

## Local Structure Theory in More Than One Dimension

**Howard A. Gutowitz**

*Laboratory of Biophysics, The Rockefeller University,  
1230 York Avenue, New York, NY 10021-6399, USA*

**Jonathan D. Victor**

*Department of Neurology, Cornell University Medical Center  
and*

*Laboratory of Biophysics, The Rockefeller University,  
1300 York Avenue, New York, NY 10021-6399, USA*

**Abstract.** The local structure theory (Gutowitz et al. [5]) is an analytical method for the determination of the statistical properties of cellular automaton evolution. In this paper we develop the local structure theory for multi-dimensional Euclidean lattices. The local structure theory for the Game of Life [1] is studied in detail. Some statistical properties of this cellular automaton are well approximated by a low order theory.

### 1. Introduction

Much of the current interest in cellular automata is due to their potential for modeling physical and biological processes ([13, 14] and references therein). These models are typically formulated with cellular automata on lattices in more than one dimension. Much of the mathematical work on cellular automata, on the other hand, concerns one-dimensional cellular automata. In this paper we begin to bridge this gap by showing how the local structure theory (Gutowitz et al. [5]) may be used to study multi-dimensional cellular automata.

The local structure theory is a generalization of the mean-field theory [11, 13] for cellular automata. It is based on the assumption that the correlations generated by cellular automaton evolution decay with distance. Previously [5], we showed that this assumption leads to the derivation of formulae for the assignment of probabilities to blocks of states of all sizes on the one-dimensional lattice, given a consistent assignment of probability to blocks of a fixed finite size. These formulae allowed us to define a sequence of measures which approximate an arbitrary shift-invariant measure. We called these measures finite block measures.

In one dimension a finite block measure is equivalent to a Markov process with memory  $n$ , also known as an  $n$ -step Markov process (Gutowitz et al. [5], Brascamp [2], Fannes and Verbeure [4]). The Markov process formulation allows a great deal of analytical machinery to be applied to these measures. Hence their properties are relatively well known. With more than one spatial dimension, it is not natural to associate spatial extension of probabilities with iteration of a temporal process. However, on a multi-dimensional Euclidean lattice, it is still possible to estimate the probability of a large block in terms of the probabilities of the smaller blocks it contains. In this paper we present a formula which uses the probabilities of blocks of a fixed finite size to estimate the probabilities of blocks of larger sizes. From this, construction of the local structure theory for cellular automata on the lattice follows immediately.

In [5] we discussed the Bayesian extension process in one dimension. We defined two operators  $L$  and  $R$  which produced an  $(n-1)$ -block by truncating an  $n$ -block from the left or right respectively. These operators allowed us to define an operator  $\pi$  which mapped an order  $n$  block probability function  $P_n$  to an  $(n+1)$ -block probability function  $P_{n+1}$  as follows:

$$\pi(P_n)(B) \equiv \begin{cases} \frac{P_n(RB)P_n(LB)}{P_n(RLB)} & \text{if } |B| = n+1 \text{ and } P_n(RLB) > 0 \\ 0 & \text{if } |B| = n+1 \text{ and } P_n(RLB) = 0 \\ P_n(B) & \text{if } |B| \leq n. \end{cases}$$

We observed that repeated application of  $\pi$  to  $P_n$  produces a collection of functions  $\{P_m\}_{m \geq n}$ . Members of the sequence  $\{P_m\}$  assign probabilities to blocks of all sizes in a manner which satisfies the Kolmogorov consistency conditions. Thus this sequence defines a shift-invariant measure on the set of all one-dimensional configurations (Denker [3]). In more than one dimension, simple generalizations of the procedure do not in general provide "block probabilities" that satisfy Kolmogorov consistency. Nevertheless, the procedure still provides approximate probabilities that may be used as the basis of a higher dimensional local structure theory. We will need some definitions to describe this more general setting.

## 2. Fundamentals, Notation

### 2.1 Lattices, Frames, Blocks

Let  $\mathbf{L}$  be a discrete Euclidean lattice of dimension  $k$ , with a translation group  $G_{\mathbf{L}}$ . Examples are the one-dimensional lattice of integers  $\mathbf{Z}$  with the translation group of left and right shifts, and the two-dimensional square lattice  $\mathbf{Z} \times \mathbf{Z}$  with the translation group of shifts left, right, up and down.

Each cell of a lattice may be labeled with one of a finite number of symbols  $\{s_0, s_1, \dots, s_l\} \in S$ , usually denoted  $\{0, 1, \dots, l-1\}$ .

A frame  $F$  is a finite subset of (not necessarily contiguous) cells of  $\mathbf{L}$ . The collection of all frames  $\mathbf{F}$  is closed under intersection  $\cap$ , union  $\cup$ ,

difference  $-$ ,  $(F_1 - F_2 \equiv F_1 \cap F_2^c)$ , and the action of  $G_L$ . Frames have a definite position in the lattice, but are unlabeled by states.

A block  $B$  is an assignment of state labels to the cells of a frame. The set of all blocks  $B$  for a given frame  $F$  will be denoted  $B_F$ . Blocks correspond to cylinder sets. If  $F_1 \subset F_2$  then a block  $B_1$  of  $F_1$  may be defined by restricting a block  $B_2$  of  $F_2$  to  $F_1$ . We write  $B_1 = B_2 \downarrow F_1$ . In addition,  $G_L$  has a natural action on blocks;  $\gamma$  in  $G_L$  maps members of  $B_F$  to members of  $B_{\gamma F}$ .

### 3. Block probability functions

A block probability function  $P_F$  for a frame  $F$  is an assignment of probabilities to all labelings and partial labelings of  $F$ . Thus, it is a map from all  $B \in B_F'$  (for each  $F' \subseteq F$ ) to the reals. It must satisfy the Kolmogorov consistency conditions:

$$P_F(B) \leq 0 \quad (3.1)$$

$$\sum_{B \in B_{F'}} P_{F'}(B) = 1 \quad (3.2)$$

and, for each  $B' \in bb_{F'}$ ,

$$P_F(B') = \sum_{B \downarrow F' = B'} P_F(B) \quad (3.3)$$

Note that it suffices to define  $P_F$  on  $B \in B_F$ ; condition (3.3) then forces unique assignments of values to partial labelings of  $F$ . Also, (3.2) is a special case of (3.3) with  $F' = \emptyset$ .

### 4. Bayesian extension in more than one dimension

The essence of Bayesian extension in one dimension is the maximum entropy estimation of block probabilities in terms of the probabilities of smaller blocks they contain. That is, Bayesian extension solves a variational problem in which the entropy is maximized subject to the Kolmogorov consistency constraints (3.1, 3.2, 3.3), and previously assigned probabilities of smaller blocks. In one dimension this variational problem solves uniquely in terms of rational functions of small block probabilities (Brascamp [2]). Moreover, extension of  $n$ -block probabilities to  $m$ -block probabilities  $m > n$  is simply obtained by recursive application of one step extension.

In more than one dimension maximum entropy extension is considerably more delicate. The variational problem translates into a system of polynomial equations, whose explicit solution in terms of radicals is in general not possible. Furthermore, extension to each larger frame typically requires solution of an entirely new (and larger) system of equations. In this paper we do not solve the more general problem. Rather, we approximate the maximum entropy extension by a formula which is a straight-forward

analog of the one-dimensional formula. As in the one dimensional case, the formula is rational in smaller block probabilities.

To describe the approximate Bayesian extension of  $k$ -dimensional blocks, let us restrict attention to blocks on frames which are parallelepipeds. On these blocks, define operators  $L_j$  and  $R_j$  that truncate from the left or right respectively in the  $j$ th dimension. The operators  $L_1, \dots, L_k$  and  $R_1, \dots, R_k$  all commute.

To define Bayesian extension along dimension  $j$ , let  $F$  be an  $(i_1 \times \dots \times i_k)$ -frame, and  $P_F$  a block probability function on  $F$ . Let  $F_{*j}$  be an extension of  $F$  by one unit along dimension  $j$ . Then the Bayesian extension of  $P_F$  along dimension  $j$  is defined as the following function on blocks of  $F_{*j}$ :

$$\pi_j(P_F)(B) \equiv \begin{cases} \frac{P_F(R_j B) P_F(L_j B)}{P_F(L_j B)} & \text{if } |B| \in \mathbf{B}_{F_{*j}} \text{ but } B \in \mathbf{B}_F \\ 0 & \text{if } P_F(R_j L_j B) = 0 \\ P_F(B) & \text{if } B \in \mathbf{B}_F. \end{cases}$$

In one dimension, Bayesian extension always produces functions which satisfy all of the Kolmogorov consistency conditions; i.e., Bayesian extension always produces block probability functions. In more than one dimension this is not always the case. Examples may be constructed in which either the normalization condition (3.2) or local shift invariance is violated (Schlijper [9, 10]) and personal communication). If correlations exist in at most one dimension, then the problem may be reformulated as a one-dimensional problem in a larger state space, and no such pathologies exist. However, when correlations exist in two or more dimensions, (4, above) must be modified such that condition (3.2) is satisfied. To satisfy the requirement of shift invariance within the extended frame, it suffices to average block probabilities on frames which are related by a shift. In this paper we will use the term Bayesian extension to refer to extension via formula (4, above) appropriately adjusted so that the normalization and local shift invariance conditions are satisfied. The adjustment used will be described in conjunction with the numerical experiments.

## 5. Finite block measures

Consider a parallelepiped frame  $F$  and an infinite sequence of frames beginning with  $F$  which extends along each dimension in turn. This sequence of frames grows without bound in all dimensions. Now suppose it possible to use Bayesian extension to define a block probability function  $P_F$  on each frame in the sequence such that each  $P_F$  is consistent with block probability functions on frames contained in  $F$ . In this case it is possible to assign probabilities in a consistent fashion to all frames. By the Kolmogorov consistency theorem (Denker [3]), these probability assignments define a unique measure on the entire lattice  $\mathbf{L}$ , which we will denote  $\mu(P_F)$ . As in [5] we call such measure a finite block measure. If some infinite extension exists then a shift-invariant extension exists by averaging an (infinite) measure

with respect to the shift. Since shift-invariance is enforced at each stage of the Bayesian extension process, if  $P_F$  is shift-invariant, so is  $\mu(P_F)$ .

Given a block probability function  $P_F$ , is it possible to define  $\mu(P_F)$ ? In more than one dimension, the question is in general undecidable (Schlijper [9]). However, if correlations exist in at most one dimension the construction reduces to the one dimensional case. Furthermore, at least some sets of block probabilities with non-trivial correlation in more than one dimension may be defined using the methods of Julesz et al. [6] and Victor [12]. General properties of entropy (and linearity of the Kolmogorov consistency constraints) imply that if any extension to a finite block measure exists then a unique maximum-entropy extension exists. Thus, although Bayesian extension,  $n$ -step Markov processes and maximum entropy extensions coincide in one dimension, these concepts are distinct in higher dimensions.

## 6. The local structure theory

Here we summarize the development of the local structure theory in the present context. For a fuller discussion see [5]. Let  $B$  be a block on some frame in  $\mathbf{F}_L$ . The action of a cellular automaton  $\tau$  on an arbitrary  $G_F$ -invariant measure  $\mu$  may be defined (Lind [7]) by

$$\tau\mu(B) = \mu(\tau^{-1}(B)). \quad (6.1)$$

Since a cellular automaton is a local map, the inverse image of a block is a collection of finite, though typically larger, blocks. Since the inverse image of a finite block is finite, we can define a finite approximation to (6.1). Let us first assume that the restriction of  $\mu$  to a frame  $F$  is a block probability function which may be infinitely extended. That is, we assume that  $\mu$  may be approximated by a finite block measure  $\mu_F$ . We may operate on  $\mu_F$  with the cellular automaton, and then again restrict the resulting measure to  $F$ . If this new assignment of probability to blocks on  $F$  is extendable, we may repeat the process.

We introduce an operator,  $\sigma_F$ , which operates on measures.  $\sigma_F$  is called the scramble operator for a frame  $F$ . It maps a measure  $\mu$  to the finite block measure  $\sigma_F(\mu)$  which agrees with  $\mu$  on all blocks of the frame  $F$ . The scramble operator, combined with the action of the cellular automaton map on measures, defines the local structure operator  $\Delta_F(\tau)$  by

$$\Delta_F(\tau)(\mu) = \sigma_F \tau \sigma_F(\mu). \quad (6.2)$$

$\Delta_F(\tau)$  maps the set of finite block measures on  $F$  into itself, in accordance with the cellular automaton rule.

In order to define  $\Delta_F(\tau)(\mu)$ , an operator which approximates the action of a cellular automaton on measures on infinite configurations, we assumed that a block probability function on  $F$  could be consistently extended to block probability functions on frames of all larger size. In more than one dimension, there is no guarantee that a block probability function is extendable to a finite block measure  $\mu(P_F)$  on infinite configurations, and

$\sigma_F(\mu)$  is not necessarily defined. However, cellular automata can also be characterized by how they map (finite) block probability functions to other block probability functions. This motivates an interpretation of the local structure operator  $\Delta_F(\tau)$  as an operator on block probability functions  $P_F$  rather than on finite block measures  $\mu(P_F)$ .

Given a cellular automaton  $\tau$ , a block probability function  $P_F$  and a block  $B$  on  $F$ , it is natural to assign  $\tau(P_F)(B)$  to be the probability of  $\tau^{-1}(B)$ , as suggested by equation (6.1).  $\tau^{-1}(B)$  is a union of blocks  $B'$  on a larger frame  $F'$  which contains  $F$ . The probability of each such  $B'$  of  $F'$  is not given directly by  $P_F$ , but may be estimated by extension of  $P_F$  to a block probability function on  $F'$ . Thus a finite extension suffices to estimate the action of  $\tau$  on a block probability function  $P_F$ , for any finite frame  $F$ . This construction forms the basis of the computations described below.

## 7. Local structure theory on the square lattice

### 7.1 The mean-field theory

The mean-field theory for cellular automata assumes that no correlation between states of cells is generated by cellular automata evolution. This assumption leads to a simple formula for the density of 1's at time  $t+1$  in terms of the density at time  $t$  (Schulman and Seiden [11]). Let  $\#0(B)$  and  $\#1(B)$  be the number of 0's and 1's in a block  $B$ , and  $p_t$  be the density of 1's at time  $t$ . The mean-field theory on the square lattice states that

$$p_{t+1} = \sum_B \tau(B) p_t^{\#1(B)} (1 - p_t)^{\#0(B)}, \quad (7.1)$$

where the sum is taken over all  $(3 \times 3)$ -blocks. The density at time  $t+1$  is a polynomial function of the density at time  $t$ . The fixed points of this polynomial recursion equation are taken as an approximation to the density in the limit of large time. As we showed previously (Gutowitz et al. [5]), the local structure theory on a  $(1 \times 1)$ -frame is identical with the mean-field theory.

## 8. The $(1 \times n)$ -Theory on the Square Lattice

Let  $B$  be a (horizontal)  $(1 \times 2)$ -block on the square lattice, and  $\tau$  a cellular automaton which operates on the lattice. The inverse image of  $B$  under  $\tau$  is a union of  $(3 \times 4)$ -blocks. To implement the  $(1 \times 2)$ -order local structure theory, we need to compute by Bayesian extension the probability of the  $(3 \times 4)$ -blocks in terms of probabilities assigned to  $(1 \times 2)$ -blocks at a given generation. Then, by summing over the probabilities of the  $(3 \times 4)$ -blocks in the inverse image of  $B$ , we obtain an estimate of the probability of  $B$  at the next generation.

Let us call a sequence of frames  $\{F_i\}$  a Bayesian sequence if each  $F_{i+1}$  is an extension of  $F_i$  along some single dimension, and write  $\dots \triangleleft F_i$

$\triangleleft F_{i+1} \triangleleft \dots$ . By using the Bayesian sequence  $(1 \times 2) \triangleleft (2 \times 2) \triangleleft (2 \times 3) \triangleleft (3 \times 3) \triangleleft (3 \times 4)$ , we obtain an expression for the probability of a  $(3 \times 4)$ -block in terms of the probabilities of its  $(1 \times 2)$ -sub-blocks. Since only one-dimensional correlation may be expressed in terms of  $(1 \times 2)$ -blocks, (3) needs no adjustment. We will use the following notation: The rows of a  $(3 \times 4)$ -block are numbered from top to bottom and the columns from left to right. A (horizontal)  $(1 \times 2)$ -block whose leftmost cell is at position  $i, j$  in the  $(3 \times 4)$ -block is denoted  $B_{ij}^h$ . Likewise, a (singleton)  $(1 \times 1)$ -block at position  $i, j$  is denoted  $B_{ij}^s$ . Using this notation, the probability of a  $(3 \times 4)$ -block is estimated by

$$\prod_{i=1}^3 \left[ \frac{\prod_{j=1}^3 P(B_{ij}^h)}{\prod_{j=2}^3 P(B_{ij}^s)} \right]. \quad (8.1)$$

In the  $(1 \times 2)$ -theory rows do not interact via the block probability function, only via the cellular automaton rule itself. In essence, the block probability is the product of three one-dimensional functions. Thus we may predict that the  $(1 \times 2)$ -theory could only account for the generation of correlation in one direction. This prediction is supported by the experiments discussed below.

The  $(1 \times n)$ -theory for general  $n$  is constructed in the same way as the  $(1 \times 2)$ -theory.

## 9. The $(2 \times 2)$ -theory

The inverse image of a  $(2 \times 2)$ -block under a nearest neighbor cellular automaton on the square lattice is a union of  $(4 \times 4)$ -blocks. A Bayesian extension sequence leading from a  $(2 \times 2)$ -block to a  $(4 \times 4)$ -block is:  $(2 \times 2) \triangleleft (2 \times 3) \triangleleft (3 \times 3) \triangleleft (3 \times 4) \triangleleft (4 \times 4)$ . We will denote a  $(2 \times 2)$ -block whose upper-left cell is at position  $i, j$  in a  $(4 \times 4)$ -block by  $B_{ij}$ . A (vertical)  $(2 \times 1)$ -block whose top cell is at position  $i, j$  is denoted  $B_{ij}^v$ .  $B_{ij}^h$  and  $B_{ij}^s$  have the previously assigned meanings. Using this notation, the probability of a  $(4 \times 4)$ -block is estimated in terms of  $(2 \times 2)$ -block probabilities by

$$\frac{\left[ \prod_{i=j}^3 \prod_{i=j}^3 P(B_{ij}) \right] \left[ \prod_{i=2}^3 \prod_{i=2}^3 P(B_{ij}^s) \right]}{\left[ \prod_{i=1}^3 \prod_{j=2}^3 P(B_{ij}^v) \right] \left[ \prod_{i=j}^3 \prod_{j=1}^3 P(B_{ij}^h) \right]} \quad (9.1)$$

That is, we multiply together all possible  $(2 \times 2)$ -blocks in the  $(4 \times 4)$ -block, divide by the probability of the  $(1 \times 2)$ - and  $(2 \times 1)$ -block intersections, and then multiply by the probabilities of the intersections of the  $(1 \times 2)$  and  $(2 \times 1)$ -blocks, which are the central  $(1 \times 1)$ -blocks.

In the one-dimensional theories, there is no possibility of violation of the normalization condition (3.2). In the  $(2 \times 2)$ -theory, there is this possibility



(Schlijper [9,10]). In the experiments described below, normalization is preserved as follows: Equation (8) is used to assign values to all  $(4 \times 4)$ -blocks save the block composed of all 0's. The sum of these values subtracted from 1 is then assigned to the 0  $(4 \times 4)$ -block. Since the 0  $(4 \times 4)$ -block is a predecessor only of the 0  $(2 \times 2)$ -block, we can make the same adjustment by computing the probabilities of the inverse images of all  $(2 \times 2)$ -blocks save that of the 0  $(2 \times 2)$ -block. Since the 0  $(2 \times 2)$ -block is the  $(2 \times 2)$ -block with the largest number of predecessors, this maneuver markedly reduces the amount of computation needed at each iteration of the local structure theory.

In the experiments described below the amount of adjustment needed is quite small. If we used the fixed-point  $(2 \times 2)$ -block probabilities of the  $(2 \times 2)$ -theory (see below) to estimate the probabilities of all  $(4 \times 4)$ -blocks via equation (8), the sum of these estimates is 1.02. At small time less correlation has been generated by cellular automaton evolution, and less adjustment is needed.

## 10. Empirical studies of the Game of Life

The Game of Life [1] is an outer totalistic [8] cellular automaton on the square lattice. The neighborhood of a cell includes the eight cells with touch the given cell. Under the Game of Life rule, if a cell is in state 1 and has either two or three neighbors also in state 1, then it remains in state 1. If a cell is in state 0 and has exactly three neighbors in state 1, then it goes to state 1 at the next generation. In all other situations, the cell goes to state 0 at the next generation. As is well known [1, 8, 11], this rule has very complicated dynamical behavior. In this section we study the evolution of low order local structure operators for Game of Life rule. We compare the probabilities assigned to blocks by these operators at each generation to Monte Carlo estimates of the same probabilities.

Figure 1a shows the density as a function of time as determined by Monte Carlo sampling, and the  $(1)$ ,  $(1 \times 2)$ ,  $(1 \times 3)$ , and  $(2 \times 2)$  local structure theory. The Monte Carlo results are derived from sampling 20  $(100 \times 100)$ -blocks with periodic boundary conditions. These blocks were generated to be unbiased and uncorrelated. The large time density is 0.37 for the first order (mean-field) theory, 0.30 for the  $(1 \times 2)$  and  $(1 \times 3)$  theories, and 0.036 for the  $(2 \times 2)$ -theory. The Monte Carlo density appears to be still decreasing after 512 generations. Shulman and Seiden [11], using (presumably more extensive) Monte Carlo computations report the large time density as  $0.029 \pm 0.009$ . The one-dimensional theories are only slightly better at predicting limit density than the (0-dimensional) mean-field theory. The fully two-dimensional theory is much better, however. Its predicted limit density is within experimental error of Schulman and Seiden's empirical estimate.

There are similar improvements in the estimates for blocks on frames of larger size. Figure 1b and 1c show the evolution of two selected  $(2 \times 2)$ -



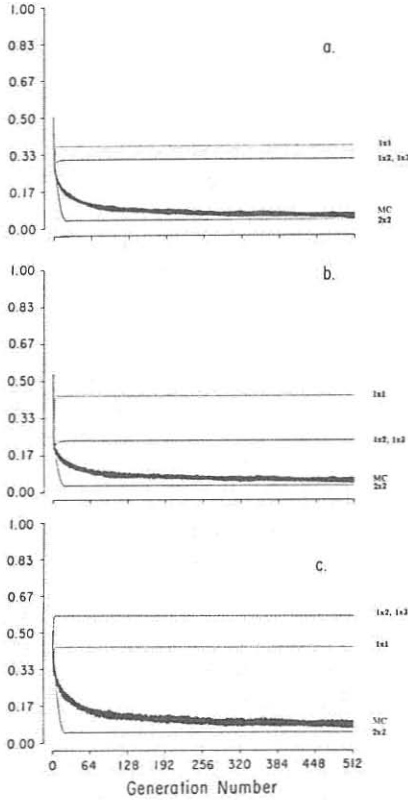


Figure 1: The evolution of the density and two  $(2 \times 2)$ -block probabilities under Monte Carlo simulations of the Game of Life rule, and the  $(1)$ ,  $(1 \times 2)$ ,  $(1 \times 3)$ , and  $(2 \times 2)$  - local structure theory. (a) the density  $P(1)$ , (b), the  $\begin{pmatrix} 10 \\ 01 \end{pmatrix}$ -block, and (c) the  $\begin{pmatrix} 00 \\ 11 \end{pmatrix}$ -block. Probabilities are followed for 512 generations from unbiased, uncorrelated initial conditions. The  $(2 \times 2)$ -block probabilities are scaled by  $2^4$ .  $\pm 1$  standard deviation is indicated at each generation.

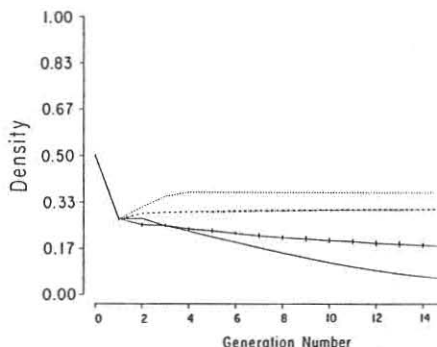


Figure 2: Figure 1a on an expanded time scale. Evolution of the density for 16 generations under the Game of Life rule.

block probabilities: the  $\binom{10}{01}$ -block and the  $\binom{00}{11}$ -block. These data serve to emphasize several points concerning the density estimate. First, that the  $(1 \times 2)$  and  $(1 \times 3)$ -theories are hardly distinguishable. Second, that the  $(2 \times 2)$ -theory is much better at predicting large time probabilities than the lower dimensional theories. While the one-dimensional theories are typically better than the zero-dimensional theory, this is not always the case (figure 1c). All orders of theory accurately predict the density at the first generation (figure 2). The density at the first generation is approximately 0.27. This is simply the fraction of  $(3 \times 3)$ -blocks which yield 1 upon application of the cellular automaton. As correlations develop, all orders of theory studied depart from the empirical curve. The zero and one-dimensional theories rise to their final values. The  $(2 \times 2)$ -theory proceeds in the correct direction (downward), but at a much faster rate than the cellular automaton itself.

The  $(2 \times 2)$ -theory is the first in the sequence of theories which connects correlation in the two dimensions. Bayesian extension followed by an ad hoc adjustment to ensure normalization results in a dramatic improvement in the block probability estimates in the  $(2 \times 2)$  case even though the  $(1 \times n)$ -theory rests on a rigorous maximum-entropy extension and the  $(2 \times 2)$ -theory does not. It is illuminating to consider two dimensional configurations which are obtained by simply copying a one dimensional vertical configuration to all positions horizontally. The Game of Life maps the set of such "bar-code" configurations into itself. On these configurations the Game of Life acts as the one-dimensional  $r = 1$  rule 22 on vertical cross-sections. We have previously demonstrated [5] that the one-dimensional local structure theory of low order accurately predicts the statistical behavior of rule 22. This does not say however that the  $(1 \times n)$ -theory is an accurate representation of the action of the Game of Life on this special set of configurations. The one dimensional theories cannot tell that the con-

figurations have a uniform structure in the horizontal direction, however sensitive they may be to correlation in the vertical direction. Hence, even an approximate method of relating structure in the two dimensions may result in improved predictions.

## 11. Conclusions

Clearly, this paper does not exhaust the possibilities inherent in its subject matter. Topics of current research interest include the following: What do the higher order theories for complicated rules like the Game of Life look like? How does the geometry of a lattice shape the behavior of cellular automata on that lattice? What is the relationship between exact maximum-entropy extension and its approximation by Bayesian extension?

When cellular automata are used as models of physical processes, the question arises whether the model adequately captures the behavior of the physical system. In the same way, one is faced with the question of whether the local structure theory adequately captures the behavior of the model. If it does, benefits may accrue from working with the local structure approximations rather than the directly with the cellular automaton itself in that analytical techniques may replace empirical techniques.

## Acknowledgements

This work was supported by grants EY 6871, EY 1428 and NS 877 from the NIH, the Hartford Foundation, the McKnight Foundation, and a Rockefeller University Graduate Fellowship. The authors are indebted to Dr. A.G. Schlijper for his insightful comments on an earlier draft.

## References

- [1] E.R. Berlekamp, J.H. Conway, and R.K. Guy, *Winning ways for your mathematical plays, Volume II* (Academic Press, N.Y., 1982).
- [2] H.J. Brascamp, "Equilibrium states for a one dimensional lattice gas", *Communications in Mathematical Physics* **21** (1971) 56.
- [3] M. Denker, "Ergodic theory on compact spaces", *Lecture Notes and Mathematics* **527**, Springer-Verlag. (1976).
- [4] M. Fannes and A. Verbeure, "On solvable models in classical lattice systems", *Communications in Mathematical Physics* **96** (1984) 115.
- [5] H.A. Gutowitz, J.D. Victor, and B.W. Knight, *Local structure theory for cellular automata*, *Physica D*, submitted (1986).
- [6] B. Julesz, E.N. Gilbert, and J.D. Victor, "Visual discrimination of textures with identical third-order statistics", *Biological Cybernetics* **31** (1978) 137.

- [7] D.A. Lind, "Applications of ergodic theory and Sofic systems to cellular automata", *Physica D* **10** (1984) 36.
- [8] N. Packard and S. Wolfram, "Two-dimensional cellular automata", *Journal of Statistical Physics* **38** (1985) 901.
- [9] A.G. Schlijper, "Variational approximation in classical lattice systems", Thesis. University of Groningen, Netherlands (1985).
- [10] A.G. Schlijper, "On some variational approximations in two-dimensional classical lattice systems", *Journal of Statistical Physics* **40** (1985) 1.
- [11] L.S. Schulman and P.E. Seiden, "Statistical mechanics of a dynamical system", based on Conway's Game of Life, *Journal of Statistical Physics* **19** (1978) 293.
- [12] J.D. Victor, "Complex visual textures as a tool for studying the VEP", *Vision Research* **25** (1985) 1811.
- [13] S. Wolfram, "Statistical mechanics of cellular automata", *Reviews of Modern Physics* **55** (1983) 601.
- [14] S. Wolfram, (editor) *Theory and applications of cellular automata*, (World Scientific, 1986).