

## Some Rigorous Results for the Greenberg–Hastings Model

Richard Durrett<sup>1</sup> and Jeffrey E. Steif<sup>2</sup>

Received April 9, 1990; revised July 5, 1990

---

In this paper, we obtain some rigorous results for a cellular automaton known as the Greenberg–Hastings Model. The state space is  $\{0, 1, 2\}^{\mathbb{Z}^d}$ . The dynamics are deterministic and discrete time. A site which is 1 changes to 2, a site which is 2 changes to 0, and a site which is 0 changes to a 1 if one of its  $2d$  neighbors is a 1. In one dimension, we compute the exact asymptotic rate at which the system dies out when started at random and compute the topological entropy. In two or more dimensions we show that starting from a nontrivial product measure, the limit exists as  $3m \rightarrow \infty$  and is Bernoulli shift. Finally, we investigate the behavior of the system on a large finite box.

---

**KEY WORDS:** Cellular automata; topological entropy; Bernoulli shifts.

### 1. INTRODUCTION

We consider the following cellular automaton, known as the *Greenberg–Hastings Model* (see Refs. 6 and 12), which is a simplified model for cells in an excitable medium. The state space is  $X = \{0, 1, 2\}^{\mathbb{Z}^d}$ . Sites  $x \in \mathbb{Z}^d$  represent cells which can be excited (1), tired (2), or rested (0). With these interpretations in mind we consider the following deterministic discrete time dynamics on  $X$ . An excited cell is always tired at the next time step. A tired cell always becomes rested. Finally a rested site becomes excited if and only if at least one of its  $2d$  neighbors is excited.

Although our dynamics are completely deterministic, we can obtain a stochastic process by starting with an initial probability distribution on  $X$  and letting the system evolve. Let  $\eta_n \in X$  denote the state of the process at time  $n$ , i.e.,  $\eta_n(x)$  denotes the state of the cell at location  $x$  at time  $n$ . Our

---

<sup>1</sup> White Hall, Department of Mathematics, Cornell University, Ithaca, New York 14853.

<sup>2</sup> White Hall, Department of Mathematics, Cornell University, Ithaca, New York 14853.

first step is to investigate the behavior of  $\eta_n^*$ , the system starting from a product measure in which the states 0, 1, and 2 each have probability  $\frac{1}{3}$ . In Sec. 2 we prove the following theorem:

**Theorem 1.** In  $d = 1$ ,  $\text{Prob}\{\eta_n^*(0) = 1\} \sim \sqrt{2/(27\pi n)}$ .

Here,  $a_n \sim b_n$  is short for  $\lim_{n \rightarrow \infty} a_n/b_n = 1$ . In view of the dynamics,  $\text{Prob}\{\eta_{n+1}^*(0) = 2\} = \text{Prob}\{\eta_n^*(0) = 1\}$  and hence

$$\text{Prob}\{\eta_n^*(0) \neq 0\} \sim \sqrt{8/(27\pi n)}$$

Note that this implies that  $\eta_n^* \rightarrow 0$  in distribution as  $n \rightarrow \infty$ .

In Ref. 4, an analogous result is obtained for a very similar system called the *cyclic cellular automaton*. The method of using random walks to analyze certain cellular automata is introduced in Ref. 4.

From the last result the reader might leap to the conclusion that there are no nontrivial stationary distributions in one dimension. However, there are a large number of them. To see this, observe that a configuration that looks like

...01200012000012012000...

i.e., a collection of 12s separated by at least one 0, just moves to the left at rate 1, so if the locations of the 1s are a stationary sequence we have a stationary distribution. Faced with this bewildering variety of stationary distributions, it is natural to try to identify some with special properties. To do this we turn to the notion of topological entropy, which we will now describe.

Let  $T$  denote the transformation  $T: \eta_0 \rightarrow \eta_1$ . It is clear that  $T$  is continuous when  $X$  is given the product topology. Because of this we can define a notion called the *topological entropy*, denoted by  $h(X, T)$ .  $h(X, T)$  is the supremum over all the stationary distributions of the entropy of the corresponding stationary sequence  $\{\eta_n, n \geq 0\}$ . To compute the topological entropy of  $T$ , we will use an equivalent definition that counts, in some sense, the number of possible evolutions of the system.

For  $d = 1$ , let  $a_{n,m}$  be the number of 0, 1, 2-valued configurations  $\sigma$  on  $[-m, m] \times [0, n - 1]$  that can be extended to a possible evolution of the dynamics, that is, such that there exists  $\tilde{\sigma} \in \{0, 1, 2\}^{\mathbb{Z}^2}$  with  $\tilde{\sigma} = \sigma$  on  $[-m, m] \times [0, n - 1]$  and with  $T$  ( $i$ th row of  $\tilde{\sigma}$ ) =  $i + 1$ st row of  $\tilde{\sigma}$ . The topological entropy is equivalently defined to be  $\sup_m \limsup_{n \rightarrow \infty} (\ln a_{n,m})/n$ . Notice that we do not divide by  $m$ .

In Sec. 3, we prove the following theorem:

**Theorem 2.** In  $d=1$ , the topological entropy of the Greenberg–Hastings Model,  $h(X, T)$ , is  $2 \ln \rho$  where

$$\rho = 1/3 \left( \sqrt[3]{\frac{29 + 9\sqrt{31/3}}{2}} + \sqrt[3]{\frac{29 - 9\sqrt{31/3}}{2}} + 1 \right)$$

is the positive root of  $x^3 - x^3 - 1$ .

Some of the ideas of the proof of Theorem 2 come from Ref. 8 where a somewhat similar system is analyzed. To explain the source of  $\rho$ , consider the stationary distributions we described above. A general result implies that within this class the largest entropy occurs for a Markov chain (see Ref. 11). The only available parameter for the Markov chain is  $p$  = the probability the chain stays at 0. Optimizing over  $p$  gives a stationary measure with entropy  $\ln \rho$ . Intuitively, the entropy can be doubled by putting a copy of the Markov chain on  $x > 0$  and an independent copy of its reflection on  $x < 0$ . The distribution just described is not stationary but there is a sequence of stationary distributions that converge to it and have entropy approaching  $2 \ln \rho$ .

In Sec. 4, we show that the qualitative behavior of the system in  $d \geq 2$  is much different than in  $d = 1$ . Let  $\mu$  be an arbitrary translation invariant product measure on  $X$  in which each state has positive probability and suppose that  $\eta_0$  has distribution  $\mu$ . It is easy to see that in  $d = 1$  the system dies out, i.e., converges in distribution to the all 0s state. Our next result implies that if  $d \geq 2$ , then the system settles into a nontrivial cycle with period three. This theorem is completely analogous to a theorem proved in Ref. 5 for the cyclic cellular automaton. To state this result we need a definition that is copied from Ref. 5. We call a set of distinct points  $x_0, x_1, \dots, x_{n-1}, x_n = x_0$  a *clock* for  $\eta$  if  $x_{i+1}$  is a neighbor of  $x_i$  for each  $i$  and  $\eta(x_i) \equiv i \pmod{3}$  for each  $i$ . Clearly,  $n$  must be a multiple of 6. It is easy to see that a clock has period 3 no matter what the rest of the configuration looks like, and it is not much harder to show

**Lemma 1.** If  $\eta_0$  contains a clock then  $\lim_{n \rightarrow \infty} \eta_{3n}$  exists.

Lemma 1 implies that if  $\eta_0$  contains a clock with probability 1, then  $\eta_{3n}$  has a limit a.s. The last result gives an enormous collection of period 3 limit states, so it is again natural to try to single out some that are special. The definition of topological entropy generalizes in a straightforward way to  $d > 1$  but is  $\infty$  for trivial reasons, so we turn to another approach to construct nice limits.

If  $\eta_0$  has distribution  $\mu$  where  $\mu$  is a translation invariant product measure, then Lemma 1 implies that  $\eta_{3n}$  has a limit a.s., which we denote

by  $\nu$ . Our next result says that the  $\nu$ 's obtained in this way are very nice measures. To state that result we need a definition. Let  $\mathcal{B}$  denote the Borel field on  $X$  generated by its topology and let  $\mu_{p_0, p_1, p_2}$  be product measure on  $X$  in which the  $i$ 's have density  $p_i$ . A translation-invariant measure  $m$  on  $X$  is a *Bernoulli shift* if for some  $p_0, p_1, p_2$ , there is a bimeasurable measure-preserving transformation

$$f: (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, \mu_{p_0, p_1, p_2})$$

which commutes with all special translations.

**Theorem 3.** The limit measure  $\nu$  is a Bernoulli shift.

If  $\mu$  is any translation invariant product measure, we conjecture that the corresponding limit distribution  $\nu$  satisfies  $T\nu = \nu$ , that is,  $\nu$  is actually stationary. In support of this conjecture, we note that computer simulations indicate that the density of 0s, 1s, 2s approach  $\frac{1}{3}$  under iteration.

Our final topic, which is treated in Sec. 5, was motivated by simulating the system on a computer. If we consider the model on  $B_L = \{0, \dots, L - 1\}^d$  with periodic boundary conditions and initial state  $\mu_{1/3, 1/3, 1/3}$  then with high probability the system settles rapidly into a period 3 state. To get more interesting pictures, it is natural to start with  $\mu_{1-2\delta, \delta, \delta}$  where  $\delta$  is small. A little experimentation reveals that depending on the size of  $\delta$  we get one of three limiting behaviors: (a) The system dies out, (b) it approaches a period 4 state, or (c) it approaches a period 3 state. Denote the events in (a), (b), and (c) by  $D, E_4,$  and  $E_3$ , and let  $F = (D \cup E_3)^c$ . Writing  $\text{Prob}_L$  to indicate the dependence on  $L$ , our result can be stated as the following:

**Theorem 4.** Let  $d \geq 2$ ,  $\beta(d) = 12\binom{d}{2} + 60\binom{d}{3}$  and  $L \rightarrow \infty$  with  $\delta_L \rightarrow 0$ .

If  $\delta_L L^{d/4} \rightarrow \lambda \in (0, \infty]$ , then

$$\begin{aligned} \text{Prob}_L(E_3) &\rightarrow 1 - e^{-\beta(d)\lambda^4} \\ \text{Prob}_L(F) &\rightarrow e^{-\beta(d)\lambda^4} \end{aligned}$$

If  $\delta_L L^{d/4} \rightarrow 0$  and  $\delta_L L^{d/2} \rightarrow \lambda \in [0, \infty]$ , then

$$\begin{aligned} \text{Prob}_L(F) &\rightarrow 1 - e^{-2d\lambda^2} \\ \text{Prob}_L(D) &\rightarrow e^{-2d\lambda^2} \end{aligned}$$

To explain the first pair of results, we observe that the shortest clock has length 6 (when  $L \geq 4$ ) and requires four nonzero sites. It is easy to see

that there are  $C(d)L^d$  ways to find a clock in  $B_L$ . Since the occurrence of clocks in disjoint sets of sites are independent, it should be easy to believe that when  $\delta_L \sim \lambda/L^{d/4}$  the limiting number of clocks is Poisson, and the Chen–Stein method as explained in Ref. 1 allows us to easily prove this. Finally we show that  $E_3$  occurs if and only if there is a clock initially.

To prove the second pair, we show that the system dies out if there are no 12 pairs (i.e., a 1 and a 2 at adjacent sites). Calculating the probability of this event and applying the Chen–Stein method gives the rest of the conclusion. We believe that this result holds with  $F$  replaced by  $E_4$ . The last conjecture is supported by simulations but we do not know how to prove it. Finally, it is interesting to compare the result in Theorem 4 with the pictures in Ref. 10. On p. 50 of Ref. 10, the authors consider  $L = 256$  and  $\delta = .03$  in  $d = 2$ . In this case  $\delta L^{1/2} = 0.48$ , so our theorem predicts a Poisson number of clocks with mean about  $12/16$ , while the picture shows two clocks.

The remainder of the paper is devoted to proofs. Sections 2–5 are almost independent of each other and can be read in any order.

## 2. ASYMPTOTIC RATE OF DYING OUT FOR $d = 1$

In this section, we prove Theorem 1. Before starting the proof, we need to review some results concerning general random walks on  $\mathbf{Z}$ . Let  $F$  be a distribution on  $\mathbf{Z}$  which has mean 0 and finite variance. Let  $\{X_1, X_2, \dots\}$  be an i.i.d. sequence of random variables with distribution  $F$ . Then  $S_0 = 0, S_1 = X_1, S_2 = X_1 + X_2, \dots$  is called the random walk with step size distribution  $F$ . By combining Theorems XII.7.1a, XII.7.4, XIII.5.5, and XVIII.5.1 in Ref. 3, one obtains the following theorem:

**Theorem 5.** If  $S_n$  is the random walk with step size distribution  $F$ , then  $\text{Prob}\{S_1 > 0, S_2 > 0, \dots, S_n > 0\} \sim \sigma((2\pi n)^{1/2} E[S_N])^{-1}$  where  $\sigma$  is the standard deviation of the distribution  $F$ ,  $N$  is the stopping time corresponding to the first entrance into  $\{1, 2, \dots\}$ , and  $E$  denotes expected value.

We can now proceed with the proof of Theorem 1.

*Proof of Theorem 1.* We write  $\eta_n$  for  $\eta_n^*$ . We first need to construct an auxiliary random walk. The distribution for the steps of this walk, which we will still denote by  $F$ , will be the distribution of  $1 - \#$  of 01s between two successive 10s when we choose a random element  $\eta_0$  from  $X$ . Clearly this distribution is concentrated on  $\{1, 0, -1, -2, \dots\}$ . The mean and variance of  $F$  can be explicitly computed. This is stated in the following lemma which will be proved later.

**Lemma 2.** The mean and variance of  $F$  are 0 and  $\frac{2}{3}$ , respectively.

Let  $\bar{S}_n$  denote the random walk whose steps are distributed according to  $F$ . Clearly  $\bar{S}_N = 1$  where  $N$  is the first entrance into  $\{1, 2, \dots\}$ . It then follows from Theorem 5 that

$$\text{Prob}\{\bar{S}_1 > 0, \bar{S}_2 > 0, \dots, \bar{S}_n > 0\} \sim 1/\sqrt{3\pi n} \tag{2.1}$$

Returning to our original system, if there is a 1 at lattice point 0 at time  $n$ , it clearly must have originated either from lattice point  $-n$  or  $n$ . Let  $E_n$  denote the event that there is a 1 at lattice point 0 at time  $n$  that originated from lattice point  $-n$ . So

$$E_n = \{\eta_0(-n) = 1, \eta_1(-n+1) = 1, \eta_2(-n+2) = 1, \dots, \eta_n(0) = 1\}$$

We first show that

$$\text{Prob}(E_n) \sim 1/\sqrt{54\pi n} \tag{2.2}$$

which is the heart of the argument.

The method by which we will analyze our system is by constructing an induced process on the shifted lattice  $\mathbf{Z} + \frac{1}{2}$  that takes values in  $\{r, l, 0\}$ ,  $r$  for moving right and  $l$  for moving left. Given an element  $\eta$  of  $X$ , we construct a configuration on the shifted lattice with values in  $\{r, l, 0\}$  as follows. At  $n + \frac{1}{2}$ , we place an  $r$  if  $\eta(n) = 1$  and  $\eta(n+1) = 0$ , and  $l$  if  $\eta(n) = 0$  and  $\eta(n+1) = 1$ , and a 0 otherwise. One can easily see that the events  $E_n$  and

{at time 0, there is an  $r$  at  $-n + \frac{1}{2}$  and there are always strictly more  $r$ 's than  $l$ 's on  $[-n + \frac{1}{2}, n - \frac{3}{2}]$  counting from the left}

are the same. When an element  $\eta_0$  is chosen at random from  $X$ , we obtain an induced measure  $\nu$  on the  $\{r, l, 0\}$  process. These configurations lie on the shifted lattice, but by moving everything over by  $\frac{1}{2}$ , our induced measure sits on  $\{r, l, 0\}^{\mathbf{Z}}$ . Clearly  $\nu$  is a stationary ergodic measure but moreover it can be easily seen to be a Markov renewal process with the process restarting itself whenever an  $r$  or  $l$  is hit. Let  $Y_k$  denote this Markov renewal process taking values in  $\{r, l, 0\}$ . From the above discussion and translation invariance, we have that

$$\text{Prob}(E_n) = \text{Prob}\{Y_0 = r \text{ and on } [0, 2n - 2], \text{ the } Y_k \text{ process always has strictly more } r\text{'s than } l\text{'s counting from the left}\}$$

Conditioning on  $Y_0 = r$ , we let  $\tau_0 = 0$  and  $\tau_1, \tau_2, \tau_3, \dots$  be the successive times at which an  $r$  occurs. Since the process restarts itself whenever an  $r$

is hit, it is clear that conditional on  $Y_0 = r$ , the differences  $\tau_{i+1} - \tau_i$ 's form an i.i.d. sequence. Since  $\text{Prob}\{Y_0 = r\} = \frac{1}{9}$ , it follows from a standard argument, using the stationarity, that the mean of the differences is 9. Moreover the common distribution of the differences is clearly stochastically dominated by a geometric random variable and hence the moment generating function exists in a neighborhood of 0. It follows from large-deviation theory that for all  $\varepsilon > 0$ , there is a  $C_\varepsilon > 1$  such that, for all  $n$ ,

$$\text{Prob} \left\{ \left| \frac{\tau_n}{n} - 9 \right| \geq \varepsilon \mid Y_0 = r \right\} \leq C_\varepsilon^{-n}$$

Let  $S_n = \#$  of  $r$ 's  $- \#$  of  $l$ 's on  $[0, n]$ . We have that

$$\text{Prob}(E_n) = \text{Prob}\{S_0, S_1, \dots, S_{2n-2} > 0\}$$

and so we want to compute the asymptotics for  $\text{Prob}\{S_0, S_1, \dots, S_n > 0\}$ . Let  $\varepsilon > 0$ . We then have that

$$\begin{aligned} & \text{Prob}\{S_0, S_1, \dots, S_{\lceil 9m + m\varepsilon \rceil} > 0\} \\ &= \text{Prob}\{Y_0 = r\} \text{Prob}\{S_0, S_1, \dots, S_{\lceil 9m + m\varepsilon \rceil} > 0 \mid Y_0 = r\} \\ &\leq \frac{1}{9} \text{Prob}\{|\tau_m - 9m| \geq m\varepsilon \mid Y_0 = r\} \\ &\quad + \frac{1}{9} \text{Prob}\{S_0, S_{\tau_1-1}, \dots, S_{\tau_m-1} > 0 \mid Y_0 = r\} \\ &\leq \frac{1}{9} C_\varepsilon^{-m} + \frac{1}{9} \text{Prob}\{\bar{S}_1, \bar{S}_2, \dots, \bar{S}_m > 0\} \end{aligned}$$

Hence by (2.1)

$$\limsup_{m \rightarrow \infty} 9(3\pi m)^{1/2} \text{Prob}\{S_0, S_1, \dots, S_{\lceil 9m + m\varepsilon \rceil} > 0\} \leq 1 \tag{2.3}$$

Now going the other way around,

$$\begin{aligned} & \text{Prob}\{S_0, S_1, \dots, S_{\lfloor 9m - m\varepsilon \rfloor} > 0\} \\ &= \text{Prob}\{Y_0 = r\} \text{Prob}\{S_0, S_1, \dots, S_{\lfloor 9m - m\varepsilon \rfloor} > 0 \mid Y_0 = r\} \\ &\geq -\frac{1}{9} \text{Prob}\{|\tau_m - 9m| \geq m\varepsilon \mid Y_0 = r\} \\ &\quad + \frac{1}{9} \text{Prob}\{S_0, S_{\tau_1-1}, \dots, S_{\tau_m-1} > 0 \mid Y_0 = r\} \end{aligned}$$

since the events  $\{S_0, S_1, S_2, \dots, S_{\tau_m} > 0\}$  and  $\{S_0, S_{\tau_1-1}, \dots, S_{\tau_m-1} > 0\}$  are the same. (The  $S_i$ 's taken their lowest values one step before the  $\tau_m$ 's.) This last sum is in turn

$$\geq -\frac{1}{9} C_\varepsilon^{-m} - \frac{1}{9} \text{Prob}\{\bar{S}_1, \bar{S}_2, \dots, \bar{S}_m > 0\}$$

Hence

$$\liminf_{m \rightarrow \infty} 9(3\pi m)^{1/2} \text{Prob}\{S_0, S_1, \dots, S_{\lfloor 9m - m\varepsilon \rfloor} > 0\} \geq 1 \tag{2.4}$$

Since  $\varepsilon > 0$  is arbitrary and using the substitution  $n = 9m$ , Eq. (2.3) and (2.4) give

$$\text{Prob}\{S_0, S_1, \dots, S_n > 0\} \sim 1/\sqrt{27\pi n}$$

and hence

$$P(E_n) = \text{Prob}\{S_0, S_1, \dots, S_{2n-2} > 0\} \sim 1/\sqrt{54\pi n}$$

which demonstrates (2.2).

Let  $F_n$  denote the event that there is a 1 at lattice point 0 at time  $n$  that has come from lattice point  $n$ . Clearly  $\text{Prob}(E_n) = \text{Prob}(F_n)$  and  $\{\eta_n(0) = 1\} = E_n \cup F_n$  and we therefore have that  $\text{Prob}\{\eta_n(0) = 1\} = 2 \text{Prob}(E_n) - \text{Prob}(E_n \cap F_n)$ . If  $\text{Prob}(E_n \cap F_n)$  is of smaller order than  $n^{-1/2}$ , we would obtain  $\text{Prob}\{\eta_n(0) = 1\} \sim \sqrt{2/(27\pi n)}$ , as desired. In this regard, for  $n$  even, we let  $E'_n$  be the event that there is a 1 at lattice point  $-n/2$  at time  $n/2$  that has come from  $-n$ . So

$$E'_n = \{\eta_0(-n) = 1, \eta_1(-n+1) = 1, \eta_2(-n+2) = 1, \dots, \eta_{n/2}(-n/2) = 1\}$$

Analogously, we let  $F'_n$  be the event that there is a 1 at lattice point  $n/2$  at time  $n/2$  that has come from  $n$ . Since  $E_n \subseteq E'_n$ ,  $F_n \subseteq F'_n$ , and  $E'_n$  and  $F'_n$  are independent events, it follows that  $\text{Prob}(E_n \cap F_n) \leq (\text{Prob}(E'_n))^2$ . Now, the asymptotic formula  $\sqrt{n} \text{Prob}(E_n) \rightarrow a$  limit implies that  $\text{Prob}(E'_n) \leq c/n^{1/2}$  for some constant  $c$ . For  $n$  odd, a simple modification can be made.  $\square$

*Proof of Lemma 2.* Instead of computing the distribution of 1 - # of 01's between two successive 10s, we compute just the distribution of  $Z = \#$  of 01s between two successive 10s, which has the same variance. We will show that  $\text{Prob}(Z = n)$  is  $\frac{1}{4}$  if  $n = 0$  and  $\frac{9}{16}(\frac{1}{4})^{n-1}$  for  $n \geq 1$ . From this, it is straightforward to compute the mean and variance of  $F$  to be 0 and  $\frac{2}{3}$ , respectively.

Let  $p$  be the probability that the first 10 or 01 after a 10 is a 10, which is just  $\text{Prob}(Z = 0)$ . By conditioning on the lattice point after 10 one obtains

$$p = (\frac{1}{3})p + (\frac{1}{3})0 + (\frac{1}{3})\frac{1}{2}$$

which gives  $p = \frac{1}{4}$ . Next, since 10s and 01s are renewals, one obtains using an obvious symmetry that for  $n \geq 1$

$$\text{Prob}(Z = n) = (1 - p)^2 p^{n-1} = \frac{9}{16}(\frac{1}{4})^{n-1} \quad \square$$



### 3. TOPOLOGICAL ENTROPY FOR $d = 1$

In this section, we compute the topological entropy of the one-dimensional Greenberg–Hastings Model. We have a space  $X = \{0, 1, 2\}^{\mathbb{Z}}$  and a transformation  $T$  defined on it. If  $(M, S)$  is any pair consisting of a compact metric space  $M$  together with a continuous transformation  $S$  from  $M$  to  $M$ , there is a notion of the *topological entropy* of such a system, which is defined in Ref. 11. We will not give this general definition here but mention that, for the case  $(X, T)$ , it coincides with that given in the introduction. We denote this number by  $h(X, T)$ .

If  $F$  is a finite set,  $S$  a continuous mapping from  $F^{\mathbb{Z}}$  to itself, and  $W$  a closed  $S$ -invariant subset of  $F^{\mathbb{Z}}$ , then we can define the topological entropy of such a system in a way which is completely analogous to our definition of the topological entropy for the Greenberg–Hastings Model. We denote this number by  $h(W, S)$ .

As discussed in the introduction, if we consider the closed  $T$ -invariant subset  $L$  of  $X$  consisting of all configurations where every 1 has a 2 to its right, every 2 has a 0 to its right and every 0 has a 0 or a 1 to its right, then it is clear that  $T$  acting on this subset is exactly the same as shifting the configuration to the left, and hence the notation  $L$ .

A topological Markov shift is a system  $(W, S)$  defined by a 0, 1 matrix  $A = \{a_{i,j}\}_{i,j=1}^k$  as follows.  $W$  is the subset of  $\{1, \dots, k\}^{\mathbb{Z}}$  consisting of sequences  $\{x_n\}_{n \in \mathbb{Z}}$  with the property that  $a_{x_n, x_{n+1}} = 1$  for all  $n$  and  $S$  is simply the transformation that shifts a sequence one unit to the left. By Theorem 7.13 of Ref. 11, the entropy of any irreducible topological Markov shift is  $\ln(\lambda)$  where  $\lambda$  is the largest positive eigenvalue of the corresponding matrix. This theorem follows fairly easily from the spectral radius formula.

Now,  $L$  above is exactly the set of configurations which are associated to the topological Markov shift

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Since  $T$  acting on  $L$  just shifts a configuration to the left, the above theorem gives  $h(L, T) = \ln(\rho)$  where  $\rho$  is the positive root of  $x^3 - x^2 - 1$  and in particular  $h(X, T) \geq \ln \rho$ .

*Proof of Theorem 2.* One first notes that if  $\sigma \in \{0, 1, 2\}^{\mathbb{Z}^2}$  is a possible evolution of the system, then each row of  $\sigma$  is in the eventual image of  $T$ ,  $\bigcap_{n=1}^{\infty} T^n(X) \equiv Y$ . Hence we have  $h(X, T) = h(Y, T)$ . The following lemma is easy to verify and left to the reader.

**Lemma 3.**  $\eta \in X$  is in the eventual image  $Y$  of  $T$  if and only if there exists  $n \in \mathbf{Z} \cup \{-\infty, \infty\}$ , such that:

at  $n$  and to the left on  $n$ ,

- every 1 has a 2 to its left
- every 2 has a 0 to its left
- and every 0 has a 0 or 1 to its left

and to the right of  $n$ ,

- every 1 has a 2 to its right
- every 2 has a 0 to its right
- and every 0 has a 0 or a to its right

Note that the configurations in  $L$  define above are those with  $n = -\infty$  in this lemma. Now, let  $Y'$  be the closed subset of  $\{0, r, l\}^{\mathbf{Z}}$  consisting of all configurations for which there exists an  $n \in \mathbf{Z} \cup \{-\infty, \infty\}$  such that at  $n$  and to the left of  $n$ , there are only  $r$ 's and 0s with at least two 0s between  $r$ 's and to the right of  $n$ , there are only  $l$ 's and 0s with at least two 0s between  $l$ 's. Let  $T'$  map  $\{0, r, l\}^{\mathbf{Z}}$  to itself by having the  $r$ 's move right, the  $l$ 's move left, and an  $r$  and an  $l$  annihilate each other when they meet or cross. It is easy to see that  $Y'$  is invariant under  $T'$ . There is a natural map  $U$  from  $Y$  to  $Y'$  as follows. Let

$$U(\eta)(x) = \begin{cases} r & \text{if } \eta(x) = 2 \text{ and } \eta(x + 1) = 1 \text{ or } 2 \\ l & \text{if } \eta(x) = 2 \text{ and } \eta(x - 1) = 1 \text{ or } 2 \\ 0 & \text{otherwise} \end{cases}$$

This is a slight modification of what we did in Sec. 2. Since  $UT = T'U$  and  $U$  is surjective,  $h(Y', T') \leq h(Y, T)$  by Theorem 7.2 in Ref. 11. On the other hand using Theorem 17 of Ref. 2, one can show that  $h(Y', T') = h(Y, T)$ . Hence we need only consider the simpler system  $(Y', T')$ .

**Definition 1.** Let  $B_{n,m}$  be the collection of  $r, l, 0$ -valued configurations  $\sigma$  on  $[-m, m] \times [0, n - 1]$  that can be extended to a possible evolution of the dynamics for  $T'$ , that is, such that there exists  $\tilde{\sigma} \in \{r, l, 0\}^{\mathbf{Z}^2}$  with  $\tilde{\sigma} = \sigma$  on  $[-m, m] \times [0, n - 1]$  and with  $T'$  ( $i$ th row of  $\tilde{\sigma}$ ) =  $i + 1$ st row of  $\tilde{\sigma}$ .

Our problem has been reduced to computing  $\sup_m \limsup_{n \rightarrow \infty} (\ln b_{n,m})/n$  where  $b_{n,m} = |B_{n,m}|$ . The fact that  $\sup_m \limsup_{n \rightarrow \infty} (\ln b_{n,m})/n = 2 \ln \rho$  is contained in Proposition 1 and 2 where the two inequalities are proven. □

Let  $\gamma_{n,g}$  be the number of ways of putting down  $g$   $l$ 's on  $[0, n - 1]$  with the  $l$ 's separated by at least two 0s and let  $\gamma_n = \sum_{g=0}^n \gamma_{n,g}$ . By the discussion given earlier in this section, we have that

$$\limsup_{n \rightarrow \infty} \frac{\ln \gamma_n}{n} = \ln \rho$$

**Lemma 4.**

$$\limsup_{n \rightarrow \infty} \frac{\ln \sum_{g=0}^n \gamma_{n,g}^2}{n} = 2 \ln \rho$$

*Proof.* Since

$$\sum_{g=0}^n \gamma_{n,g}^2 \leq \left( \sum_{g=0}^n \gamma_{n,g} \right)^2 \leq (n + 1) \sum_{g=0}^n \gamma_{n,g}^2$$

it follows that

$$\limsup_{n \rightarrow \infty} \frac{\ln \sum_{g=0}^n \gamma_{n,g}^2}{n} = \limsup_{n \rightarrow \infty} \frac{\ln (\sum_{g=0}^n \gamma_{n,g})^2}{n}$$

which is  $2 \ln \rho$  by above, recalling that  $\sum_{g=0}^n \gamma_{n,g} = \gamma_n$ . □

**Proposition 1.**  $h(Y', T') \leq 2 \ln \rho$ .

*Proof.* Fix  $m$ . We shall show that  $\limsup_{n \rightarrow \infty} (\ln b_{n,m})/n \leq 2 \ln \rho$ . Let  $C_{n,m}$  be the set of configurations  $\sigma$  in  $\{0, r, l\}^{[-m-n+1, m+n-1]}$  such that  $\sigma$  extends to a configuration in  $Y'$  and such that there are no  $l$ 's to left of  $-m$  and no  $r$ 's to the right of  $m$ .

We shall first show that there is a surjection of  $C_{n,m}$  onto  $B_{n,m}$  and then we will obtain a bound on  $|C_{n,m}|$ . There is an obvious map  $f$  from  $C_{n,m}$  to  $B_{n,m}$  as follows. If  $\sigma \in C_{n,m}$ , then we evolve  $\sigma$  up to time  $n - 1$  and restrict to  $[-m, m] \times [0, n - 1]$ . The fact that  $\sigma$  is defined on  $[-m - n + 1, m + n - 1]$  implies that the evolution of  $\sigma$  in  $[-m, m]$  is defined up to time  $n - 1$ . Next given  $\hat{\sigma} \in B_{n,m}$ , by definition there is some  $\sigma \in \{0, r, l\}^{[-m-n+1, m+n-1]}$  which extends to an element of  $Y'$  and which has the property that when evolved and restricted to  $[-m, m] \times [0, n - 1]$ ,  $\hat{\sigma}$  is obtained. We can now modify  $\sigma$  so that it lies in  $C_{n,m}$  and such that its evolution restricted to  $[-m, m] \times [0, n = 1]$  is still  $\hat{\sigma}$ . If in  $\sigma$  there are  $x$   $l$ 's to the left of  $-m$ , then remove these  $l$ 's together with the  $x$  rightmost  $r$ 's or whatever number of  $r$ 's there are if there are

fewer than  $x$ . In a similar fashion, remove all  $r$ 's from  $\sigma$  that are to the right of  $m$ . It is clear that this modification lies in  $C_{n,m}$  and that its evolution in  $[-m, m] \times [0, n-1]$  is not changed. Hence  $f$  is surjective. Finally it is clear that

$$|C_{n,m}| \leq (\gamma_{n+2m})^2$$

and hence

$$\limsup_{n \rightarrow \infty} (\ln b_{n,m})/n \leq \limsup_{n \rightarrow \infty} (\ln \gamma_{n+2m}^2)/n = 2 \ln \rho \quad \square$$

**Proposition 2.**  $h(Y', T') \geq 2 \ln \rho$ .

*Proof.* Let  $\varepsilon > 0$ . Using Lemma 4 choose  $m$  such that

$$\frac{\ln(\sum_{g=0}^m \gamma_{m,g}^2)}{m+2} \geq 2 \ln \rho - \varepsilon$$

Let  $n = t(m+2)$  with  $t$  a positive integer. Then  $b_{n,m}$  is at least the number of ways of putting down 0s,  $r$ 's, and  $l$ 's on  $[-n, n]$  such that there are 0s at 0,  $h(m+2) + m + 1$ ,  $h(m+2) + m + 2$ ,  $-h(m+2) - m - 1$ , and  $-h(m+2) - m - 2$  for  $h=0, \dots, t-1$ , there are no  $l$ 's to the left of 0, no  $r$ 's to the right of 0, both the  $l$ 's and  $r$ 's are separated by at least two 0s, and the number of  $r$ 's in  $[-h(m+2) - m, -h(m+2) - 1]$  is the same as the number of  $l$ 's in  $[h(m+2) + 1, h(m+2) + m]$  for  $h=0, \dots, t-1$ . This is because different such configurations when evolved differ at some point in  $[-m, m] \times [0, n-1]$  as long as  $m \geq 1$ . The latter number is given by  $(\sum_{g=0}^m \gamma_{m,g}^2)^t$ . Hence

$$\frac{\ln b_{n,m}}{n} \geq \frac{\ln(\sum_{g=0}^m \gamma_{m,g}^2)^t}{n} \geq 2 \ln \rho - \varepsilon$$

and so  $h(Y', T') \geq \limsup_{n \rightarrow \infty} (\ln b_{n,m})/n \geq 2 \ln \rho - \varepsilon$ . □

#### 4. THE LIMIT DISTRIBUTIONS FOR $d \geq 2$ ARE BERNOULLI SHIFTS

Let  $\mu$  denote a fixed translation invariant product measure on  $X = \{0, 1, 2\}^{\mathbb{Z}^d}$  in which each state has positive probability. Restricting ourselves to  $d \geq 2$ , we demonstrate in this section the existence of a non-trivial limit measure when  $\mu$  is the initial distribution and show that this measure has the strongest possible ergodic behavior with respect to spatial translations.

We now give the proof of Lemma 1 which follows exactly the proof given in Ref. 5 for the cyclic cellular automaton.

*Proof of Lemma 1.* Let  $\eta_0$  contain the clock  $x_0, x_1, \dots, x_{n-1}$  as defined in the introduction. Clearly each of these  $x_i$  cycles around at period 3. We show all lattice points eventually cycle at period 3. If not, there is a closest point  $y$  to the clock which does not eventually cycle at period 3. Since it is closest, it has a neighbor  $x$  which does eventually cycle at period 3. If  $y$  does not cycle at period 3, then at some point in time,  $x$  will be a 1 and  $y$  will be a 0. At this point, it is clear that  $y$  will be “caught by  $x$ ” and then also begin to cycle at period 3.  $\square$

We let  $T^k\mu$  denote the distribution on  $X$  at time  $k$  when  $\mu$  is the original distributions. Since under  $\mu$  a clock will exist a.s., Lemma 1 gives the following:

**Theorem 6.**  $\lim_{n \rightarrow \infty} T^{3n}\mu$  exists.

Letting  $\nu$  denote the above limit measure, we clearly have  $T^3\nu = \nu$  and  $\nu$  is translation-invariant. Theorem 3 states that  $\nu$  is a Bernoulli shift. Being a Bernoulli shift is the strongest ergodic theoretic property a measure can have. In particular, it implies mixing of all orders and ergodicity (see Ref. 11).

Before proceeding with the proof of Theorem 3, we define a metric on translation-invariant measures on  $X$ . It is called the  $\bar{d}$ -metric and has been discussed in many places (see, e.g., Ref. 7).

**Definition 2.** If  $m_1$  and  $m_2$  are translation-invariant measures on  $X$ , then

$$\bar{d}(m_1, m_2) = \inf \text{Prob}^m \{ \eta(0) \neq \delta(0) \}$$

where the infimum is taken over all translation invariant couplings  $m$  of  $m_1$  and  $m_2$  and  $\text{Prob}^m$  denotes probability with respect to the measure  $m$ .

*Proof of Theorem 3.* First, factors of Bernoulli shifts are Bernoulli shifts (see Ref. 7); that is, if  $(X, \mathcal{B}, m_1)$  is a Bernoulli shift and

$$f: (X, \mathcal{B}, m_1) \rightarrow (X, \mathcal{B}, m_2)$$

is measurable, measure-preserving (but not necessarily invertible), and commutes with spacial translations then  $(X, \mathcal{B}, m_2)$  is a Bernoulli shift. Hence for all  $n$ ,  $T^{3n}\mu$  is a Bernoulli shift. Next, a  $\bar{d}$ -limit of Bernoulli shifts is a Bernoulli Shift (see Ref. 7). Hence, to prove the theorem, we need only show that the weak limit  $\nu$  is actually a  $\bar{d}$ -limit. This however is very easy.

There is a canonical coupling of  $T^{3n}\mu$  and  $\nu$ . Pick a configuration  $\eta$  from  $X$  according to  $\mu$ . Let  $\delta = T^{3n}\eta$  and let  $\gamma = \lim_{n \rightarrow \infty} T^{3n}\eta$ , which we know exists from Lemma 1. Clearly the distributions for  $\delta$  and  $\gamma$  are  $T^{3n}\mu$  and  $\nu$ , respectively. Moreover, Lemma 1 also implies that these couplings demonstrate the  $\bar{d}$ -convergence. □

**5. CRITICAL PROBABILITIES FOR  $d \geq 2$**

Theorem 4 demonstrates the fact that there are three distinct regimes where distinct behavior occurs. The transition points are  $\delta_L = c/(L^{d/4})$  and  $\delta_L = c/(L^{d/2})$ , where  $c$  is any constant. Above  $c/(L^{d/4})$ , the system eventually enters a periodic 3 state. Below  $c/(L^{d/2})$ , the system eventually dies out (enters the only fixed point of all 0s.) Above  $c/(L^{d/2})$  but below  $c/(L^{d/4})$ , the system does not die out nor enter a periodic 3 state.

We will prove Theorem 4 by establishing a number of lemmas. The first two, Lemmas 5 and 6, are purely deterministic statements concerning how one determines the eventual fate for a certain large collection of configurations. Lemmas 7 and 8 establish a Poisson approximation for dependent events using the Chen–Stein method. The remaining lemmas allow us to put together a proof of Theorem 4. First, clocks are defined exactly as they were in the previous section. The only difference is that while clocks on the infinite lattice must have length a multiple of 6, the length need only be a multiple of 3 on the finite lattice because of the periodic boundaries. We give another relevant definition here. As in Ref. 5, we call a set of distinct points  $x_0, x_1, \dots, x_{n-1}, x_n = x_0 \in B_L$  a *defect* for  $\eta$  if  $x_{i+1}$  is a neighbor of  $x_i$  for each  $i$  and

$$\frac{1}{3} \sum_{k=0}^{n-1} \{ \eta(x_{k+1}) - \eta(x_k) \} \neq 0$$

where the summands are all chosen (mod 3) to be  $-1, 0$ , or  $1$  and the sum is ordinary addition, not (mod 3). Intuitively, if we view 0, 1, and 2 as sitting on a circle, then as we go along the points  $x_0, x_1, \dots, x_{n-1}$  and return to  $x_0$ , the corresponding path on the circle has nonzero index. For easy reference, we put together a list of definitions, some of which are already given.

**Definitions 3.** We have the following:

Let  $E_3$  denote the event that a periodic 3 state is eventually entered.

Let  $D$  denote the event that the system eventually dies out.

Let  $F = (D \cup E_3)^c$ .

Let  $C_k$  denote the event that there is a clock of length  $3k$  initially.

Let  $C = \bigcup_{i=1}^{\infty} C_i$ .

Let  $R$  denote the event that there is a  $2 \times 2$  box with a defect initially.

Let  $G$  denote the event that there is an adjacent 12 pair initially.

**Lemma 5.** A configuration eventually dies out if and only if it does not contain a defect.

*Proof.* We first show that for any  $\eta$  and any distinct  $x_0, x_1, \dots, x_{n-1}, x_n = x_0$  with  $x_i$  and  $x_{i+1}$  neighbors for  $i = 0, \dots, n-1$ ,

$$\frac{1}{3} \sum_{k=0}^{n-1} \{\eta(x_{k+1}) - \eta(x_k)\}$$

is invariant under the dynamics; that is,

$$\frac{1}{3} \sum_{k=0}^{n-1} \{T\eta(x_{k+1}) - T\eta(x_k)\} = \frac{1}{3} \sum_{k=0}^{n-1} \{\eta(x_{k+1}) - \eta(x_k)\} \quad (5.1)$$

Here, as in the definition of a defect, the summands are all chosen (mod 3) to be  $-1, 0$ , or  $1$  and the sum is ordinary addition, not (mod 3). One proves this as follows. Let  $S$  be the subset of  $\{x_0, x_1, \dots, x_{n-1}\}$  consisting of those points  $x$  with  $T\eta(x) = \eta(x)$ . Clearly it must be the case that  $\eta(x) = 0$  for  $x \in S$ . Now, it is easy to check by cases that

$$\begin{aligned} & (T\eta(x_{k+1}) - T\eta(x_k)) - (\eta(x_{k+1}) - \eta(x_k)) \\ &= \begin{cases} 0 & \text{if } x_k, x_{k+1} \in S \text{ or } x_k, x_{k+1} \notin S \\ 1 & \text{if } x_k \in S, x_{k+1} \notin S \\ -1 & \text{if } x_{k+1} \in S, x_k \notin S \end{cases} \end{aligned}$$

Since the number of pairs  $(x_k, x_{k+1})$  with  $x_k \in S$  and  $x_{k+1} \notin S$  is the same as the number of pairs with  $x_k \notin S$  and  $x_{k+1} \in S$ , (5.1) follows. We therefore obtain the fact that  $\eta$  has a defect if and only if  $T\eta$  has a defect. It follows that if  $\eta$  eventually dies out, it cannot contain a defect.

For the converse, we need to show that a periodic configuration  $\eta$  which is not all 0s has a defect. In this regard, one first notes that every 1 must have a neighbor which is a 2 (since  $\eta = T\delta$  for some  $\delta$ ), and every 2 must have a neighbor which is a 0 (since  $\eta = T^2\delta$  for some  $\delta$ ). At the end, we show that every 0 can reach a 1 by moving through 0s. Assuming this, one constructs a defect as follows. Let  $x_0$  be such that  $\eta(x_0) = 1$ . Let  $x_1$  be a neighbor of  $x_0$  such that  $\eta(x_1) = 2$ . Let  $x_2$  be a neighbor of  $x_1$  such that  $\eta(x_2) = 0$ . From  $x_2$ , jump along neighbors which are 0 until a 1 is reached.

Continue the process until a 1 is reached twice using the fact that  $B_L$  is finite. Removing an initial segment and other intermediate segments if necessary, one obtains a defect.

Finally, to clean up the last point, we show that if  $\eta$  is periodic, then there cannot be a set  $A$  such that  $\eta$  is 0 on  $A$  and 2 on  $\partial A$  where  $\partial A$  is the set of points which are not in  $A$  but are adjacent to some point in  $A$ . Letting  $A$  be the set of 0s which can be reached by moving only through 0s from some fixed 0, we have as a consequence of this fact that for a periodic configuration which is not all 0s, every 0 can reach a 1 by moving only through 0s.

Choose  $\delta$  so that  $\delta$  is periodic and  $T\delta = \eta$ .  $\delta$  must be 1 on  $\partial A$ , 2 on those points in  $A$  which are adjacent to some point in  $\partial A$ , and 0 or 2 elsewhere in  $A$ . Since every 2 is adjacent to a 0 in a periodic configuration,  $\delta$  must be 0 somewhere in  $A$ . Letting  $B$  be those points in  $A$  where  $\delta$  is 0, we have that  $\delta$  is 0 on  $B$  and 2 on  $\partial B$  with  $|B| < |A|$ . Iterating this, we finally obtain a periodic configuration which has a 0 all of whose neighbors are 2. But such a configuration cannot be  $T^{3\gamma}$  for any  $\gamma$ . □

**Lemma 6.** A configuration eventually reaches a periodic 3 state if and only if it contains a clock initially. If a configuration initially contains a  $2 \times 2$  box with a defect but not a clock, then the system does not die out nor reach a periodic 3 state. If there are no adjacent 12s in a configuration initially, then it dies out. In the above notation,  $E_3 = C$ ,  $C^c \cap R \subseteq F$ , and  $G^c \subseteq D$ .

*Proof.* First, if there is a clock, then the system eventually reaches a periodic 3 state. This is just a finite version of Lemma 1. Next, let  $\eta$  be a periodic 3 configuration. We claim that  $\eta$  has a clock. One first needs to observe that any 0 must have a 1 next to it (since  $\eta$  has period 3). As in the proof of Lemma 5 any 1 in  $\eta$  must have a 2 next to it and any 2 in  $\eta$  must have a 0 next to it. Now letting  $x_1$  be an arbitrary point in  $B_L$ , let  $x_2$  be a neighbor of  $x_1$  such that  $\eta(x_2) \equiv \eta(x_1) + 1 \pmod{3}$ . Similarly, let  $x_3$  be a neighbor of  $x_2$  such that  $\eta(x_3) \equiv \eta(x_2) + 1 \pmod{3}$ . Continuing in this way, we get a sequence  $x_1, x_2, \dots$ . Since  $B_L$  is finite, we can stop the first time a lattice point is repeated. If  $x_j = x_{j+k}$  is the repeated site, then  $k$  is a multiple of 3 and moreover some rotation of  $\{x_j, x_{j+1}, \dots, x_{j+k-1}\}$  is a clock for  $\eta$ . One next shows that if  $T\eta$  has a clock, then  $\eta$  must have a clock at the same lattice points. Let  $x_0, x_1, \dots, x_{n-1}$  be a clock for  $T\eta$ . We claim that  $x_1, \dots, x_{n-1}, x_0$  is a clock for  $\eta$ . If  $T\eta(x_i)$  is 1 or 2, then clearly  $\eta(x_i)$  must be 0 or 1, respectively. If  $T\eta(x_i) = 0$ , then  $T\eta(x_{i-1}) = 2$  and so  $\eta(x_{i-1}) = 1$ . Therefore if  $\eta(x_i)$  were 0, then  $T\eta(x_i)$  would be a and hence it must be the case that  $\eta(x_i)$  is 2. Hence  $x_1, \dots, x_{n-1}, x_0$  is a clock for  $\eta$ .



Hence clocks cannot be created unless they are there initially and so if a configuration eventually reaches a periodic 3 state, it must have had a clock initially.

The second statement follows from the first statement and Lemma 5. Finally, if  $\eta$  has no adjacent 12s pairs, it is easy to see that  $\eta$  has no defects and hence dies out by Lemma 5.  $\square$

The next lemma gives a Poisson approximation for the number of 1, 2 pairs in the initial configuration.

**Lemma 7.** Fix  $\delta_L \in (0, 1)$  and let  $L \geq 3$ . Let the random variable  $X_L$  denote the number of adjacent 1, 2s in  $B_L$  where the distribution on  $\{0, 1, 2\}^{B_L}$  is product measure with 1 and 2s each having density  $\delta_L$  and 0s having density  $1 - 2\delta_L$ . Then for some constant  $K$  (depending only on  $d$ ),  $\|X_L - Y\| \leq KL^d \delta_L^3$  where  $Y$  has a Poisson distribution with mean  $2dL^d \delta_L^2$  and  $\|\cdot\|$  denotes the total variation norm given by  $\|W - Z\| = 2 \sup_{A \subseteq \mathbb{Z}} |P(W \in A) - P(Z \in A)|$ .

*Proof.* Let  $I$  be the set of pairs of adjacent points in  $B_L$ . Clearly,  $|I| = dL^d$  if  $L \geq 3$ . For  $i \in I$ , let  $S_i$  be the subset of  $I$  consisting of pairs of adjacent points which have at least one point in common with  $i$  and let  $C = |S_i|$ , which is independent of  $i$ . For  $i \in I$ , let  $A_i$  be the event  $\{i \text{ has both a 1 and 2 initially}\}$ . Clearly,  $\text{Prob}_L(A_i) = 2\delta_L^2$ .

It is easy to check that

$$b_1 \equiv \sum_{i \in I} \sum_{j \in S_i} \text{Prob}_L(A_i) \text{Prob}_L(A_j) = 4dCL^d \delta_L^4$$

and that

$$b_2 \equiv \sum_{i \in I} \sum_{j \neq i \in S_i} \text{Prob}_L(A_i \cap A_j) = 2d(C - 1)L^d \delta_L^3 \leq 2dCL^d \delta_L^3$$

We clearly also have that  $A_i$  and  $\sigma(A_j: j \in S_i^c)$  are independent. Since  $X_L = \sum_{i \in I} I_{A_i}$  and  $\sum_{i \in I} \text{Prob}_L(A_i) = 2dL^d \delta_L^2$ , it follows from Theorem 1 in Ref. 1 that

$$\|X_L - Y\| \leq 2(4dCL^d \delta_L^4 + 2dCL^d \delta_L^3) \leq 12dCL^d \delta_L^3$$

where  $Y$  is Poisson with mean  $2dL^d \delta_L^2$ . Now let  $K = 12dC$ .  $\square$

We next want to obtain a Poisson approximation for the number of clocks of length 6. Any such clock either sits inside a  $2 \times 3$  rectangle or inside a  $2 \times 2 \times 2$  cube if  $d > 2$ . There is however a problem in applying the Chen–Stein method as before because the degree of dependence between

different clocks can become too large. This is because it is possible that two clocks intersect in five places; for example, in  $d=3$ , there might be clocks at both  $((0, 0, 0), (1, 0, 0), (1, 1, 0), (1, 1, 1), (1, 0, 1), (0, 0, 1))$  and at  $((0, 0, 0), (1, 0, 0), (1, 1, 0), (1, 1, 1), (0, 1, 1), (0, 0, 1))$ , differing only in the fifth location. This problem does not occur for  $d=2$ . The way to get around this is to count the number of  $2 \times 3$  rectangles and the number of  $2 \times 2 \times 2$  cubes which contain a clock rather than counting the total number of clocks. In this way, if we have

$$\{\eta(0, 0, 0) = 1, \eta(1, 0, 0) = 0, \eta(0, 1, 0) = 0, \eta(1, 1, 0) = 2, \\ \eta(0, 0, 1) = 2, \eta(1, 0, 1) = 0, \eta(0, 1, 1) = 0, \eta(1, 1, 1) = 1\}$$

we count this only once even though it actually contains 4 clocks.

**Lemma 8.** Fix  $\delta_L \in (0, 1)$  and let  $L \geq 4$ . Let the random variable  $X_L$  denote the number of  $2 \times 3$  rectangles plus the number of  $2 \times 2 \times 2$  cubes which contain a clock of length 6 where the distribution on  $\{0, 1, 2\}^{B_L}$  is product measure with 1 and 2s each having density  $\delta_L$  and 0s having density  $1 - 2\delta_L$ . Then for some constant  $K'$  (depending only on  $d$ ),  $\|X_L - Y\| \leq K' L^d \delta_L^5$ , where  $Y$  has a Poisson distribution with mean  $12 \binom{d}{2} L^d \delta_L^4 (1 - 2\delta_L)^2 + \binom{d}{3} L^d z_L$ , where

$$6\delta_L^4 (1 - 2\delta_L)^2 \leq z_L \leq (16) 6\delta_L^4 (1 - 2\delta_L)^2$$

and  $\lim_{\delta_L \rightarrow 0} z_L / \delta_L^4 = 60$ .

*Proof.* Let  $I$  be the set of  $2 \times 3$  rectangles and  $2 \times 2 \times 2$  cubes in  $B_L$ . We leave to the reader to show that  $|I| = (2 \binom{d}{2} + \binom{d}{3}) L^d$  (the second term being 0 by convention if  $d=2$ ) since there are  $2 \binom{d}{2} L^d$  rectangles and  $\binom{d}{3} L^d$  cubes. Here one uses the fact that  $L \geq 4$ . Let  $\alpha(d)$  denote  $2 \binom{d}{2} + \binom{d}{3}$ . For  $i \in I$ , let  $S_i$  be the subset of  $I$  consisting of sets which have at least one lattice point in common with  $i$ .  $|S_i|$  can only take to two possible values depending on whether  $i$  is a rectangle or a cube, and we let  $C'$  be the larger. For  $i \in I$ , let  $A_{i\#}$  be the event  $\{i \text{ has a clock initially}\}$ . Since there are six ways to put a clock in a rectangle,  $\text{Prob}_L(A_i) = 6\delta_L^4 (1 - 2\delta_L)^2$  when  $i$  is a rectangle. Let  $z_L = \text{Prob}_L(A_i)$  when  $i$  is a cube. Since there are 16 possible subsets of a cube where a clock can sit (a fact easily checked), we have that  $6\delta_L^4 (1 - 2\delta_L)^2 \leq z_L \leq (16) 6\delta_L^4 (1 - 2\delta_L)^2$ . Hence

$$b_1 \equiv \sum_{i \in I} \sum_{j \in S_i} \text{Prob}_L(A_i) \text{Prob}_L(A_j) \leq (16)^2 36\alpha(d) C' L^d \delta_L^8$$

The computation of

$$b_2 \equiv \sum_{i \in I} \sum_{j \neq i \in S_i} \text{Prob}_L(A_i \cap A_j)$$

is as follows.

One can see geometrically that for  $j \neq i \in S_i$ ,

$$A_i \cap A_j \subseteq A_i \cap \{\text{there is a 1 or 2 in } j \cap i^c\}$$

Hence

$$\text{Prob}_L(A_i \cap A_j) \leq (16) 6\delta_L^4(1 - 2\delta_L)^2 10\delta_L$$

by independence and so

$$b_2 \leq 960\alpha(d) C' L^d \delta_L^5$$

We clearly also have that  $A_i$  and  $\sigma(A_j; j \in S_i^c)$  are independent. Since  $X_L = \sum_{i \in I} I_{A_i}$  and

$$\sum_{i \in L} \text{Prob}_L(A_i) = 12 \binom{d}{2} L^d \delta_L^4 (1 - 2\delta_L)^2 + \binom{d}{3} L^d z_L$$

it follows again from Theorem 1 in Ref. 1 that

$$\|X_L - Y\| \leq 2[(16)^2 36\alpha(d) C' L^d \delta_L^8 + 960\alpha(d) C' L^d \delta_L^5] \leq 20, 352\alpha(d) C' L^d \delta_L^5$$

where  $Y$  is Poisson with mean  $12\binom{d}{2} L^d \delta_L^4 (1 - 2\delta_L)^2 + \binom{d}{3} L^d z_L$ . Now let  $K' = 20, 352\alpha(d) C'$ . Last, we need to show that  $\lim_{L \rightarrow \infty} z_L / \delta_L^4 = 60$ . Neglecting terms of the form  $(1 - 2\delta_L)$  which approach 1, it is clear that  $z_L$  is a polynomial in  $\delta_L$  with terms involving only  $\delta_L^4$ ,  $\delta_L^5$  and  $\delta_L^6$ , and hence this limit is just the coefficient of  $\delta_L^4$ . Now, this coefficient is the number of ways of assigning 0s, 1s, and 2s to the vertices of a cube such that there are four 0s, two 1s, and two 2s and such that there is a clock. We claim there are 60 such ways.

First, the two 1s must be either adjacent or on opposite corners in order to have a clock. There are 12 ways to have them adjacent as there are 12 edges. Then for one of these 1s, it must have a 0 and 2 as its other two neighbors. Each of these extend uniquely to an allowable assignment, which gives 24. Next, there are four ways to put the 1s on opposite corners. Then for one of these 1s, it must have two 0s and a 2 as neighbors, giving 12 ways. Finally, any way of filling in the final three places with a 2 and two 0s gives an allowable assignment, giving 36 of this type and hence 60 altogether.  $\square$

**Lemma 9.** Of  $\lim_{L \rightarrow \infty} \delta_L = 0$  and  $\lim_{L \rightarrow \infty} \text{Prob}_L(G) < \gamma$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(R) = \gamma$  where  $G$  and  $R$  are defined in Definitions 3.

*Proof.* Straightforward. □

**Lemma 10.** If  $\lim_{L \rightarrow \infty} \delta_L = 0$  and  $\lim_{L \rightarrow \infty} \delta_L L^{d/2} = \infty$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(R) = 1$ .

*Proof.* Applying Lemma 9, it suffices to show that if  $\lim_{L \rightarrow \infty} \delta_L = 0$  and  $\lim_{L \rightarrow \infty} \delta_L L^{d/2} = \infty$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(G) = 1$ . If  $\lim_{L \rightarrow \infty} \delta_L^3 L^d = 0$ , then this is immediate from Lemma 7. If the  $\delta_L$ 's are larger such that  $\limsup_{L \rightarrow \infty} \delta_L^3 L^d > 0$ , it is even easier to have an adjacent 12 pair and  $\lim_{L \rightarrow \infty} \text{Prob}_L(G) = 1$  is proved in this case by a simple coupling argument using the first case. □

**Lemma 11.** If  $\lim_{L \rightarrow \infty} \delta_L = 0$  and  $\lim_{L \rightarrow \infty} \delta_L L^{d/4} = \infty$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(E_3) = 1$ .

*Proof.* If  $\lim_{L \rightarrow \infty} \delta_L^5 L^d = 0$ , then using Lemmas 6 and 7, we have

$$\text{Prob}_L(E_3) \geq \text{Prob}_L(C_2) \text{ which } \rightarrow 1 \text{ as } L \rightarrow \infty$$

where we recall that  $C_2$  is the event that there is clock of length 6 initially. If  $\delta_L^5 L^d$  does not tend to 0, then, as in Lemma 10, it is obvious that it is even easier to have 1s and 2s in the desired locations and so, in this case, we also have  $\lim_{L \rightarrow \infty} \text{Prob}_L(E_3) = 1$ . □

**Lemma 12.** If  $\lim_{L \rightarrow \infty} \delta_L L^{d/4} = \lambda \in (0, \infty)$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(E_3) = 1 - e^{-\beta(d)\lambda^4}$  and  $\lim_{L \rightarrow \infty} \text{Prob}_L(F) = e^{-\beta(d)\lambda^4}$  where  $\beta(d)$  is defined in Theorem 4.

*Proof.* We first note that when the length  $L$  of  $B_L$  is at least 4, then there cannot be a clock of length 3; that is,  $C_1$  is impossible. Using Lemmas 6 and 8, we have

$$\text{Prob}_L(E_3) \geq \text{Prob}_L(C_2) \text{ which } \rightarrow 1 - e^{-\beta(d)\lambda^4} \text{ as } L \rightarrow \infty$$

and

$$\text{Prob}_L(E_3) \leq \text{Prob}_L(C_2) + \text{Prob} \left( \bigcup_{i=3}^{\infty} C_i \right)$$

Hence, to prove the first claim, we need only show that

$$\lim_{L \rightarrow \infty} \text{Prob}_L \left( \bigcup_{i=3}^{\infty} C_i \right) = 0$$

which follows from

$$\begin{aligned} \sum_{k=3}^{\infty} \text{Prob}_L(C_k) &\leq \sum_{k=3}^{\infty} L^d (2d)^{3k} 6\delta_L^{2k} = 6L^d \sum_{k=3}^{\infty} ((2d)^3 \delta_L^2)^k \\ &= 6L^d ((2d)^3 \delta_L^2)^3 (1 - (2d)^3 \delta_L^2)^{-1} \rightarrow 0 \quad \text{as } L \rightarrow \infty \end{aligned}$$

The second claim follows immediately from Lemmas 6 and 10.  $\square$

**Lemma 13.** If  $\lim_{L \rightarrow \infty} \delta_L L^{d/2} = \infty$  but  $\lim_{L \rightarrow \infty} \delta_L L^{d/4} = 0$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(F) = 1$ .

*Proof.* First  $\lim_{L \rightarrow \infty} \delta_L L^{d/4} = 0$  implies that  $\lim_{L \rightarrow \infty} \text{Prob}_L(E_3) = 0$  by following the proof of Lemma 12. Then Lemmas 6 and 10 again finish the proof.  $\square$

**Lemma 14.** If  $\lim_{L \rightarrow \infty} \delta_L L^{d/2} = \lambda \in [0, \infty)$ , then  $\lim_{L \rightarrow \infty} \text{Prob}_L(F) = 1 - e^{-2d\lambda^2}$  and  $\lim_{L \rightarrow \infty} \text{Prob}_L(D) = e^{-2d\lambda^2}$ .

*Proof.* By Lemmas 6 and 7,

$$\text{Prob}_L(D) \geq \text{Prob}_L(G^c) \rightarrow e^{-2d\lambda^2} \quad \text{as } L \rightarrow \infty$$

On the other hand, by Lemmas 6 and 9,

$$\text{Prob}_L(F \cup E_3) \geq \text{Prob}_L(C \cup R) \geq \text{Prob}_L(R)$$

which  $\rightarrow 1 - e^{-2d\lambda^2}$  as  $L \rightarrow \infty$ . Since  $\lim_{L \rightarrow \infty} \text{Prob}_L(E_3) = 0$  by the proof of Lemma 12, the inequalities derived above prove the lemma.  $\square$

## ACKNOWLEDGMENTS

The first author was partially supported by the National Science Foundation, National Security Agency, and the Army Research Office through the Mathematical Sciences Institute at Cornell University. The second author was partially supported by the Army Research Office through the Mathematical Sciences Institute at Cornell University and a National Science Foundation Mathematical Sciences Postdoctoral Research Fellowship.

We thank the referee for a number of helpful suggestions and in particular for a simplification of the proof of Lemma 2.

## REFERENCES

1. Arratia, R., Goldstein, L., and Gordon, L. (1989). Two moments suffice for Poisson approximations: The Chen–Stein Method, *Ann. Prob.* **17**, 9–25.
2. Bowen, R. (1971). Entropy for group endomorphisms and homogeneous spaces, *AMS Trans.* **153**, 401–414.
3. Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II. Wiley, New York.
4. Fisch, R. Clustering in the one-dimensional 3-color cyclic cellular automaton. Submitted for publication.
5. Fisch, R., Gravner, J., and Griffeath, D. Cyclic cellular automata in two dimensions. In *Proceedings of the Conference in Honor of Ted Harris*. Birkhauser. To appear.
6. Greenberg, J. M., and Hastings, S. P. (1978). Spatial patterns for discrete models of diffusion in excitable media. *Siam. J. Appl. Math.* **34**, 515–523.
7. Ornstein, D. S. (1974). *Ergodic Theory, Randomness and Dynamical Systems*. Yale University Press, New Haven.
8. Smillie, J. *Properties of the Directional Entropy Function for Cellular Automata*. Lecture Notes in Mathematics, Vol. 1342, Springer-Verlag, Berlin, pp. 689–705.
9. Tamayo, P., and Hartman, H. (1988). Cellular automata, reaction-diffusion systems and the origin of life. In *Artificial Life*. Addison–Wesley, Reading, MA, pp. 105–124.
10. Toffoli, T., and Margolus, N. (1987). *Cellular Automata Machines*, MIT Press, Cambridge, MA.
11. Walters, Peter (1975). *An Introduction to Ergodic*. Springer–Verlag, New York.
12. Winfree, A. T., Winfree, E. M., and Seifert, H. (1985). Organizing centers in a cellular excitable medium, *Physica* **17D**, 109–115.