty)$ satisfies $t^{-1/2}\psi(t) \downarrow 0$ then $\psi(n)W_n \to \infty$ if and only if*

$$\int_2^\infty \psi(t)^{d-2} t^{-d/2} \, dt = \infty$$

Later Kesten [16] showed that this holds for any genuinely d-dimensional random walk. If we let $\psi(t) = t^{1/2} \log^{-\alpha}(t)$ then the integral is

$$\int_2^\infty t^{-1} \log(t)^{-(d-2)\alpha} \, dt$$

which diverges if $\alpha(d-2) < 1$. If $d = 4$ this holds if $\alpha = 1/3$. This implies that

Lemma 4.6. *If two random walks hit at time t then the other one is with high probability at least a distance $t^{1/2} \log^{-1/3} t$ away.*

To prove Lemma 4.4 now we have to check (i), (ii), and (iii) from the previous proof. To check (i), it again suffices to prove that this holds when $b_i < a_{i+1}$ for $1 \le i \le n - 1$. The argument is almost the same as before. We are considering independent random walks so if $t \to \infty$ then $V(t)/\sqrt{t \log N}$ has a limiting multivariate normal distribution. Condition on the path of the random walk V_t up to time $N^{b(n-1)}$. With high probability $|V(N^{b(n-1)})| \le N^{b(n-1)/2} \log N$. When this is true, $|V(N^{a(n)}) - V(N^{b(n-1)})| \gg |S(N^{b(n-1)})|$, so the conditional probability of a return to 0 in $[N^{a(n)}, N^{b(n)}]$ is in the limit, independent of the value of $S(N^{b(n-1)})$.

Condition (ii) follows from the derivation of (12) since in this calculation we are computing an expected value and don't have to worry about the correlation between the three differences. To check (iii) we start with the observation that (11) implies that after a return to 0 there will not be one for the next $\sqrt{\log N}$ units of time. Using Lemma 4.6 we see that during these $\sqrt{\log N}$ steps there will be no collision with the other random walk. Using the proof of (12) again we see that the conditional probability of another return to 0 by time N^{t+h} is $\le Ch$. $\square$

Lemma 4.7.

$$Ep_{N,M}(0|v_1, v_1 + v_2) \to \frac{e^{-b} - e^{-3b}}{2}$$

Proof. By Lemma 4.4 the time of the first collision is exponential with rate $3b$. All three pairs has aysmptotically the same probability to coalesce. Using Lemma 4.6 we see that the time to a collision between the coalesced pair and the remaining particle is exponential with rate b so

$$Ep_{N,M}(0|v_1, v_1 + v_2) \to \frac{1}{3} \int_0^1 3be^{-3bu} e^{-b(1-u)} \, du$$

$$= \frac{e^{-b}}{2} \int_0^1 2be^{-2bu} \, du = \frac{e^{-b}}{2} \cdot (1 - e^{-2b})$$

which gives the desired result. $\square$

Convergence to the limiting PDE. With the convergence of the dual to branching Brownian motion, the convergence to the PDE is the same as in [10], [11], and in Section 2.6–2.10 of [4].

5   Proof of local CLT

Proof of 5. Let Y be a random variable with $P(Y \in \theta \mathbf{Z}) = 1$ and $\psi(t) = E \exp(itY)$. It follows from part (iii) of Exercise 3.3.2 in [8] that

$$P(Y = x) = \frac{1}{2\pi/\theta} \int_{-\pi/\theta}^{\pi/\theta} e^{-itx} \psi(t)\, dt$$

Using this formula with $\theta = 1/\sqrt{n \log N}$, $\psi(t) = E \exp(itS_n/\sqrt{n \log N}) = \varphi^n(t/\sqrt{n \log N})$, and then multiplying each side by $1/\theta$ gives

$$(n^{1/2} \log^{1/2} N) p_n(x) = \frac{1}{2\pi} \int_{-\pi\sqrt{n \log N}}^{\pi\sqrt{n \log N}} e^{-itx} \varphi^n(t/\sqrt{n \log N})\, dt$$

Using the inversion formula for continuous densities, Theorem 3.3.5 in [8], on $n(x)$, which has ch.f. $\exp(-\sigma^2 t^2/2)$, gives

$$n(x) = \frac{1}{2\pi} \int e^{-itx} \exp(-\sigma^2 t^2/2)\, dt$$

Subtracting the last two equations gives (recall $\pi > 1$, $|e^{-itx}| \leq 1$)

$$\left| (n^{1/2} \log^{1/2} N) p_n(x) - n(x) \right| \leq \int_{-\pi\sqrt{n \log N}/h}^{\pi\sqrt{n \log N}/h} |\varphi^n(t/\sqrt{n \log N}) - \exp(-\sigma^2 t^2/2)|\, dt$$
$$+ \int_{\pi\sqrt{n \log N}/h}^{\infty} \exp(-\sigma^2 t^2/2)\, dt$$

The right-hand side is independent of x, so to prove the theorem it suffices to show that it approaches 0. The second integral clearly $\to 0$. To estimate the first integral, we observe that $\varphi^n(t/\sqrt{n \log N}) \to \exp(-\sigma^2 t^2/2)$, so the integrand goes to 0.

To prove that the integral converges to 0, we will divide the integral into three pieces. The bounded convergence theorem implies that for any $A < \infty$ the integral over $(-A, A)$ approaches 0. To estimate the integral over $(-A, A)^c$, we let $\bar{\varphi}(t) = \phi(t/\sqrt{\log N})$ be the characteristic function of $Y_i X_i/\sqrt{\log N}$ and note that since $EY_i = 0$ and $EY_i^2 = \sigma^2$, formula (3.3.3) from [8] and the triangle inequality imply that

$$|\bar{\varphi}(u)| \leq |1 - \sigma^2 u^2/2| + \frac{u^2}{2} E(\min(|u| \cdot |Y|^3, 6|Y|^2))$$

The last expected value $\to 0$ as $u \to 0$ uniformly in N. This means we can pick $\delta > 0$ so that if $|u| < \delta$, it is $\leq \sigma^2/2$ and hence

$$|\bar{\varphi}(u)| \leq 1 - \sigma^2 u^2/4 \leq \exp(-\sigma^2 u^2/4),$$

since $1 - x \leq e^{-x}$. Applying the last result to $u = t/\sqrt{n}$ we see that for $t \leq \delta\sqrt{n}$

$$(*) \qquad\qquad |\varphi(t/\sqrt{n})^n| \leq \exp(-\sigma^2 t^2/4)$$

So the integral over $(-\delta\sqrt{n}, \delta\sqrt{n}) - (-A, A)$ is smaller than

$$2\int_A^{\delta\sqrt{n}} \exp(-\sigma^2 t^2/4)\, dt$$

which is small if A is large.

To estimate the rest of the integral we observe that if $t \in -[-\pi, \pi]$

$$\bar{\varphi}(u) \approx \int_{-c}^c \cos(tx)\frac{dx}{2c} = \frac{\sin(tc)}{tc}$$

where the error in $\approx$ comes from the difference between the uniform and the distribution of Y_i, and hence is $O(1/\sqrt{\log N})$. From this it follows that there is an $\eta < 1$ so that $|\bar{\varphi}(u)| \le \eta < 1$ for $|u| \in [\delta, \pi]$, uniformly in N. Letting $u = t/\sqrt{n}$ again, we see that the integral over $[-\pi\sqrt{n}, \pi\sqrt{n}] - (-\delta\sqrt{n}, \delta\sqrt{n})$ is smaller than

$$2\int_{\delta\sqrt{n}}^{\pi\sqrt{n}/h} \eta^n + \exp(-\sigma^2 t^2/2)\, dt$$

which $\to 0$ as $n \to \infty$. This completes the proof. $\qquad\square$

References

[1] Archetti, M., Ferraro, D.A., Christofori, G. (2015) Heterogeneity for IGF-II production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *Proc. Natl. Acad. Sci.* 112, 1833–1838

[2] Cox, J.T. (2010) Intermediate range migration in the two-dimensional stepping stone model. *Ann. Appl. Probab.* 20, 785-805

[3] Cox, J.T. and Durrett, R. (2016) Evolutionary games with weak selection on the torus. *Stoch. Proc. Appl.* 126, 2388–2409

[4] Cox, J.T., Durrett, R., and Perkins, E. (2011) Voter model perturbations and reaction diffusion equations. *Astérisque*, 343 or arXiv:1103.1676

[5] Cox, J.T., Merlke, M., and Perkins, E. (2011) Coexistence in a two-dimensional Lotka-Volterra model. *Electronic J. Probability* 15, 1198—1266

[6] Durrett, R. (1992) Multicolor particle systems with large threshold and range. *J. Theoretical Prob.*, 5, 127-152

[7] Durrett, R. (2014) Spatial evolutionary games with small selection coefficients. *Electronic J. Probability.* 19 (2014), paper 121

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

[9] Durrett, R., Liggett, T.M., and Zhang, Y. (2014) The contact process with fast voting. *Electronic Journal of Probability.* 18, paper 28

[10] Durrett, R., and Neuhauser, C. (1994) Particle systems and reaction diffusion equations. *Ann. Probab.* 22 (1994), 289-333

[11] Durrett, R., and Zahle, I. (2007) On the width of hybrid zones. *Stoch. Proc. Appl.* 117 (2007), 1751–1763

[12] Dvoretsky, A., and Erdös, P. (1951) Some problems on random walk in space. *Proc. 2nd Berkeley Symp.* 353–367

[13] Griffeath, D.S. (1978) *Additive and Cancellative Interacting Particle Systems.* Springer Lecture notes in Math 724.

[14] Harris, T.E. (1976) On a class of set-valued Markov processes. *Ann. Probab.* 4, 175–194

[15] Hofbauer, J., and Sigmund, K. (1998) *Evolutionary Games and Population Dynamics.* Cambridge U. Press

[16] Kesten, H. (1978) Erickson's conjecture on the rate of escape of d-dimensional random walk. *Transactions of the AMS.* 240, 65–113

[17] Liggett, T.M. (21983) *Interacting Particle Systems.* Springer, New York

[18] Molofsky, J., Durrett, R., Dushoff, J., Griffeath, D., and Levin, S.A. (1999). Local frequency dependence and global coexistence. *Theoret. Pop. Biol.*, 55, 270–282

[19] Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M.A. (2006) A simple rule for the evolution of cooperation on graphs and social networks. *Nature.* 441, 502–505