LARGE DEVIATIONS FOR MEAN FIELD MODELS OF PROBABILISTIC CELLULAR AUTOMATA

P. BALISTER[1], B. BOLLOBÁS[2], AND R. KOZMA[1]

Abstract. Probabilistic cellular automata (PCA) form a very large and general class of stochastic processes. These automata exhibit a wide range of complex behavior and are of interest in a number of fields of study, including mathematical physics, percolation theory, computer science, and neurobiology. Very little has been proved about these models, even in simple cases, so it is common to compare the models to mean field models. It is normally assumed that mean field models are essentially trivial. However, we show here that even the mean field models can exhibit surprising behavior. We prove some rigorous results on mean field models, including the existence of a surrogate for the 'energy' in certain non-reversible models. We also briefly discuss some differences that occur between the mean field and lattice models.

1. Probabilistic Cellular Automata

Let A be a finite set of *states* and let Γ be a finite subset of $\mathbb{Z}^d$ containing the origin $\mathbf{0} = (0, \dots, 0)$. A *probabilistic cellular automaton* or *PCA* on $\mathbb{Z}^d$ is a stochastic process giving rise to a sequence of configurations $\Phi_t \colon \mathbb{Z}^d \to A$, where the state $\Phi_t(\mathbf{x})$ of $\mathbf{x} \in \mathbb{Z}^d$ at time t is determined randomly with probabilities dependent on the states of the points of the neighborhood $\mathbf{x}+\Gamma = \{\mathbf{x}+\mathbf{y} : \mathbf{y} \in \Gamma\}$ of $\mathbf{x}$ at time $t-1$. To be more precise, fix a function $p \colon A^\Gamma \times A \to [0,1]$ that assigns for each configuration $\phi \colon \Gamma \to A$ and each $a \in A$ a probability $p_{\phi,a}$ with $\sum_{a \in A} p_{\phi,a} = 1$ for all ϕ. We set $\Phi_{t+1}(\mathbf{x}) = a$ independently for each $\mathbf{x} \in \mathbb{Z}^d$ with probability $p_{\phi,a}$, where $\phi(\mathbf{y}) = \Phi_t(\mathbf{x}+\mathbf{y})$ is the restriction to Γ of (a translate of) Φ_t. We start the process with Φ_0 chosen randomly from some specified probability distribution on the set of all states $\Phi \colon \mathbb{Z}^d \to A$.

Usually we define Γ to be the closed neighborhood of $\mathbf{0}$, consisting of all points unit distance from $\mathbf{0}$ together with $\mathbf{0}$ itself, or more generally, all points within some fixed distance of $\mathbf{0}$. Similarly we can define probabilistic cellular automaton on the d-dimensional torus $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}$, or on the corresponding finite grid (with suitable boundary conditions imposed). These models have also been referred to as *contact processes* and have been studied in some simple cases on infinite graphs [21, 22].

Probabilistic cellular automata generalize deterministic cellular automata such as Conway's game of life [17], and bootstrap percolation [1, 3, 8, 13, 18, 25, 27], where the update rules are deterministic, but the initial configuration is random. Oriented

Date: October 1, 2009.

[1]Research supported by NSF grant EIA-0130352.

[2]Research supported by NSF grants DMS-9970404 and EIA-0130352 and DARPA grant F33615-01-C1900.

percolation models in $\mathbb{Z}^d$ (see [4, 6, 19, 22]) can also be thought of as $(d-1)$-dimensional probabilistic cellular automata, with time representing the direction of orientation. Given the complexity of deterministic automata it should not come as a surprise that random automata display extremely complex behavior which is very difficult to analyze rigorously.

In the models described, we have updated all the sites in $\mathbb{Z}^d$ simultaneously at each step. One can ask what happens if the sites are updated asynchronously? Equivalently, run the process in continuous time t and change each site according to independent exponential random variables whose rates $r_{\phi,a}$ (defined for $a \neq \phi(\mathbf{0})$) depend on the neighborhood of the site. For a finite grid the asynchronous model can be obtained as a limit of synchronous models as follows. Given transition rates $r_{\phi,a}$ for all $\phi\colon \Gamma \to A$, define for sufficiently small $\varepsilon > 0$ the random automaton with probabilities

$$p_{\phi,a} = \begin{cases} \varepsilon r_{\phi,a} & \text{if } \phi(\mathbf{0}) \neq a; \\ 1 - \sum_{b \neq a} \varepsilon r_{\phi,b} & \text{if } \phi(\mathbf{0}) = a. \end{cases} \tag{1.1}$$

Then as $\varepsilon \to 0$ this automaton approximates the asynchronous model with rates $\varepsilon r_{\phi,a}$, since the chances of neighbors being updated at the same time tends to zero. For the infinite lattice $\mathbb{Z}^d$ one expects the same result to hold — the asynchronous model being approximated increasingly well by the synchronous model as $\varepsilon \to 0$ (with suitable re-scaling of t). As a result, asynchronous models can be considered as a special limiting case of suitable synchronous models. However the synchronous models display a richer variety of behavior, such as small period oscillations, that are not possible in asynchronous models. Also it is not clear that, for example, critical exponents at phase transitions for asynchronous models necessarily match those of the approximating synchronous models. There are however reasons for believing that the asynchronous models are easier to analyze — for example, the Kinetic Ising model can be more easily described by an asynchronous model (see Section 4).

2. The Mean Field Approximation

From now on we shall restrict our attention to 2-state processes with $A = \{-, +\}$. In this case we write p_ϕ for $p_{\phi,+}$ and so $p_{\phi,-} = 1 - p_\phi$.

We start by considering p of a certain special form. Assume p_ϕ depends only on the cardinality of set of neighbors that are in state $+$, $|\{\mathbf{x} \in \Gamma : \phi(\mathbf{x}) = +\}|$, and on the state of the site itself, $\phi(\mathbf{0})$. We shall call these models *semi-totalistic*. We write p_r^+ (resp. p_r^-) in place of p_ϕ when $|\phi^{-1}(+)| = r$ and $\phi(\mathbf{0}) = +$ (resp. $\phi(\mathbf{0}) = -$). Semi-totalistic models are substantially more restrictive than the general case, but they still have complex behavior, sometimes including spontaneous symmetry breaking (see Section 4). We shall call the model *totalistic* if $p_r^- = p_r^+ = p_r$ for all r. In this case, the site itself is treated on the same basis as its neighbors.

By interchanging $+$ and $-$ throughout we obtain a new model with $p'_\phi = 1 - p_{-\phi}$ ($p'^{\pm}_r = 1 - p^{\mp}_{|\Gamma|-r}$ for semi-totalistic models). We call the model *symmetric* if it is the same under interchange of $+$ and $-$, so $p_\phi = 1 - p_{-\phi}$.

On a finite grid or torus, we can compare probabilistic automata with the corresponding *mean field* models. In the mean field model, instead of taking $|\Gamma| - 1$ specified neighbors, we take $|\Gamma| - 1$ elements of the grid at random (with replacement). Since there is no ordering of these neighbors, the transition probabilities depend only on the number of + states in this neighborhood and the state of the site itself, so may be given by probabilities $p_r^{\pm}$ as in the semi-totalistic case. Because of this, we shall only consider semi-totalistic models from now on. However, we shall not in general assume our models are totalistic, since even for a totalistic asynchronous model, the synchronous approximations described above are only semi-totalistic, and we may wish to compare asynchronous models to their corresponding mean field versions.

It is clear that the mean field model does not depend on the topology of the grid, and the only information of relevance in Φ_t is given by the cardinality of +-states, $|\Phi^{-1}(+)| = |\{\mathbf{x} : \Phi_t(\mathbf{x}) = +\}|$. We define ρ_t to be $|\Phi^{-1}(+)|/N$ where N is the size of grid. Thus $\rho_t \in [0, 1]$ gives the density of + sites in Φ_t. Write $B(k, p)$ for a binomial random variable giving the sum of k independent Bernoulli random variables, each of which is 1 with probability p.

Theorem 2.1. *If there are N points in the grid then ρ_t is given by*

$$N\rho_{t+1} = B(N\rho_t, f_m^+(\rho_t)) + B(N(1-\rho_t), f_m^-(\rho_t)), \tag{2.1}$$

where the functions $f_m^{\pm}$ are defined by

$$f_m^+(x) = \sum_r \tbinom{|\Gamma|-1}{r-1} p_r^+ x^{r-1}(1-x)^{|\Gamma|-r}, \tag{2.2}$$

$$f_m^-(x) = \sum_r \tbinom{|\Gamma|-1}{r} p_r^- x^r(1-x)^{|\Gamma|-r-1}. \tag{2.3}$$

Hence ρ_{t+1} is given by a distribution with mean $f_m(\rho_t)$ and variance $g_m(\rho_t)/N$ where

$$f_m(x) = xf_m^+(x) + (1-x)f_m^-(x) = \sum_r \tbinom{|\Gamma|}{r} p_r x^r(1-x)^{|\Gamma|-r}, \tag{2.4}$$

$$g_m(x) = xf_m^+(x)(1-f_m^+(x)) + (1-x)f_m^-(x)(1-f_m^-(x)), \tag{2.5}$$

and $p_r = (rp_r^+ + (|\Gamma| - r)p_r^-)/|\Gamma|$.

Proof. Clear. □

For large N, ρ_{t+1} is approximately $f_m(\rho_t)$ with high probability. Therefore, to a first approximation, the behavior of this model can be described in terms of the iterates of f_m. Note that f_m is unchanged if we replace the random automaton by the totalistic model with parameters p_r, however $f_m^{\pm}$ and g_m may be different, and as we shall see, the dynamics can be significantly altered.

Using Theorem 2.1, we can consider the mean field model to be a special case of a more general stochastic model where +'s and −'s at time t become +'s at time $t + 1$ with probabilities $f_m^+(r/N)$ and $f_m^-(r/N)$ respectively, where r is the total number of + states at time t. We can then forget the $p_r^{\pm}$ and just consider $f_m^{\pm}$ to be arbitrary continuous functions from $[0, 1]$ to $[0, 1]$. The functions f_m and g_m are then defined in

terms of $f_m^{\pm}$ by (2.4) and (2.5). Our main results (Theorem 2.2 and Theorem 3.1) are proved in this more general setting.

Since $f_m(0) \geq 0$ and $f_m(1) \leq 1$, there must be at least one fixed point of f_m in $[0,1]$. Iterating the map f_m can result in stable fixed points, stable limit cycles, chaotic behavior, or a combination of these, depending on the initial value ρ_0.

Both Φ_t and ρ_t are Markov processes on a finite state space. Thus they have stationary distributions. For large N one expects these to be heavily concentrated near stable fixed points and limit cycles when these exist, whereas if iterates of f_m behave chaotically one expects the stationary distribution of ρ_t to be more complex. When there are several attractors, the stationary distribution may be heavily concentrated about just one of them, so the process spends most of its time near just one fixed point or limit cycle even when several exist for the iterates of f_m. In the case when f_m is monotonically increasing (so the only attractors are fixed points) we have the following result.

Theorem 2.2. *Assume $f_m^{\pm}(x)$ are continuous, $f_m^{\pm}(x) \neq 0, 1$, and $f_m(x)$ is increasing with a finite number of fixed points. Then there exists a continuous function $\lambda\colon [0,1] \to [0,\infty)$, independent of N, such that in the stationary distribution,*

$$\mathbb{P}(\rho_t = r/N) = e^{-N\lambda(r/N)+o(N)}. \tag{2.6}$$

Furthermore, if I is an interval on which $f_m(x) > x$ (resp. $f_m(x) < x$) then λ is a strictly decreasing (resp. increasing) function on I. Also, $\lambda(x) = 0$ implies (but is not implied by) x is a stable fixed point of f_m.

The function $\lambda(x)$, which in general is somewhat awkward to calculate, acts as a surrogate for the 'energy' of the system, even though there is no Hamiltonian defined for non-reversible Markov chains.

To prove Theorem 2.2, we shall need the following 'large deviations' estimate for the Binomial distribution.

Lemma 2.3. *If $0 \leq r \leq N$ and $0 < p < 1$, then*

$$\mathbb{P}(B(N,p) = r) = \exp(-Nh(r/N,p) + O(\log N)) \tag{2.7}$$

where

$$h(z,p) = z \log \tfrac{z}{p} + (1-z) \log \tfrac{1-z}{1-p}. \tag{2.8}$$

Note that in the case when $z = 0$ or 1 we use the convention that $0 \log 0 = 0$.

Proof. $\mathbb{P}(B(N,p) = s) = \binom{N}{s} p^s (1-p)^{N-s}$, and for $p = z = r/N$, this is maximized when $s = r$. Since there are only $N+1$ possible values for s, we have

$$\tfrac{1}{N+1} \leq \binom{N}{r} z^r (1-z)^{N-r} \leq 1.$$

Hence

$$\begin{aligned} \log\left(\binom{N}{r} p^r (1-p)^{N-r}\right) &= \log\left(\binom{N}{r} z^r (1-z)^{N-r}\right) - r \log \tfrac{z}{p} - (N-r) \log \tfrac{1-z}{1-p} \\ &= O(\log N) - Nh(z,p) \end{aligned}$$

as required. □

Corollary 2.4. *With the notation and assumptions as in Theorem 2.2,*

$$\mathbb{P}(\rho_{t+1} = x \mid \rho_t = y) = \exp(-Nt(y,x) + O(\log N)), \tag{2.9}$$

where $x = r/N$, $0 \le r \le N$,

$$t(y,x) = \inf_{yz_+ + (1-y)z_- = x} \{yh(z_+, f_m^+(y)) + (1-y)h(z_-, f_m^-(y))\}, \tag{2.10}$$

and the infimum is over all $(z_+, z_-) \in [0,1]^2$ *such that* $yz_+ + (1-y)z_- = x$.

Note that here and below, the constant implicit in the $O()$ term is bounded independently of the parameters x and y, although it may depend on the functions $f_m^\pm$.

Proof. Write $p_\pm = f_m^\pm(y)$. By Theorem 2.1, the probability that we want is given by

$$\begin{aligned}&\textstyle\sum_{r_+ + r_- = r} \mathbb{P}(B(Ny, p_+) = r_+)\mathbb{P}(B(N(1-y), p_-) = r_-)\\ &\quad = \textstyle\sum_{r_+ + r_- = r} \exp\{-Nyh(z_+, p_+) - N(1-y)h(z_-, p_-) + O(\log N)\}\end{aligned}$$

where $z_+ = \frac{r_+}{Ny}$ and $z_- = \frac{r_-}{N(1-y)}$. By the addition of another $O(\log N)$ term in the exponent, the sum can be replaced by a maximum over $r_\pm$. By the assumptions of Theorem 2.2, $p_\pm$ are continuous in y, and are never 0 or 1 for $y \in [0,1]$. Hence $p_\pm$ and $1 - p_\pm$ are bounded away from zero, independently of y. Thus, if we vary $z_\pm$ by at most $1/M$, the functions $h(z_\pm, p_\pm)$ given by (2.8) vary by at most $O((\log M)/M)$. (The worst case is when $z_\pm$ varies between 0 and $1/M$, or between 1 and $1 - 1/M$). Hence the maximum over $r_\pm$ can be replaced by a supremum over all $z_\pm \in [0,1]$ with the addition of another $O(\log N)$ term in the exponent. The result now follows. □

Before we continue, it is worth describing the function $t(y,x)$ in a little more detail. First we note that $\frac{\partial h}{\partial z}(z,p) = \log \frac{z(1-p)}{p(1-z)}$ tends to $\pm\infty$ as $z \to 1$ or 0. Hence the infimum in (2.10) occurs with $z_\pm$ in the interior of $[0,1]$ when $y \in (0,1)$. Indeed, by the method of Lagrange multipliers, the infimum occurs when

$$yz_+ + (1-y)z_- = x \quad \text{and} \quad \log \tfrac{z_+(1-p_+)}{p_+(1-z_+)} = \log \tfrac{z_-(1-p_-)}{p_-(1-z_-)} = \beta, \text{ say.} \tag{2.11}$$

For $y = 0$ or 1 either z_+ or z_- is undetermined by (2.10), however any solution of (2.11) still achieves the infimum in (2.10). Since $z_\pm$, and hence x, increase monotonically with β, these equations uniquely determine $z_\pm$, and hence $t(y,x)$, in terms of x for all $x \in (0,1)$. Moreover, the limiting solutions of (2.11) as $x \to 0$ or 1 ($\beta \to \mp\infty$) give the correct value of $t(y,x)$ for $x = 0, 1$, and

$$\frac{\partial t}{\partial x}(y,x) = \beta \qquad \text{for all } x \in (0,1),\ y \in [0,1]. \tag{2.12}$$

Since $p_\pm$ are continuous and bounded away from 0 and 1, $t(y,x)$ is continuous in both variables for all $(y,x) \in [0,1]^2$. Since $[0,1]^2$ is compact, it is also uniformly continuous.

The minimum value of $t(y,x)$ as x varies is 0 and occurs when $\beta = 0$, $z_\pm = p_\pm$, and $x = f_m(y)$. Indeed, if $x \approx f_m(y)$ then one can check that $\beta \approx (x - f_m(y))/g_m(y)$ and $t(y,x) \approx (x - f_m(y))^2/2g_m(y)$, which is what one would expect from the central limit theorem. For $x > f_m(y)$ we have $\beta > 0$, and so $t(y,x)$ is strictly increasing in x.

Similarly, if $x < f_m(y)$ then $t(y,x)$ is strictly decreasing in x. Finally, $\frac{\partial \beta}{\partial z_\pm} \geq 2$, so $\frac{\partial \beta}{\partial x} \geq 2$ and thus by (2.12), $t(y,x) \geq (x - f_m(y))^2$. Hence for any $\varepsilon > 0$ there is a $\delta > 0$, independent of x and y, such that

$$t(y,x) < \delta \qquad \Rightarrow \qquad |x - f_m(y)| < \varepsilon. \tag{2.13}$$

The construction of $\lambda(x)$ is given in the proof of Theorem 2.2 below, however we note that $\lambda(x)$ satisfies

$$\lambda(x) = \inf_y \{\lambda(y) + t(y,x)\}, \qquad \inf_x \lambda(x) = 0. \tag{2.14}$$

Unfortunately, (2.14) alone does not quite specify $\lambda(x)$, so we need to be a bit more careful in its definition.

Proof of Theorem 2.2. We prove Theorem 2.2 in several steps.

Step 1. Approximate the log of the distribution at finite time.

Let N be large and consider the distribution $\mu_{x_0,i}$ of the density of the mean field model after i steps starting with density x_0. The distribution $\mu_{x_0,i}$ is given by

$$\mu_{x_0,i}(x) = \begin{cases} 1 & i = 0,\ x = x_0; \\ 0 & i = 0,\ x \neq x_0; \\ \sum_y \mu_{x_0,i-1}(y) \mathbb{P}(\rho_{t+1} = x \mid \rho_t = y) & i > 0. \end{cases}$$

where the sum is over $y = r/N$. By Corollary 2.4

$$\log \mu_{x_0,i}(x) = \log \textstyle\sum_{y=r/N} \exp\{\log \mu_{x_0,i-1}(y) - Nt(y,x) + O(\log N)\},$$

where the constant implicit in the $O(\log N)$ term depends only on the functions $f_m^\pm$, and is independent of x, y, x_0, and i. There are only $N+1$ values of y, so we can, by the addition of another $O(\log N)$ term, write this as

$$\log \mu_{x_0,i}(x) = \sup_{y=r/N} \{\log \mu_{x_0,i-1}(y) - Nt(y,x) + O(\log N)\}.$$

Extend the definition of $\mu_{x_0,i}(x)$ to arbitrary $x \in [0,1]$ by, for example, linear interpolation of $\log \mu_{x_0,i}$ between $x = s/N$ and $x = (s+1)/N$. Since $t(y,x)$ is uniformly continuous in both x and y, we can extend the supremum over all $y \in [0,1]$ to get

$$\log \mu_{x_0,i}(x) = \sup_{y \in [0,1]} \{\log \mu_{x_0,i-1}(y) - Nt(y,x) + o(N)\},$$

where the $o(N)$ term is bounded independently of x, x_0, and i.

Define for each $x_0 \in [0,1]$ and $i \geq 0$ the function $\lambda_{x_0,i}$ inductively so that

$$\lambda_{x_0,i}(x) = \begin{cases} 0 & i = 0,\ x = x_0; \\ +\infty & i = 0,\ x \neq x_0; \\ \inf_{y \in [0,1]} \{\lambda_{x_0,i-1}(y) + t(y,x)\} & i > 0. \end{cases} \tag{2.15}$$

Note that $\lambda_{x_0,i}$ is independent of N, and is continuous in x and x_0 for $i > 0$.

Since $\log \mu_{x_0,0}(x) = -N\lambda_{x_0,0}(x)$, we have by induction,

$$\log \mu_{x_0,i}(x) = -N\lambda_{x_0,i}(x) + o(iN). \tag{2.16}$$

Unfortunately, one cannot just take the limit as $i \to \infty$ in (2.16) to obtain the stationary distribution. Indeed, although $\mu_{x_0,i}$ tends to a limit which is independent of x_0, the limit of the $\lambda_{x_0,i}$ does depend on x_0. The problem is that there may exist long lived metastable states at certain densities, so that the convergence of $\mu_{x_0,i}$ to the stationary distribution is very slow (see Theorem 3.1). The $o(iN)$ error term is then too large for (2.16) to be useful. Instead, we shall piece together $\lambda(x)$ from the limits of $\lambda_{x_0,i}$ for various different values of x_0.

Step 2. Prove $\lambda_{x_0,i}(x)$ converges as $i \to \infty$.

Assume x_0 is a fixed point of f_m, so $f_m(x_0) = x_0$ and $t(x_0, x_0) = 0$. Then $\lambda_{x_0,1}(x) \leq \lambda_{x_0,0}(x)$ for all x, so by induction on i, $\lambda_{x_0,i}(x)$ is monotonically decreasing in i. Since it is also bounded below (by 0), it tends pointwise to a limit $\lambda_{x_0}(x)$ and

$$\lambda_{x_0}(x) = \inf_{y\in[0,1]} \{\lambda_{x_0}(y) + t(y, x)\}. \tag{2.17}$$

Since $t(y, x)$ is continuous in x, uniformly in y, any solution of (2.17) is continuous. Also, as $\lambda_{x_0,i}(x)$ decreases monotonically to a continuous function on the compact set $[0, 1]$, this convergence must be uniform in x.

Now assume x_0 is not a fixed point, but x is. By considering $y = x$ in (2.15), we see that $\lambda_{x_0,i}(x)$ is monotonically decreasing in i for this particular x (however it will generally not be decreasing at other points). The limit $\lambda_{x_0}(x)$ is therefore also defined for all fixed points x and arbitrary x_0. Define $\lambda'_{x_0}(x)$ for all $x, x_0 \in [0, 1]$ by

$$\lambda'_{x_0}(x) = \min_{f_m(y)=y} \{\lambda_y(x) + \lambda_{x_0}(y)\}.$$

By (2.15) and induction on i, $\lambda_{x_0,i}(x)$ is just the minimum over all sequences $x_0 = y_0, y_1, y_2, \ldots, y_i = x$ of $\sum_{j=0}^{i-1} t(y_j, y_{j+1})$. From this it can be deduced in general that

$$\lambda_{x_0,i+j}(x) \leq \lambda_{y,i}(x) + \lambda_{x_0,j}(y).$$

Hence $\limsup_i \lambda_{x_0,i}(x) \leq \lambda'_{x_0}(x)$. It remains to show that $\liminf_i \lambda_{x_0,i}(x) \geq \lambda'_{x_0}(x)$ since then $\lambda_{x_0,i}(x)$ will converge to $\lambda_{x_0}(x) = \lambda'_{x_0}(x)$ for all x and x_0.

Fix x and $\varepsilon > 0$, and assume i is large. Pick $x_0 = y_0, y_1, y_2, \ldots, y_i = x$ so that $\lambda_{x_0,i}(x) = \sum_{j=0}^{i-1} t(y_j, y_{j+1})$. Suppose that $|f_m(y_j) - y_j| \geq 3\varepsilon$ for all $j = 1, \ldots, i-1$. Since $\lambda_{x_0,i}(x)$ is bounded, and by (2.13), if i is sufficiently large then $|f_m(y_j) - y_{j+1}| < \varepsilon$ for all but at most $k = \varepsilon i - 2$ values of j. In particular, there must be a sequence of at least $\frac{i-1}{k+1} - 1 \geq 1/\varepsilon - 1$ consecutive values of j, $0 < j < i$, for which $|f_m(y_j) - y_{j+1}| < \varepsilon$. Without loss of generality, we may suppose $f_m(y_j) \geq y_j$ and hence $f_m(y_j) - y_j \geq 3\varepsilon$ for the first such j. Then $y_{j+1} > y_j + 2\varepsilon$ and so $f_m(y_{j+1}) - y_{j+1} \geq f_m(y_j) - y_{j+1} > -\varepsilon$ by monotonicity of f_m. Since $|f_m(y_{j+1}) - y_{j+1}| \geq 3\varepsilon$ we have $f_m(y_{j+1}) - y_{j+1} \geq 3\varepsilon$, and so by induction, $y_{j+1} > y_j + 2\varepsilon$ for each j in the sequence. Thus we obtain a sequence of y_j of length $1/\varepsilon$, each more than 2ε larger than the last. This is clearly impossible in $[0, 1]$, so there must be some y_j, $0 < j < i$, with $|f_m(y_j) - y_j| < 3\varepsilon$.

Hence, provided ε is small enough, some such y_j is close to a fixed point y. But then $\lambda_{x_0,i}(x) = \lambda_{y_j,i-j}(x) + \lambda_{x_0,j}(y_j)$ is close to $\lambda_{y,i-j}(x) + \lambda_{x_0,j}(y) \geq \lambda_y(x) + \lambda_{x_0}(y) \geq \lambda'_{x_0}(x)$, and we are done.

Thus $\lambda_{x_0,i}(x)$ converges uniformly in (x, x_0) to a continuous function $\lambda_{x_0}(x)$, which is the minimum of some expressions of the form $\lambda_y(x) + k_y$ where $f_m(y) = y$ and $k_y = \lambda_{x_0}(y)$ are constants depending on x_0.

Step 3. Approximate the stationary distribution.

Fix $\varepsilon > 0$. Then there is an M such that $|\lambda_{x_0,M}(x) - \lambda_{x_0}(x)| < \varepsilon$ for all x_0 and x. Hence, for sufficiently large N, (2.16) implies

$$\left| \tfrac{\log \mu_{x_0,M}(x)}{N} + \lambda_{x_0}(x) \right| < 2\varepsilon.$$

If μ is the stationary distribution of the mean field model, then we have $\mu(x) = \sum_{x_0} \mu(x_0)\mu_{x_0,M}(x)$. Hence

$$\left| \tfrac{\log \mu(x)}{N} + \lambda^{(N)}(x) \right| < 3\varepsilon,$$

where $\lambda^{(N)}(x) = \inf_{x_0}\{\lambda_{x_0}(x) - \log \mu(x_0)/N\}$ (once again approximating a log of a sum by an supremum). Since $\lambda_{x_0}(x)$ can be written in terms of $\lambda_y(x)$ where $f_m(y) = y$, we can write

$$\lambda^{(N)}(x) = \min_{f_m(y)=y} \{\lambda_y(x) + k_y^{(N)}\}. \tag{2.18}$$

The choice of constants $k_y^{(N)}$ corresponds to a point in $\mathbb{R}^k$ where k is the (finite) number of fixed points of f_m. Indeed, we can clearly take this point in the bounded region $[0, K]^k$ where $K = \sup_{x,y} t(y, x)$. Hence as $N \to \infty$ there will be accumulation points (k_y) and functions

$$\lambda(x) = \min_{f_m(y)=y} \{\lambda_y(x) + k_y\} \tag{2.19}$$

such that $\mu(x) = \exp(-N\lambda(x) + o(N))$ for infinitely many N. We shall fix some increasing sequence N_i for which $k_y^{(N_i)} \to k_y$. We shall show that we can determine k_y uniquely, thus showing that there is only one accumulation point and $\mu(x) = \exp(-N\lambda(x) + o(N))$ for all N.

Step 4. Balancing conditions.

We now find relationships between the k_y. Fix some fixed point x_0 of f_m and consider $\lambda_{x_0,i}(x)$. We shall show by induction on i that $\lambda_{x_0,i}(x)$ is increasing for $x > x_0$. Recall that $\lambda_{x_0,i}(x) \leq \lambda_{x_0,i-1}(x)$, and by (2.15), $\lambda_{x_0,i}(x) \leq \lambda_{x_0,i-1}(y_0)$ where $f_m(y_0) = x$. Set $\tilde{x} = \min\{x, y_0\}$ so that $\lambda_{x_0,i}(x) \leq \lambda_{x_0,i-1}(\tilde{x})$. Suppose $x' > x$ and $\lambda_{x_0,i}(x') < \lambda_{x_0,i}(x)$. Then since $\lambda_{x_0,i-1}(y)$ is assumed to be increasing for $y \geq \tilde{x} > x_0$, the infimum in (2.15) for x' must occur at some $y' < \tilde{x}$. But $t(y', z)$ is strictly increasing for $z > f_m(y')$ and $f_m(y') \leq f_m(\tilde{x}) \leq x$. Thus $t(y', x') > t(y', x)$, and the infimum in (2.15) is smaller for x than for x', a contradiction. Hence $\lambda_{x_0,i}(x)$ is increasing for $x > x_0$. Taking limits, we see that $\lambda_{x_0}(x)$ is increasing for $x > x_0$. Similarly $\lambda_{x_0,i}(x)$ is decreasing for $x < x_0$.

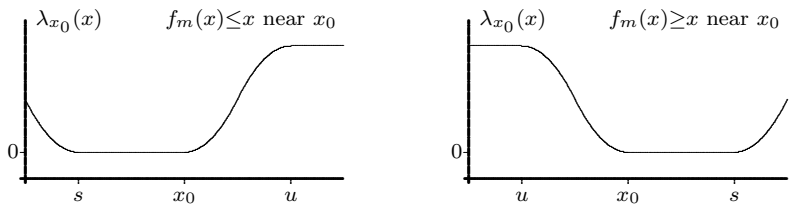


FIGURE 1. x_0 lying between stable (s) and unstable (u) fixed points.

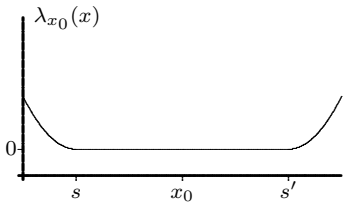


FIGURE 2. Unstable x_0 lying between stable points s, s'.

We can in fact say more. If $x > x_0$ and $f_m(x) < x' < x$ then by (2.17), $\lambda_{x_0}(x) - \lambda_{x_0}(x') \geq \inf_{y\leq x}(t(y,x) - t(y,x')) > 0$. Thus $\lambda_{x_0}(x)$ is strictly increasing on intervals where $f_m(x) < x$. Conversely, if $x < f_m(x)$ then $\lambda_{x_0}(x) \leq \lambda_{x_0}(f_m(x)) \leq \lambda_{x_0}(x)$ and so $\lambda_{x_0}(x)$ is locally constant. Similarly, if $x < x_0$ then $\lambda_{x_0}(x)$ is strictly decreasing with increasing x when $f_m(x) > x$, and locally constant when $f_m(x) < x$.

If $f(x_0) = x_0$, then x_0 is a stable fixed point if $f(x) - x$ changes from positive to negative as x increases, and is an unstable fixed point if $f(x) - x$ changes from negative to positive. If $f(x) - x$ does not change sign, we shall call x_0 an *indifferent* fixed point.

Suppose that x_0 is an indifferent fixed point and consider the mean field model for large N which is initially in a state with distribution given approximately by $\exp\{-N\lambda_{x_0}(x)\}$. If $f_m(x) \leq x$ near $x = x_0$ then we shall show that as t increases, there will a net 'flow' from the region $(x_0 - \varepsilon, 1]$ to $[0, x_0 - \varepsilon)$, i.e., the probability that ρ_t lies in $[0, x_0 - \varepsilon)$ will increase as t increases. Let s be the largest stable fixed point less than x_0, i.e., $s = \sup\{x \in [0, x_0] : f_m(x) > x\}$. Then $\lambda_{x_0}(x) = 0$ iff $x \in [s, x_0]$ (see Figure 1). Hence, if N is large, the distribution is concentrated almost entirely in $[s, x_0]$. Pick $\varepsilon > 0$ so that there is no fixed point in $[x_0 - \varepsilon, x_0)$. For some points x just below $x_0 - \varepsilon$, the expression $\lambda_{x_0}(y) + t(y,x)$ is minimized for some $y = f_m^{-1}(x) > x_0 - \varepsilon$. However, for $x \geq x_0 - \varepsilon$, $\lambda_{x_0}(y) + t(y,x)$ is bounded below by some positive constant for $y \leq x_0 - \varepsilon$, so there is no equivalent matching flow back across $x_0 - \varepsilon$. Hence this cannot be the stationary distribution. In the real stationary distribution, there must be a counter flow given in $\lambda(x)$ by some $\lambda_{x_1}(x) + k_{x_1}$ at $x = x_0$. For this to be sufficient for all N_i and all $\varepsilon > 0$ we need $\lambda_{x_1}(x_0) + k_{x_1} \leq \lambda_{x_0}(x_0) + k_{x_0} = k_{x_0}$. However, $\lambda_{x_1}(x) \leq \lambda_{x_1}(x_0) + \lambda_{x_0}(x)$, so $\lambda_{x_1}(x) + k_{x_1} \leq \lambda_{x_0}(x) + k_{x_0}$ for all x. Hence the $\lambda_{x_0}(x) + k_{x_0}$ term is redundant in (2.19). A similar argument holds if $f_m(x) \geq x$ near $x = x_0$.

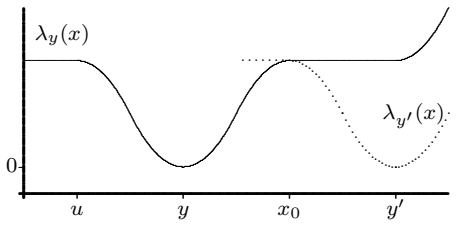


FIGURE 3. Stable y, y' and unstable u, x_0.

Now assume x_0 is an unstable fixed point of $f_m(x)$ (see Figure 2). Using a similar argument to above, there will be a net drift out of the region $(x_0-\varepsilon, x_0+\varepsilon)$ since for $x \neq x_0$ the the expression $\lambda_{x_0}(y)+t(y,x)$ is only minimized for $|y-x_0| < |x-x_0|$. To counter this flow in the stationary distribution we must have $\lambda_{x_1}(x_0) + k_{x_1} \leq \lambda_{x_0}(x_0) + k_{x_0}$ for some fixed point $x_1 \neq x_0$. But then as before, $\lambda_{x_1}(x) + k_{x_1} \leq \lambda_{x_0}(x) + k_{x_0}$ and the $\lambda_{x_0}(x) + k_{x_0}$ term in (2.19) can be removed.

Inductively removing all redundant y's from the minimum in (2.19), we can assume the only contributing expressions occur from stable fixed points. (Note that the x_1 in both the cases above can always be chosen to be a fixed point that has not been eliminated so far).

Now consider an unstable fixed point x_0. If the value of $\lambda(x_0)$ is given by $\lambda_y(x_0)+k_y$ where y is a stable point with $y < x_0$, then as before we see that there is a leftward flux near x_0 (see Figure 3). To counter this we must have a term $\lambda_{y'}(x) + k_{y'}$ with y' a stable fixed point $y' > x_0$ and $\lambda_{y'}(x_0) + k_{y'} \geq \lambda_y(x_0) + k_y$. By symmetry, we see that the minimum in (2.19) must be attained for at least one stable fixed point y with $y < x_0$, *and* at least one stable fixed point y with $y > x_0$.

Let $s_1 < s_2 < \cdots < s_r$ be the stable fixed points and $u_1 < \cdots < u_{r-1}$ the unstable fixed points of $f_m(x)$. Then

$$0 < s_1 < u_1 < s_2 < \cdots < u_{r-1} < s_r < 1.$$

Since at u_i the minimum $\lambda(x) = \min_y\{\lambda_y(x) + k_y\}$ occurs for some $y \geq s_{i+1}$ and some $y \leq s_i$, we see that on the interval (s_i, u_i), $\lambda(x) = \min_{y\leq s_i}\{\lambda_y(x) + k_y\}$, and so is strictly increasing, and on (u_i, s_{i+1}), $\lambda(x) = \min_{y\geq s_{i+1}}\{\lambda_y(x) + k_y\}$, and so is strictly decreasing. For each i, the sequence of values $a_{i,j} = \lambda_{s_i}(u_j)$ satisfies $a_{i,1} > \cdots > a_{i,i-1}$ and $a_{i,i} < \cdots < a_{i,r-1}$. Theorem 2.2 now follows from the uniqueness result in the following lemma, together with the fact that $\inf_x \lambda(x)$ must be zero. □

Lemma 2.5. *Assume $a_{i,j} \in \mathbb{R}$, $i = 1, \ldots, r$, $j = 1, \ldots, r-1$, and for all i, $a_{i,1} > \cdots > a_{i,i-1}$ and $a_{i,i} < \cdots < a_{i,r-1}$. Then there exist $k_i \in \mathbb{R}$ so that, setting $b_{i,j} = a_{i,j} + k_i$, for each j there are integers $i_0 \leq j < i_1$ such that*

$$\min_i b_{i,j} = b_{i_0,j} = b_{i_1,j}. \tag{2.20}$$

Moreover, the k_i are unique up to the addition of a constant c, independent of i.

Proof. For $r = 2$ the result is trivial, so we shall assume $r > 2$ and prove the result by induction on r. First we note that if there is any solution then there is a solution with all $k_i \in [0, K]$ where $K = \sum_j (\max_i a_{i,j} - \min_i a_{i,j})$. Moreover, if we weaken the inequalities to $a_{i,1} \geq \cdots \geq a_{i,i-1}$ and $a_{i,i} \leq \cdots \leq a_{i,r-1}$, then the set of points $((a_{i,j}), (k_i)) \in \mathbb{R}^{3r-1}$ for which (k_i) is such a solution is closed. Hence it is enough to prove existence for a dense set of $(a_{i,j})$.

We consider the case when all the $a_{i,j}$ are rational numbers. Exclude all $(a_{i,j})$ for which there exists a *cycle*, i.e., distinct $j_0, \ldots, j_{s-1}$ and distinct $i_0, \ldots, i_{s-1}$ such that $\sum_{t=0}^{s-1}(a_{i_t,j_t} - a_{i_{t+1},j_t}) = 0$ (where $i_s = i_0$ and $j_s = j_0$). This can be done inductively in j, modifying the $a_{i,j}$ slightly so that none of the differences $a_{i,j} - a_{i',j}$ are equal to the finite set of possible sums of differences with smaller j's. The set of all possible $(a_{i,j})$ is still dense in the appropriate region of $\mathbb{R}^{2r-1}$. Now by linearity it is enough to prove the result when the $a_{i,j}$ are integers.

We first prove the existence of the k_i without the restriction that $i_0 \leq j < i_1$. In other words, we need to choose k_i so that for each j there are (at least) two values of i minimizing $b_{i,j} = a_{i,j} + k_i$. Initially set all the $k_i = 0$. For each j let S_j be the set of i such that $b_{i,j} = \min_l b_{l,j}$ and consider the number s of 'bad' j's for which $|S_j| = 1$. If $s = 0$ then we are done. Otherwise pick a bad j, say j_0, with $S_{j_0} = \{i_0\}$. Define a (multi-)graph G on the vertices $\{1, \ldots, r\}$ with edges $i'i''$ whenever there is a j such that $S_j = \{i', i''\}$. Clearly G has at most $r - 1 - s$ edges, so is disconnected. Also G contains no cycle, since otherwise there would exist $j_0, \ldots, j_s = j_0$ and $i_0, \ldots, i_s = i_0$ with $S_{j_t} = \{i_t, i_{t+1}\}$ and hence $\sum_t (a_{i_t,j_t} - a_{i_{t+1},j_t}) = \sum_t (b_{i_t,j_t} - b_{i_{t+1},j_t}) = 0$. Indeed, if $\{i', i''\} \subseteq S_j$ there cannot be any path from i' to i'' in G avoiding the edge given by S_j (if it exists). Thus, for each j, either every element of S_j is in a distinct component of G, or $S_j = \{i', i''\}$ and $i'i''$ is an edge of G whose removal increases the number of components in G.

Now increase k_i by 1 for all i in the component of i_0. If $|S_j| = 2$ then either both or neither k_i, $i \in S_j$, is increased, so S_j is a subset of the new S_j defined by the updated k_i. If $|S_j| > 2$ then k_i is increased for at most one $i \in S_j$, so the new S_j loses at most one element. Hence the number of bad j does not increase. The value of b_{i_0,j_0} increases by 1, but not all b_{i,j_0} increase, so the average of the b_{i,j_0} increases by less than 1. Hence if we repeat this process, eventually b_{i_0,j_0} will no longer be the unique minimal b_{i,j_0} and when this occurs, the two smallest values of b_{i,j_0} will become equal. Thus s reduces and the result now follows by induction on s.

We shall now show that this solution satisfies the extra condition that for each j, $\min b_{i,j} = b_{i_0,j} = b_{i_1,j}$ with $i_0 \leq j < i_1$. Without loss of generality we can (by adding k_i to $a_{i,j}$) assume $k_i = 0$ and $b_{i,j} = a_{i,j}$.

Pick j_0 with a minimal value of $\min_i a_{i,j_0}$ and suppose this minimum occurs at $i = i_0$. Thus a_{i_0,j_0} is the minimum $a_{i,j}$ over all i and j. Now by the inequalities on $a_{i,j}$ we get $a_{i_0,i_0} < \cdots < a_{i_0,j_0}$ if $i_0 < j_0$ and $a_{i_0,j_0} > \cdots > a_{i_0,i_0-1}$ if $i_0 > j_0 + 1$. These contradict the minimality of a_{i_0,j_0}. Hence $i_0 = j_0$ or $i_0 = j_0 + 1$. Since there must be at least two minimal a_{i,j_0}, we get $a_{j_0,j_0} = a_{j_0+1,j_0} = \min_i a_{i,j_0}$. Thus the result holds for $j = j_0$.

Now for each j, replace $a_{j_0,j}$ with $\min\{a_{j_0,j}, a_{j_0+1,j}\}$ and remove the row $i = j_0 + 1$ and column $j = j_0$ from the matrix $a_{i,j}$. The new matrix satisfies the conditions of the lemma with r replaced with $r - 1$. The cycle condition $\sum_t (a_{i_t,j_t} - a_{i_{t+1},j_t}) \neq 0$ still holds for the new matrix. Indeed, if a cycle occurred in the new matrix, then one would obtain a cycle in the original unless some $j_t = j_0$ and the terms a_{i_t,j_t} and a_{i_{t+1},j_t} in the new matrix come from distinct columns j_0 and $j_0 + 1$ in the original matrix. However, since $a_{j_0,j_0} = a_{j_0,j_0+1}$ and row j_0 was deleted in the new matrix, we can use, e.g., $a_{i_t,j_0} - a_{i_{t+1},j_0+1} = (a_{i_t,j_0} - a_{j_0,j_0}) + (a_{j_0,j_0+1} - a_{i_{t+1},j_0+1})$ to extend the cycle to a cycle in the original matrix.

By the cycle condition in the original matrix with $j_0 = j_0$, $j_1 = j$, and $i_0 = j_0$, $i_1 = j_0 + 1$, we have $a_{j_0,j} \neq a_{j_0+1,j}$ for all $j \neq j_0$. Hence $k_i = 0$ is still a solution to (2.20), since the number of i for which b_{i,j_0} is minimal is not reduced when we combined rows j_0 and $j_0 + 1$. Now by induction on r we have $\min_i a_{i,j} = a_{i_0,j} = a_{i_1,j}$ with $i_0 \leq j < i_1$ and this implies the same result for the original matrix when $j \neq j_0$.

Finally, we need to show that for arbitrary $a_{i,j}$ the solution is unique. Suppose without loss of generality $k_i = 0$ is one solution and there is one other solution $(k'_i)_{i=1}^r$. By adding a constant to all k'_i we can assume that $\min k'_i = 0$. Let $S = \{i : k'_i > 0\}$ and set $k''_i = \varepsilon$ if $i \in S$ and $k''_i = 0$ if $i \notin S$. We assume ε is smaller than any positive k'_i so that $k''_i \leq k'_i$ for all i. Defining $S_j = \{i : a_{i,j} = \min_l a_{l,j}\}$ as above, the fact that k'_i is a solution implies that $S_j \setminus S$ is either empty, or contains i_0, i_1 with $i_0 \leq j < i_1$. Thus k''_i is also a solution with the same S_j when $S_j \setminus S = \emptyset$ and $S_j \setminus S$ otherwise.

Pick a minimal $a_{i,j}$ and call it a_{i_0,j_0}. As above, $i_0 \in S_{j_0} = \{j_0, j_0+1\}$. Clearly $j_0 \in S$ iff $j_0+1 \in S$. Hence the reduction combining rows j_0 and j_0+1 and eliminating column j_0 described above gives a new matrix with solutions k''_i which are ε for some i, and 0 for some other i. Note that this is indeed a solution, since if combining rows j_0 and j_0+1 reduces the number of minimal values of $a_{i,j}+k''_i$ for $i \neq j_0$, then it must combine two minimal values with $j < i$ or two minimal values with $j \geq i$. Since there must be at least one minimal value with $j < i$ and one with $j \geq i$, this solution still satisfies the requirements of the lemma for the reduced matrix. In particular, the solution for the reduced matrix is not unique (since we can vary ε and $\emptyset \neq S \neq \{1, \ldots, r\}$). By induction on r this is impossible, hence the original solution was unique up to the addition of a constant. □

3. Consequences of Theorem 2.2

One consequence of Theorem 2.2 is that the expected time to jump from one stable fixed point of f_m to another is always exponential in N.

Theorem 3.1. *Assume the same conditions as in Theorem 2.2 and let x_0, x_1 be two distinct stable fixed points of f_m. For sufficiently small $\varepsilon > 0$, the expected time to hit some density in $[x_1 - \varepsilon, x_1 + \varepsilon]$ starting with any density in $[x_0 - \varepsilon, x_0 + \varepsilon]$ is bounded above and below by functions of the form $\exp(cN)$.*

Proof. For the upper bound, note that the probability of jumping from anywhere to $[x_1-\varepsilon, x_1+\varepsilon]$ in one step is bounded below by $\exp(-cN+o(N))$ when $c = \sup_{x,y} t(y, x)$.

Hence the expected time to hit $[x_1-\varepsilon, x_1+\varepsilon]$ is at most $\exp(cN+o(N))$. For the lower bound we show that the expected time to leave $I=[x_0-3\varepsilon, x_0+3\varepsilon]$ is exponential in N, where we assume that ε sufficiently small so that for each point $x\in I$, $|f_m(x)-x_0|\le |x-x_0|$. To see this, observe that the probability of jumping in one step to a point at least ε further away from x_0 is bounded above by $\exp(-N\delta+o(N))$, where $\delta>0$ is chosen so that (2.13) holds. On the other hand, the expected time to leave by passing through a point in $I'=[x_0-3\varepsilon, x_0-2\varepsilon]\cup[x_0+2\varepsilon, x_0+3\varepsilon]$ starting in $I''=[x_0-\varepsilon, x_0+\varepsilon]$ is at most $\exp(cN+o(N))$ where $c=\inf_{x\in I', y\in I''}(\lambda(x)-\lambda(y))>0$. □

Theorem 3.1 implies that if x is a stable fixed point of f_m and $\lambda(x)>0$, then the density x is a metastable state for the model, i.e., the model once in this state will stay there for a very long time, even though in the stationary distribution, the density x occurs with very small probability. Indeed, it was the existence of these metastable states that made the proof of Theorem 2.2 so complicated.

Theorem 2.2 holds also for asynchronous models. In this case the function λ can be defined more simply as

$$\lambda(x)=\int \log\left(\tfrac{x(1-f_m^+(x))}{(1-x)f_m^-(x)}\right)\,dx$$

with the constant of integration chosen so that $\inf_{x\in[0,1]}\lambda(x)=0$. Here the $f_m^\pm$ are defined as those for an approximating simultaneous model (1.1), the expression $\frac{x(1-f_m^+)}{(1-x)f_m^-}$ being independent of ε. The proof of this result is significantly simpler that that of Theorem 2.2, so we shall not include the details here. It can be proved by noting that in the asynchronous model, the number of + states can change by only 1 at a time, so in the stationary distribution $\mu(x)$, the frequency of transitions from density r/N to $(r+1)/N$ must match the frequency of transitions from $(r+1)/N$ to r/N.

As the parameters p_ϕ vary, one can obtain several types of phase transition. The simplest examples are due to changes in f_m. As f_m changes we can obtain bifurcations of the stable fixed points, formation of limit cycles, and even transitions to chaos. For example, define a totalistic model with $|\Gamma|=n$ odd and a simple majority rule

$$p_r=\begin{cases} p+h, & r<n/2;\\ 1-p+h, & r>n/2.\end{cases} \tag{3.1}$$

Assume first that $h=0$ so that the model is symmetric. For small p we have two stable states and one unstable state, while for p close to 0.5 we obtain only one stable state. There is a second order phase change at a critical point, which occurs when $f_m'(0.5)=1$. It can be shown that at this point

$$\begin{aligned} p=p_{\text{crit}} &= \frac{1}{2}-\frac{2^{n-2}}{n\binom{n}{0}+(n-2)\binom{n}{1}+\cdots+1\binom{n}{(n-1)/2}}\\ &= \frac{1}{2}-\frac{2^{n-2}}{n\binom{n-1}{(n-1)/2}}=\frac{1}{2}-O(n^{-1/2}). \end{aligned} \tag{3.2}$$

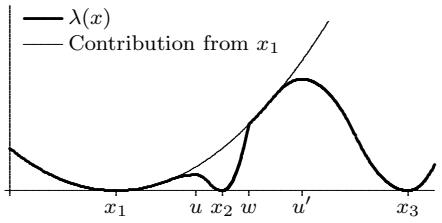
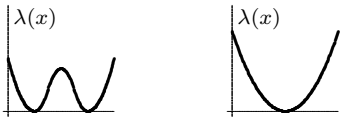
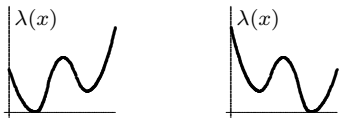


FIGURE 4



$p < p_{\text{crit}}$ $\qquad$ $p > p_{\text{crit}}$

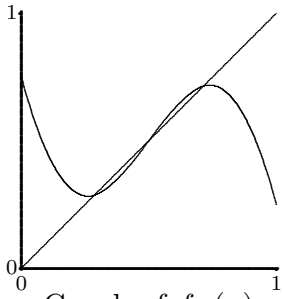
We also have transitions where the preferred fixed points or cycles change. For example, if in the above model we fixed $p < p_{\text{crit}}$, then by varying h we can obtain a first order phase transition at $h = 0$.



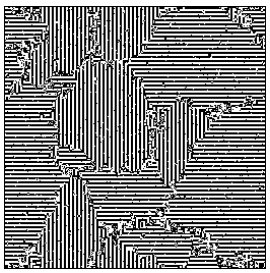
$h < 0$ $\qquad$ $h > 0$

Surprisingly, these first order transitions can occur in semi-totalistic models even with f_m fixed, for we can vary $p_r^{\pm}$ in such a way that f_m remains constant but $f_m^{\pm}$ (and hence $\lambda(x)$) varies.

Another example of strange behavior is given by the following example. For large $|\Gamma|$ we can choose the rules so as to make f_m and g_m approximate any continuous functions satisfying $f_m(x) \in [0,1]$, $h(x)|f_m(x) - x| < g_m(x) \le f_m(x)(1 - f_m(x))$, where $h(x) = \max\{f_m/x, (1 - f_m)/(1 - x)\}$. (For the upper bound on g_m choose $f_m^+ = f_m^- = f_m$, for the lower set either f_m^+ near to 1 or f_m^- near to 0.) Choose f_m so that there are three stable fixed points $x_1 < x_2 < x_3$, with $g_m(x_1)$ much smaller than $g_m(x_2)$ and $g_m(x_3)$, and such that $\lambda(x_1) = \lambda(x_2) = \lambda(x_3) = 0$. This can be achieved by making $f_m(x)$ very close to x so that there is a wide range of values available for g_m. In this case it is possible that there is a greater chance of jumping from x_1 to x_3 than from x_2 to x_3. Figure 4 shows $\lambda(x)$ (thick line), as well as $\lambda_{x_1}(x_1) + t(x_1, x)$ (thin line). This second expression gives the contribution to the infimum in equation (2.14) from $y = x_1$ and corresponds to the chance of jumping in one step from x_1 to x. The points u and u' are unstable fixed points. We see that for $x > w$ the chances are greater of jumping from x_1 to x in a single step are greater than the chances of getting from x_2 to x without going via x_1. Moreover, the function $\lambda(x)$ is not differentiable at $x = w$, despite the fact that $f_m^{\pm}(x)$ are smooth (polynomial) functions of x.



Graph of $f_m(x)$



Model on 256×256 torus

$$p_0 = \tfrac{3}{4} = 1 - p_5, \quad p_1 = p_2 = \tfrac{1}{256} = 1 - p_3 = 1 - p_4.$$

FIGURE 5. Symmetric totalistic model with non-isotropic phases.

4. Contrast with lattice models

Our real interest lies in random cellular automata on $\mathbb{Z}^d$ rather than the mean field models. So one question is how well the mean field models approximate the corresponding random automaton. Since the mean field model is equivalent to that of an semi-totalistic model, the best we could hope for is that it approximates these models. However, even totalistic models can exhibit behavior that is much richer than that of the mean field models.

For example, take the symmetric totalistic model with $p_0 = \frac{3}{4}$, $p_1 = p_2 = \frac{1}{256}$. Figure 5 shows f_m and a typical state of this model on the 256×256 torus after a few thousand iterations. From the graph it is clear that f_m has two stable fixed points and an unstable fixed point at $x = 0.5$. Thus the mean field model will have two phases, one of low density and the other of high density. The automaton does show distinct phases, but they all have densities of 0.5. Indeed there seem to be several phases, none of which are isotropic. This is an example of spontaneous symmetry breaking — the model is symmetric under interchange of coordinates, but some phases consist of horizontal lines and some phases consist of vertical lines. The lines are of width two, so for each orientation there is a choice of alignment modulo 4 in the vertical or horizontal direction, giving 8 possible phases in this model.

For other parameters, symmetric totalistic models often have behavior similar to the Ising model. This is no coincidence, since the Ising model can be represented as the stationary distribution of the asynchronous automaton with rates $r_\phi = r_{\phi,-\phi(\mathbf{0})}$ given by

$$r_\phi \;=\; \exp(2K|\{\mathbf{x} : \phi(\mathbf{x}) \neq \phi(\mathbf{0})\}| - \phi(\mathbf{0})H) \tag{4.1}$$

where K is the inverse temperature and H is the magnetization (see [5]).

For an example, consider the symmetric totalistic model (3.1) with $n = 5$, $h = 0$. For the mean field model (3.2) gives $p_{\text{crit}} = \frac{7}{30} \approx 0.233$, so this has one fixed point at $x = 0.5$ for $p \in [\frac{7}{30}, \frac{1}{2}]$, but for $p < \frac{7}{30}$ the fixed point $x = 0.5$ is unstable and there are two other fixed points which are stable.

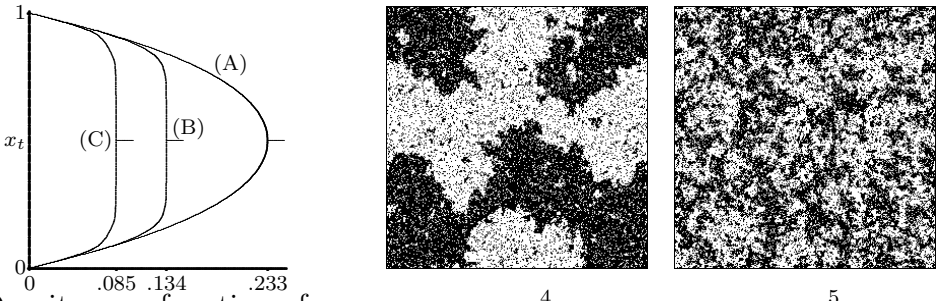


FIGURE 6. Symmetric totalistic model near phase transition. Graph (A) represents the mean field model, (B) the synchronous model, (C) the asynchronous model. Snapshots are given for the synchronous model with p slightly below and slightly above the critical value.

The probabilistic automaton on $\mathbb{Z}^2$ with $\Gamma = \{(\pm 1, 0), (0, \pm 1), (0, 0)\}$ behaves qualitatively similarly, except that the critical probability is significantly lower at about $p_c \approx 0.134$. For $p_c < p \leq 0.5$ it appears that the stationary density distribution of x_t for a sufficiently large but finite lattice is unimodal with peak at $p = 0.5$. For $p < p_c$ the distribution becomes bimodal, as one would expect from the mean field model, and there are two phases in the infinite lattice, one with high density and one with low density. The corresponding asynchronous model behaves similarly, except that the critical probability is even lower, at about $p_c \approx 0.085$ (see Figure 6).

Both of these models have behavior similar to that of the Ising model. In particular, if the modal density is plotted as a function of p, we see a bifurcation at p_c and the modal densities near p_c behave as $|\rho - 0.5| \sim (p_c - p)^\beta$. Numerical evidence suggests that the critical exponent β is close to that of the Ising model ($\beta = \frac{1}{8}$), and is different from that of the mean field model (for which $\beta = \frac{1}{2}$). Other critical exponents for these models appear to match those of the Ising model, supporting the universality hypothesis that critical exponents are exactly the same for large classes of similar models. This is in contrast to the results of [24] where different critical exponents were found in a model based on coupled chaotic maps.

In other models different types of phase transition can occur. We have seen above that a symmetry breaking phase transition can occur in some models. There are also semi-totalistic models where for values of p_ϕ close to zero or one the model seems chaotic, but the introduction of more randomness causes a phase transition to a more ordered phase.

We are lead to the question of what type of phase transition are possible and what are the values of the critical exponents for these transitions as we change the parameters p_ϕ.

References

[1] J. Adler, *Bootstrap percolation*, Physica A **171** (1991) 453–470.

[2] K. Aihara, T. Takabe, and M. Toyoda, *Chaotic neural networks*, Phys. Lett. A **144** (1990) (6–7) 333–340.

[3] M. Aizenman and Lebowitz, *Metastability effects in bootstrap percolation*, J. Phys. A **21** (1988) 3801–3813.

[4] P.N. Balister, B. Bollobás, and A.M. Stacey, *Improved upper bounds for the critical probability of oriented percolation in two dimensions*. Random Struct. Alg. **5** (1994) (4) 573–589.

[5] R.J. Baxter, *Exactly solved Models in Statistical Mechanics*, Academic Press, London (1982).

[6] B. Bollobás and A.M. Stacey, *Approximate upper bounds for the critical probability of oriented percolation in two dimensions based on rapidly mixing Markov chains*. J. Appl. Probability, **34** (1997) no. 4, 859–867.

[7] B. Bollobás, *Random Graphs*, (1985) London; Orlando: Academic Press.

[8] R. Cerf and E.N. Cirillo, *Finite size scaling in three-dimensional bootstrap percolation*. Ann. Probab., **27** (1999) no. 4 1837–1850.

[9] H. Chaté, A. Lemaitre, Ph. Marcq, P. Manneville, *Non-trivial collective behavior in extensively chaotic dynamical systems: an update*. Physica A, **224** (1996) 447–457.

[10] H. Chaté and P. Manneville, *Evidence of collective behavior in cellular automata*, Europhys. Lett., **14** (1991) (5) 409–413.

[11] H. Chaté and P. Manneville, *Collective behaviors in spatially extended systems with local interactions and synchronous updating*, Progr. Theoretical Phys., **87** (1992) (1) 1–60.

[12] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications, 2nd ed.*, (1998) Springer, New York.

[13] A.M.S. Duarte, *Simulation of a cellular automaton with an oriented bootstrap rule*, Physica A, **157** (1989) 1075–1079.

[14] W.J. Freeman, R. Kozma, and P.J. Werbos, *Biocomplexity — Adaptive Behavior in Complex Stochastic Dynamical Systems*, BioSystems, **59** (2001) 109–123.

[15] W.J. Freeman, R. Kozma, *Local-global interactions and the role of mesoscopic (intermediate-range) elements in brain dynamics*, Behav. Brain Sci., **23** (2000) (3) 401.

[16] W.J. Freeman, *Noise-induced first-order phase transitions in chaotic brain activity*, Int. J. Bifurcation & Chaos, **9** (1999) (11) 2215–2218.

[17] M. Gardner, *The Fantastic Combinations of John Conway's New Solitaire Game 'Life'*. Sci. Am., 223 (1970) 120–123.

[18] J. Gravener and E. McDonald, *Bootstrap percolation in a polluted environment*, J. of Statist. Phys., **87** (1997) nos. 3–4 915–927.

[19] G. Grimmett, *Percolation in Fundamental Principles of Mathematical Sciences*, Springer-Verlag, Berlin, (1999) *xiv*+444pp.

[20] G. Grinstein, C. Jayaprakash, and Y. He. *Statistical Mechanics of Probabilistic Cellular Automata*. Phys. Rev. Lett., 55 (1985) 2527–2530.

[21] R. Holley and T.M. Liggett, *The survival of contact processes*, Ann. Probability **6** (1978) 198–206.

[22] T.M. Liggett, *Survival of discrete time growth models, with applications to oriented percolation*, Ann. Applied Probability **5** (1995) 613–636.

[23] R. Kozma et al.,*Dynamical behavior of percolation models of cortical phase transitions*, Proc. 2001 Int. Conf. Nonlinear Theory and Applications NOLTA2001, Sendai, Japan, October 2001 (accepted).

[24] P. Marcq, H. Chaté, and P. Manneville, *Universality in Ising-like phase transitions of lattices of coupled chaotic maps*, Phys. Review E, **55** (1997) no. 3 2606–2627.

[25] T. Mountford, *Rates for the probability of large cubes being non-internally spanned in modified bootstrap percolation*, Probability Theory and Related Fields, **93** (1992) 159–167.

[26] Y. Pomeau, *Periodic behavior in cellular automata*, J. of Statistical Physics, **70** (1993) (5–6) 1379–1382.
[27] R. Schonmann, *On the behaviour of some cellular automata related to bootstrap percolation*, Ann. Probability, **20** (1992) (1) 174–193.

Department of Math Sciences, University of Memphis, Memphis, TN 38152-3240.
E-mail address, P. Balister: pbalistr@msci.memphis.edu
E-mail address, B. Bollobás: bollobas@msci.memphis.edu
E-mail address, R. Kozma: rkozma@memphis.edu