Journal of Cellular Automata, Vol. 0, pp. 1–14
Reprints available directly from the publisher
Photocopying permitted by license only

©2008 Old City Publishing, Inc.
Published by license under the OCP Science imprint,
a member of the Old City Publishing Group

Emergent Defect Dynamics in Two-Dimensional Cellular Automata

MARTIN DELACOURT[1] AND MARCUS PIVATO[2,*]

[1]*Laboratoire de l'Informatique du Parallélisme, École Normale Supérieure de Lyon, 46 Allée d'Italie, 69634 Lyon, France*
[2]*Department of Mathematics, Trent University, 1600 West Bank Drive, Peterborough, Ontario, K9J 7B8, Canada*

Received: November 1, 2007. Accepted: March 12, 2008.

A cellular automaton (CA) exhibits 'emergent defect dynamics' (EDD) if generic initial conditions rapidly coalesce into large, homogeneous 'domains' (exhibiting some spatial pattern) separated by moving 'defects'. There are many known examples of EDD in one-dimensional CA, but not in higher dimensions. We describe the results of an automated search for two-dimensional CA exhibiting EDD. We found a plethora of examples of EDD, but we also found that the proportion of CA with EDD declines rapidly with increasing neighbourhood size.

Keywords: Defect, dislocation, kink, domain boundary, emergent defect dynamics.

A well-known phenomenon in one-dimensional cellular automata (CA) is the emergence of homogeneous 'domains'—each exhibiting some regular spatial pattern—separated by *defects* (or 'domain boundaries' or 'kinks') which evolve and propagate over time, and occasionally collide. This *emergent defect dynamics* (EDD) is clearly visible, for example, in elementary cellular automata #18, #22, #54, #62, #110, and #184. EDD in one-dimensional CA has been studied both empirically [1–5, 11] and theoretically [6, 8–10, 13–16, 19]. Recently, we have begun to develop a theory of EDD in multidimensional CA [18, 20], but this theory is somewhat hampered by the dearth of known multidimensional examples.

This paper describes the results of an automated search for EDD in two-dimensional CA on the alphabet $\mathcal{A} := \{0, 1\}$. We unearthed many CA with complex and fascinating EDD, estimated the prevalence of EDD in

*E-mail: pivato@xaravve.trentu.ca

various 'spaces' of two-dimensional CA, and obtained some statistics about the qualitative properties of the EDD.

Domains and Defects. Heuristically, 'domains' and 'defects' are large-scale, emergent structures which are visible in the spacetime diagrams of CA. However, to study defects rigorously, we need a precise definition. Following [18–20], we define defects relative to some reference subshift. A *subshift* is a closed, shift-invariant subset $\mathbf{X} \subset \mathcal{A}^{\mathbb{Z}^D}$. For example, the support of a stationary Markov chain is a subshift of $\mathcal{A}^{\mathbb{Z}}$ (called a *Markov shift*). Any spatially periodic pattern (e.g. checkerboard, stripes, etc.) corresponds to a periodic subshift of $\mathcal{A}^{\mathbb{Z}^2}$. More generally, almost anything that might be called 'spatial order' corresponds to some subshift; see [12, 17].

If $\mathbf{X} \subset \mathcal{A}^{\mathbb{Z}^D}$ is a subshift, then for any finite $\mathbb{K} \subset \mathbb{Z}^D$, let $\mathbf{X}_{\mathbb{K}} := \{\mathbf{x}_{\mathbb{K}} \,;\, \mathbf{x} \in \mathbf{X}\}$ be the set of all $\mathbf{X}$-*admissible* $\mathbb{K}$-*blocks*. If $\mathsf{z} \in \mathbb{Z}^D$, then $\mathbf{X}_{\mathbb{K}} = \mathbf{X}_{\mathbb{K}+\mathsf{z}}$ (because $\mathbf{X}$ is shift-invariant). We say $\mathbf{X}$ is a *subshift of finite type* (SFT) if there is some finite $\mathbb{K}$ such that $\mathbf{X} = \{\mathbf{a} \in \mathcal{A}^{\mathbb{Z}^D} ;\, \mathbf{a}_{\mathbb{K}+\mathsf{z}} \in \mathbf{X}_{\mathbb{K}}, \forall \mathsf{z} \in \mathbb{Z}^D\}$. For example, if the elements of $\mathcal{A}$ are 'Wang tiles' (square tiles with edge-matching constraints), then the set of admissible 'tilings' of $\mathbb{Z}^2$ by $\mathcal{A}$ is an SFT of $\mathcal{A}^{\mathbb{Z}^2}$. If $\Phi : \mathcal{A}^{\mathbb{Z}^D} \longrightarrow \mathcal{A}^{\mathbb{Z}^D}$ is any CA, then the fixed-point set of Φ is an SFT of $\mathcal{A}^{\mathbb{Z}^D}$, as is the set of P-periodic points (for any fixed $P \in \mathbb{N}$).

If $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}^D}$, then the $\mathbf{X}$-*admissible region* of $\mathbf{a}$ is $\mathbb{A} := \{\mathsf{z} \in \mathbb{Z}^D;\, \mathbf{a}_{\mathbb{K}+\mathsf{z}} \in \mathbf{X}_{\mathbb{K}}\}$; this is the part of $\mathbf{a}$ which is covered by the 'homogeneous domains' exhibiting the 'regular pattern' encoded in $\mathbf{X}$. The complement $\mathbb{Z}^D \setminus \mathbb{A}$ is the *defective* part of $\mathbf{a}$. We can thus precisely define the 'defects' in $\mathbf{a}$, relative to any subshift $\mathbf{X}$ of finite type. (This definition can also be extended to subshifts *not* of finite type, at the cost of considerable technical complexity; see [18, 20] for details). The reference subshift $\mathbf{X}$ is arbitrary, but the 'correct' choice is usually obvious from context: loosely speaking, $\mathbf{X}$ should be the smallest subshift whose regular domains cover 'most' of the configuration $\mathbf{a}$.

Search Methodology. Our search method was roughly as follows:

1. Let L be large (e.g. $L := 1024$), and let $\mathbb{L} := \mathbb{Z}_{/L} \times \mathbb{Z}_{/L}$ be an $L \times L$ square lattice with toroidal boundary conditions. Let $\mathbf{a} \in \mathcal{A}^{\mathbb{L}}$ be a random initial condition.
2. Let $\Phi : \mathcal{A}^{\mathbb{L}} \longrightarrow \mathcal{A}^{\mathbb{L}}$ be a randomly generated 2-dimensional CA.
3. Compute $\mathbf{b} := \Phi^T(\mathbf{a})$ for some large T (e.g. $T := 100$)
4. Let ρ be the empirical probability distribution of $k \times k$ blocks in $\mathbf{b}$ (where e.g. $k := 3$).
5. Examine ρ for the 'statistical signature' of emergent defects.
6. Isolate likely candidates and analyze their defect structure.

Statistical signature of EDD. Let $k \in \mathbb{N}$ and let $\mathbb{K} := [0 \dots k)^2 \subset \mathbb{Z}^2$. If $|\mathcal{A}| = 2$, then $|\mathcal{A}^{\mathbb{K}}| = 2^{k^2}$. For any $\mathbf{c} \in \mathcal{A}^{\mathbb{K}}$, let

$$\rho(\mathbf{c}) := \frac{\#\{\ell \in \mathbb{L}\ ;\ \mathbf{b}_{\mathbb{K}+\ell} = \mathbf{c}\}}{L^2} \tag{1}$$

be the 'frequency' with which $\mathbf{c}$ appears in $\mathbf{b} := \Phi^T(\mathbf{a})$; thus, $\rho : \mathcal{A}^{\mathbb{K}} \longrightarrow [0, 1]$ is the 'empirical' probability distribution of $\mathbb{K}$-blocks observed in $\mathbf{b}$. To detect EDD, let $-\mathbb{N} := \{\dots, -3, -2, -1, 0\}$, and define the *log-probability histogram* $h : (-\mathbb{N}) \longrightarrow [0, 1]$ as follows:

$$h(-n) := \sum \{\rho(\mathbf{c})\ ;\ \mathbf{c} \in \mathcal{A}^{\mathbb{K}} \text{ and } 2^{-n-1} < \rho(\mathbf{c}) \leq 2^{-n}\}. \tag{2}$$

In other words, $h(-n)$ is the probability that a ρ-random $\mathbb{K}$-block $\mathbf{c}$ will have a probability $\rho(\mathbf{c})$ close to 2^{-n}.

Examples: (a) Suppose ρ is the *uniform measure*—i.e. $\rho(\mathbf{c}) = 2^{-k^2}$ for all $\mathbf{c} \in \mathcal{A}^{\mathbb{K}}$. Then $h(-k^2) = 1$, and $h(x) = 0$ for all $x \neq -k^2$.

(b) If $\mathbf{b}$ is a constant configuration, then ρ is a *point mass*—i.e. $\rho(\mathbf{c}) = 1$ for some $\mathbf{c} \in \mathcal{A}^{\mathbb{K}}$, and $\rho(\mathbf{c}') = 0$ for all other $\mathbf{c}' \in \mathcal{A}^{\mathbb{K}}$. Then $h(0) = 1$, and $h(x) = 0$ for all $x < 0$.

(c) Suppose $\mathbf{b}$ is a checkerboard, or a pattern of alternating stripes of width one (vertical or horizontal). Then ρ is supported on two distinct $\mathbb{K}$-blocks, each with probability 2^{-1}. Thus, $h(-1) = 1$, and $h(x) = 0$ for all $x \neq -1$.

(d) Let $\mathbb{P} \subset \mathbb{Z}^2$ be a subgroup of index P, and suppose $\mathbf{b}$ is a $\mathbb{P}$-periodic pattern. Then ρ is supported on P distinct $\mathbb{K}$-blocks, each with probability $1/P$. Thus, $h\lfloor -\log_2(P)\rfloor = 1$, and $h(x) = 0$ for all $x \neq \lfloor -\log_2(P)\rfloor$.

(e) Suppose $\mathbf{b}$ is a generic configuration for some stationary measure on $\mathcal{A}^{\mathbb{Z}^2}$ with entropy $\eta \in [0, 1]$. If k is large enough, then the Shannon-MacMillan-Brieman Theorem [7] says $\rho(\mathbf{c}) \approx 2^{-\eta k^2}$ for roughly $2^{\eta k^2}$ distinct $\mathbf{c} \in \mathcal{A}^{\mathbb{K}}$, and $\rho(\mathbf{c}') \approx 0$ for all other $\mathbf{c}' \in \mathcal{A}^{\mathbb{K}}$. Thus, $h\lfloor -\eta k^2\rfloor \approx 1$ while $h(x) \approx 0$ for all $x \not\approx \eta k^2$.

In each case, h has a single 'spike' at a particular log-probability value; this value crudely measures the 'complexity' of the pattern in $\mathbf{b}$. Now suppose that $\mathbf{b}$ consisted *mainly* of this pattern, but also contained a few low-probability 'defects'; this will slightly diminish the 'spike' in h, and introduce a small 'tail' of lower-probability events. This 'spike and tail' pattern is the statistical signature of defects. However, for our automated search, we used a somewhat simpler criterion: low entropy. The *entropy* of ρ is defined

$$\eta(\rho) := -\sum_{\mathbf{c} \in \mathcal{A}^{\mathbb{K}}} \rho(\mathbf{c}) \cdot \log_2[\rho(\mathbf{c})] \approx \sum_{n=0}^{\infty} n h(-n).$$

Thus, $\eta(\rho)$ is small iff $h(-n)$ is large for some small $n \in \mathbb{N}$; this indicates that large regions of $\mathbf{b}$ are admissible to some low-entropy (or zero-entropy)

subshift. In practice, such a **b** almost always contains defects. Thus, we used a cap $\eta(\rho) \leq \overline{\eta}$ as a simple heuristic test for EDD. Empirically, we found that setting $\overline{\eta} := 8.0$ produced the best results (i.e. it detects all the CA which clearly show EDD, and rejects most of the ones which don't).*

CA search spaces. For any finite 'neighbourhood' $\mathbb{H} \subset \mathbb{Z}^2$, let $\mathsf{CA}(\mathbb{H})$ be the set of all CA on $\mathcal{A}^{\mathbb{Z}^2}$ with a local rule $\phi : \mathcal{A}^{\mathbb{H}} \longrightarrow \mathcal{A}$. Clearly, $|\mathsf{CA}(\mathbb{H})| = |\mathcal{A}|^{|\mathcal{A}|^{|\mathbb{H}|}}$. Thus, if $|\mathcal{A}| = 2$ and $|\mathbb{H}| = 3$, then $|\mathsf{CA}(\mathbb{H})| = 2^{2^3} = 256$, but if $|\mathbb{H}| \geq 5$, then $|\mathsf{CA}(\mathbb{H})|$ becomes astronomical. Each neighbourhood $\mathbb{H}$ thus determines a 'space' of CA, to which we can apply our search procedure. If $|\mathbb{H}| = 3$ or 4, then it is feasible to search $\mathsf{CA}(\mathbb{H})$ exhaustively, but for larger $\mathbb{H}$, it is only feasible to search a statistically representative subset of $\mathsf{CA}(\mathbb{H})$.

We say $\mathbb{H}$ *generates* $\mathbb{Z}^2$ if any $\mathsf{z} \in \mathbb{Z}^2$ can be written as a sum of elements in $\mathbb{H}$ (thus, every cell in $\mathbb{Z}^2$ is eventually 'visible' to every other cell). We considered all neighbourhoods $\mathbb{H} \subseteq \{-1, 0, 1\}^2$ which generated $\mathbb{Z}^2$ (neglecting neighbourhoods which were duplicates under reflection or rotation). We will indicate a neighbourhood by glyph which depicts its shape in miniature. For example ▚ denotes the three-cell neighbourhood $\{(1, 0), (0, 1), (-1, -1)\}$, while ✚ is the von Neumann neighbourhood and ▦ is the Moore neighbourhood.

Log probability landscapes. Suppose we compute the log-probability histogram (2) for many CA in $\mathsf{CA}(\mathbb{H})$, and treat each histogram as a one-dimensional 'slice' of a two-dimensional surface. We must shuffle the histogram 'slices' so that they are arranged *not* in terms of any numerical code representing the underlying CA, but rather, so that similar histograms (i.e. with spikes in roughly the same locations) are placed close to one another. The result is a rugged but surprisingly regular 'landscape', whose major features correspond to clusters of CA having asymptotic measures with similar statistical properties. These landscapes can be seen as abstract depictions of the large-scale structure of the CA universe.

Figure 1(▚) shows the 'probability landscape' of all 256 CA in CA(▚), obtained using $k = 3$. The ridge (A) at the far end is caused by 38 CA which appear to almost-preserve the uniform measure: all the mass of the histogram h is concentrated on $y \in \{-9, -10\}$, which means that $\rho(\mathbf{c}) \approx 2^{-9}$ for all $\mathbf{c} \in \mathcal{A}^{\mathbb{K}}$ [Figure 2(▚A) shows a closeup of this ridge]. The 'mountain range' (B) is CA which do not preserve the uniform measure, but which do not converge rapidly to any low-entropy measure. The low ridges (C) are CA which show some indication of of EDD, but where the asymptotic probability measure is still quite complex. Finally, the ridges (D) in the right-hand

*We obtained more refined results (i.e. less 'false positives') using more multifaceted statistical criteria, but these criteria were too complicated and *ad hoc* to admit any simple scientific interpretation.

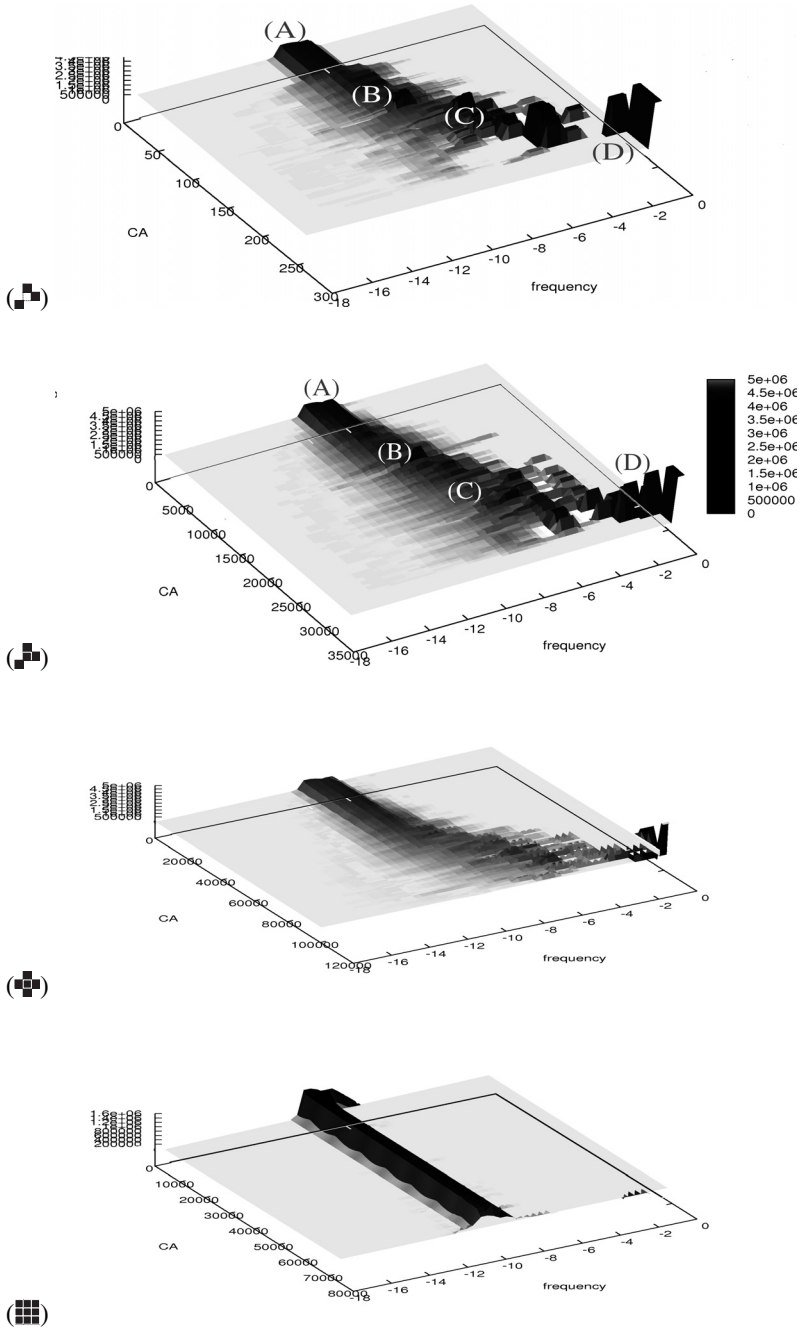
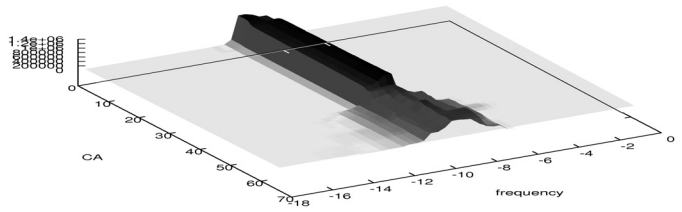
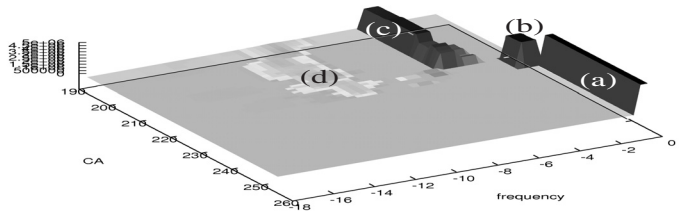


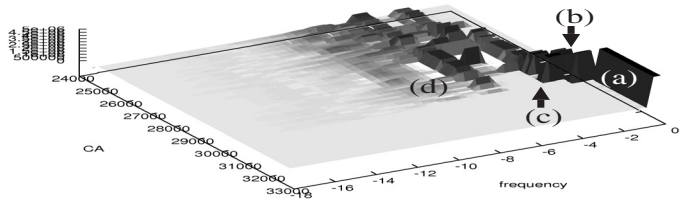
FIGURE 1
Log probability landscapes: large-scale view.



(▚A)



(▚D)



(▚D)

FIGURE 2
Log probability landscapes: closeups

corner are caused by CA which unambiguously manifest EDD; the asymptotic probability measure is almost all supported on one or more simple periodic background patterns, with the remaining probability supported on defect structures. Figure 2(▚D) shows a closeup of region (D).

The 'wall' (a) in Figure 2(▚D) is caused by 32 *nilpotent* CA, which converge to a constant (all-zeros or all-ones) configuration. The massif (b) is caused by 6 CA whose asymptotic subshift $\mathbf{X}$ has exactly 2 elements (e.g. 2-periodic flickering between all-zeros and all-ones). The ridge (c) is about 30 CA with $4 \leq |\mathbf{X}| \leq 16$ (juxtapositions of checkerboards, stripes, etc.). The low mounds (d) are caused by the low-probability defects.

Figure 1(▚) shows the probability landscape generated by the 65536 elements of CA(▚). The picture is very similar to the landscape of CA(▚); features (A)–(D) have the same interpretation. Figure 2(▚D) shows a closeup

of region (D) in Figure 1(▚). The 'wall' (a) is caused by approximately 7400 nilpotent CA. The 'teeth' (b) are caused by approximately 2000 CA whose asymptotic subshift $\mathbf{X}$ has cardinality 2 (e.g. flickering, cow pattern, width-one stripes or checkerboard). The teeth (c) are caused by approximately 600 CA where $3 < |\mathbf{X}| \leq 8$. The low mounds (d) are caused by defects.

Figure 1(✚) shows the probability landscape generated from a random sample of 100 000 out of the 4 294 967 296 elements of CA(✚). The picture is similar to the landscape for CA(▚) and CA(▞), but the proportion of CA with EDD is much smaller. Finally, Figure 1(▦) shows the probability landscape of CA(▦). Now $|\mathsf{CA}(▦)| = 2^{2^9} \approx 10^{154}$, so exhaustive coverage is out of the question; instead this landscape was generated using a random sample of size 70 000. The difference between Figure 1(▦) and Figures 1(▚, ▞, ✚) is obvious: the 'ridges' indicating EDD are totally absent. Of the 70 000 Moore CA we tested, less than 0.01% exhibited any statistical signature of EDD. Indeed, the pronounced ridge along the line $y = -9$ suggests (but does not prove) that the vast majority of Moore CA 'almost preserve' the uniform measure (or at least, its 3×3 marginal).

EDD vs. Neighbourhood size. Figures 1 and 2 indicate a striking phenomenon: *the proportion of CA exhibiting EDD declines sharply as the neighbourhood size increases.* This is illustrated clearly in Figure 3. Almost 70% of the CA in

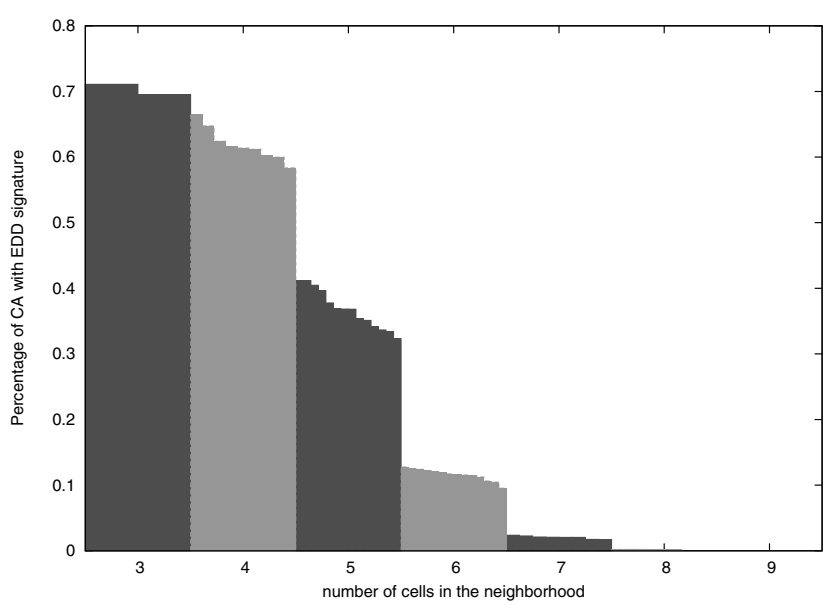


FIGURE 3
The percentage of CA($\mathbb{H}$) exhibiting $\eta(\rho) < 8.0$ (a crude test of EDD), as a function of $|\mathbb{H}|$.

CA(▚) and CA(▞) satisfy the criterion $\eta(\rho) < 8.0$ (upon visual inspection, most of these exhibit visible EDD, so the 'true' proportion of EDD is around 45%–50% in both cases). In contrast, only 35% of CA(▟) satisfy $\eta(\rho) < 8.0$ (and only 30% of CA(▙) exhibit EDD); for CA with larger neighbourhoods, the proportion is even less. The proportion of EDD in CA(▦) is virtually zero. We have no explanation for this phenomenon, but it is robust under several variations of experimental parameters:

- *Neighbourhood shape.* All neighbourhoods with the same cardinality have roughly the same proportion of EDD. In Figure 3, the separate columns over each number n correspond to all 'generating' neighbourhoods $\mathbb{H} \subseteq \{-1, 0, 1\}^2$ with $|\mathbb{H}| = n$ (modulo rotations and reflections). For example, the two columns over $n = 3$ correspond to ▞ and ▚. We arranged the columns in decreasing order of height, for clarity. There is some variation amongst neighbourhoods of cardinality n, but it is small compared to the difference between these and any neighbourhood of cardinality $n+1$.

- *Convergence time.* We looked for the criterion $\eta(\rho) < 8.0$ in CA(▟) and CA(▙) at times $T = 10, 20, \ldots, 80, 90, 100, 200, 300, 400$, and 500. The results are shown in Figure 4. Initially, as $T \nearrow 100$, there is a rapid increase in the the percent of CA with $\eta(\rho) < 8.0$; However, after $T = 100$, this levels off; this strongly suggests that almost all the convergence to EDD occurs during the first 100 iterations.

- *Block size.* We examined the probability distributions on $\mathcal{A}^{[0\ldots k)^2}$ for $k = 3$, 4 and 5. The three curves are shown in Figure 4, and clearly follow the same trajectory. Thus, using a large $\mathbb{K}$ does not seem to increase the proportion of detected EDD.

Interfaces and Dislocations. We earlier defined defects by comparison to some 'reference subshift' $\mathbf{X} \subset \mathcal{A}^{\mathbb{Z}^D}$. Many defects in $\mathbf{b}$ can then be understood as manifestations of 'global structural properties' of $\mathbf{b}$, relative to the topological dynamical system $(\mathbf{X}, \sigma)$ (where σ denotes the $\mathbb{Z}^D$-action by shifts on $\mathbf{X}$). If $\mathbf{X}$ is Φ-invariant, then Φ must preserve these 'structural properties', so Φ can neither create nor destroy defects—it can only move and combine them. The structural properties of defects were investigated in [18, 20].

For example, suppose $\mathbf{X} = \mathbf{X}_1 \sqcup \mathbf{X}_2$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are themselves nontrivial subshifts (called *transitive components* of $\mathbf{X}$). Then an element of $\mathbf{X}$ either belongs to $\mathbf{X}_1$ or $\mathbf{X}_2$. A domain boundary in $\mathbf{b}$ can then separate a $\mathbf{X}_1$-admissible region from a $\mathbf{X}_2$-admissible region, as shown in Figure 5(A); we call this an *interface*. If $\Phi(\mathbf{X}_1) = \mathbf{X}_1$ and $\Phi(\mathbf{X}_2) = \mathbf{X}_2$, then any such interface is preserved by Φ [20, Prop. 2.3].

Some structural properties can be associated with algebraic invariants. For example, let $\mathbb{P} \subset \mathbb{Z}^D$ be a subgroup of finite index, and suppose that $\mathbf{X}$

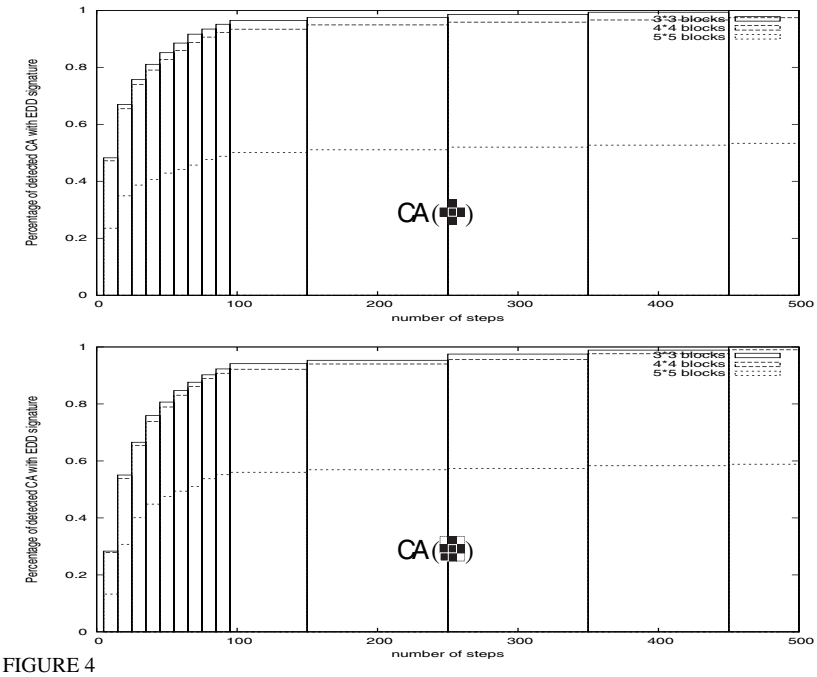
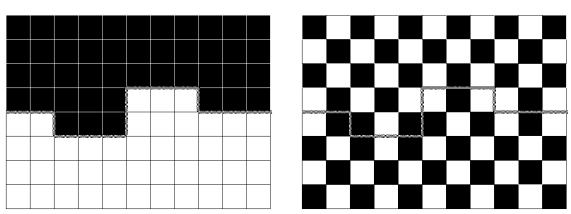


FIGURE 4
The percentage of $\mathsf{CA}(\mathbb{H})$ satisfying $\eta(\rho) < 8.0$ (i.e. potentially having EDD), as a function of time. The three curves are the generated using the probability distribution on $\mathcal{A}^{[0\ldots k)^2}$ for $k = 3$, 4 and 5.



(A) An interface **(B)** A dislocation

FIGURE 5

is $\mathbb{P}$*-periodic*—i.e. for any $\mathbf{x} \in \mathbf{X}$ and any $\mathsf{p} \in \mathbb{P}$, we have $\sigma^{\mathsf{p}}(\mathbf{x}) = \mathbf{x}$. A domain boundary in $\mathbf{b}$ can then separate two regions which are 'out of phase' with respect to this $\mathbb{P}$-periodic structure; we call this a *dislocation*. To any dislocation, we can associate a *displacement*: a unique element in the quotient group $\mathbb{Z}^D/\mathbb{P}$. Like an interface, a dislocation is Φ-invariant; furthermore, its displacement is constant over time. If two dislocations collide, then we simply add their displacements together in $\mathbb{Z}^D/\mathbb{P}$; see [20, § 3.1].

For example, if $\mathbf{X} \subset \mathcal{A}^{\mathbb{Z}^2}$ is the checkerboard, then $\mathbb{P}$ is the subgroup of all 'diagonal' shifts, generated by $(1, 1)$ and $(1, -1)$. Thus, $\mathbb{Z}^2/\mathbb{P} \cong \mathbb{Z}_{/2}$; this

	# of periodic transitive components in EDD background pattern											
CA	1		2		3		4		5		6	
	40	15.6%	25	9.77%	0	0%	3	1.17%	5	1.95%	0	0%
	10790	16.5%	202	0.31%	8	0.012%	14	0.021%	6	0.0091%	12	0.018%
*	3370	11.2%	63	0.21%	7	0.023%	0	0%	0	0%	0	0%

TABLE 1
The number (and percent) of CA exhibiting EDD and a background pattern with some number of periodic transitive components.
(∗) Statistics generated from a sample of 30000 vN CA.

reflects the fact that there are only two distinct 'phases' of the checkerboard. A dislocation in a checkerboard configuration is a domain boundary between one phase and the other, as in Figure 5(B).

Given a configuration $\mathbf{b} \in \mathcal{A}^{\mathbb{L}}$, it is possible to automatically identify the smallest subshift $\mathbf{X}$ such that 'most' of $\mathbf{b}$ is $\mathbf{X}$-admissible. We can then count the number of periodic transitive components of $\mathbf{X}$; if there are N transitive components, then there are $N(N-1)/2$ distinct types of interfaces. Table 1 presents this data for all members of CA( ) and CA( ), as well as a representative sample of CA( ).

Filtering images to reveal defects. Let $\mathbf{b} := \Phi^T(\mathbf{a})$ and $\mathbb{K} := [0 \ldots k)^2$, and let ρ be as in Eqn. (1). For all $\ell \in \mathbb{L}$, let $g_\ell := \rho(\mathbf{b}_{\ell+\mathbb{K}})$; this yields a configuration $\mathbf{g} \in [0,1]^{\mathbb{L}}$. If we visualize $\mathbf{g}$ as a 'greyscale' pixel-map image, then regular domains in $\mathbf{b}$ will appear as light grey (high probability) areas, whereas defects will appear as dark grey (low probability) regions. Figure 6 shows some typical examples of this procedure. Furthermore, the interfaces and/or dislocations in $\mathbf{b}$ can be revealed by automatically colour-coding each connected component of the light-grey region according to its transitive component and/or phase; animations generated using this process are available at http://xaravve.trentu.ca/DefectAnims.

A partial catalogue of EDD CA. Clearly it is not possible (or very useful) here to list the many thousands of CA we have found which exhibit EDD. Furthermore, such a list would be very incomplete, both due to our inherently imprecise heuristics for identifying EDD, and the smallness of our dataset compared to the vast CA-spaces they sample.

It *is* possible to give a fairly complete catalogue of EDD in CA( ) (which has only 256 members). Space constraints make it impossible to provide detailed information about each EDD CA; instead we have crudely classified them according to the number of distinct periodic patterns which they manifest. CA with one pattern exhibit dislocation defects; CA with two or more patterns exhibit both dislocations and interfaces. We only list CA with 'interesting' dynamics (e.g. not nilpotent, not eventually fixed/periodic).

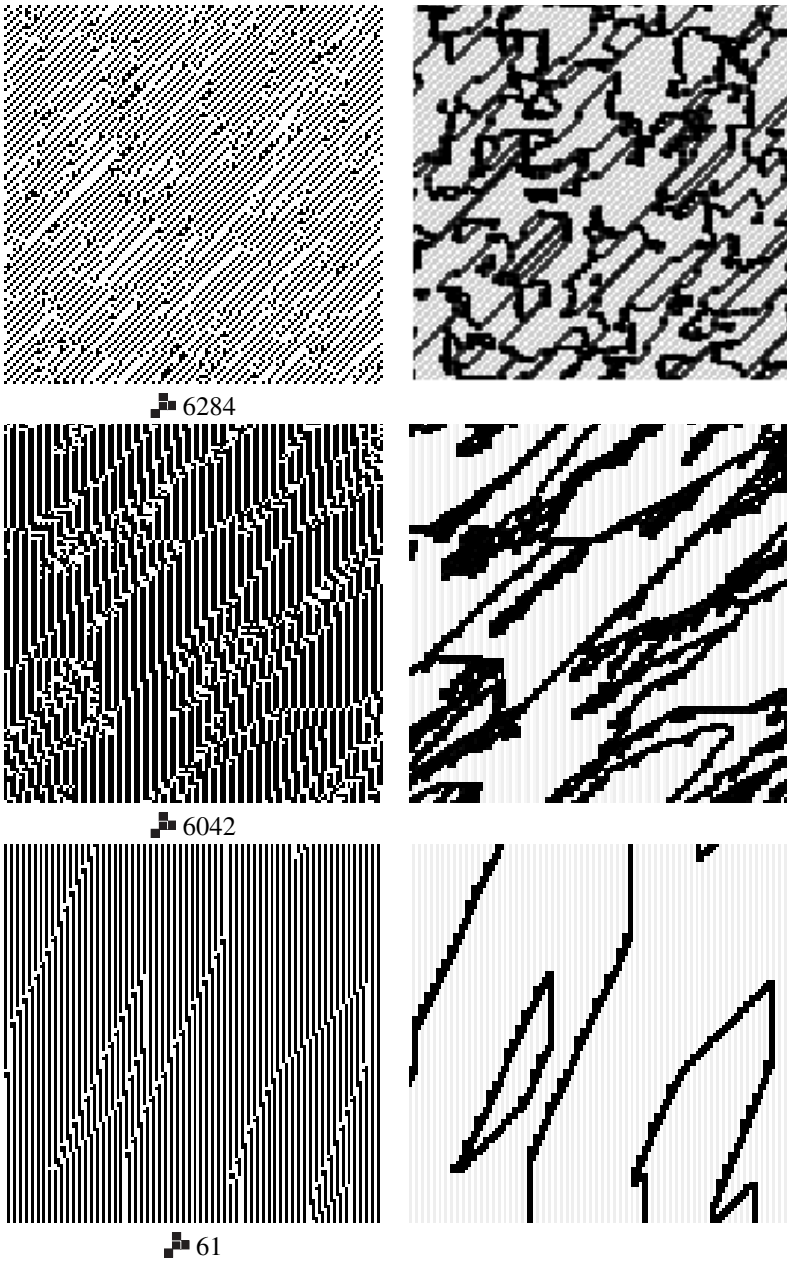


FIGURE 6
Left: Raw image. **Right:** Probabilistically filtered image, revealing defects.

We encode each CA using a Wolfram-style number, as follows: First the cells with coordinates (i, j) in the neighbourhood are numbered in increasing order of $(3j+i)$. Thus, the cells $(-1, -1)$, $(1, 0)$, and $(0, 1)$ of are numbered 0, 1, and 2, respectively. This maps $\{0, 1\}^{\text{[diagram]}} \longrightarrow \{0, 1\}^{\{0,1,2\}} \cong \{0, \ldots, 7\}$ by translating each element of $\{0, 1\}^{\text{[diagram]}}$ into a 3-bit binary number in the obvious way. Thus, any local rule (a map $\mathcal{A}^{\text{[diagram]}} \longrightarrow \{0, 1\}$) is translated into a map $\{0, \ldots, 7\} \longrightarrow \{0, 1\}$, which is then treated as an 8-bit binary number, between 0 and 255. By translating this numerical code back into a local rule, it is easy to simulate the following CA on a computer and observe their dynamics.

Two patterns. 11, 13, 14, 31, 35, 43, 47, 49, 50, 55, 59, 69, 77, 79, 81, 84, 87, 93, 113, 115, 117, 142, 143, 178, 179, 212, 213.

Four patterns. 232.

Five patterns. 62, 94, 118, 131, 133, 145.

For CA( ) (with 65536 members), we will provide only a short list of interesting examples.

One pattern.

Checkerboard. 7, 71, 199, 263, 311, 327, 423, 455, 519, 567, 583, 711, 775, 791, 807, 823, 839, 855.

Stripes of width 2. 61, 93, 110, 125, 126, 234, 281, 29, 317, 345, 385, 403, 46, 554, 574, 61, 618, 62, 642, 4714.

Stripes of width 3. 537, 601, 6042.

Two patterns. 119, 407, 679.

Checkerboard vs. Irregular. 55.

Three patterns. 599.

Four patterns. 359.

Five patterns. 295, 535.

For larger CA spaces, we are somewhat like astronomers trying to catalogue and classify the stars in the sky. Compiling such a catalogue will be a fascinating but endless project.

Conclusions. Emergent defect dynamics seems to be ubiquitous in two-dimensional CA. Even in the simplest classes of two-dimensional CA, our automated search uncovered a menagerie of examples. This preliminary exploration raises many questions.

- We have surveyed only 2-dimensional boolean CA (i.e. $\mathcal{A} = \{0, 1\}$). What is the prevalence of EDD in CA with larger alphabets? What about higher dimensions?

- Why is EDD common in CA with smaller neighbourhoods (e.g. and), yet very rare in CA with larger neighbhourhoods (e.g.)?
- In many of the CA we discovered, the defects exhibit complex emergent behaviour. For example, domain boundaries often move and evolve in complex ways. Is it possible to develop a mathematical description of domain boundary motion?
- In CA which converge to a spatially periodic structure, we automatically identified and classified the resulting dislocations. However, most CA with EDD do not converge to a simple spatially periodic structure. The theory of dislocations in [20] also applies to subshifts with *quasiperiodic* structure (e.g. Sturmian shifts, substitution shifts, Toeplitz shifts, etc.); however, we presently have no method to automatically identify such quasiperiodic structures and the corresponding dislocations.
- The theory in [18] describes an entirely different class of '(co)homological defects', which can exist even without spatially periodic structure. Theoretically, these defects can have quite complex properties and interactions. However, as yet, we have no algorithm to identify and classify such defects.

REFERENCES

[1] N. Boccara, J. Naser and M. Roger. Particle-like structures and their interactions in spatiotemporal patterns generated by one-dimensional deterministic cellular automata. *Physical Review A* **44**(2) (1991), 866–875.

[2] James P. Crutchfield and James E. Hanson. The attractor-basin portrait of a cellular automaton. *J. Statist. Phys.* **66**(5–6) (1992), 1415–1462.

[3] James P. Crutchfield and James E. Hanson. Attractor vicinity decay for a cellular automaton. *Chaos* **3**(2) (1993), 215–224.

[4] James P. Crutchfield and James E. Hanson. Turbulent pattern bases for cellular automata. *Phys. D* **69**(3–4) (1993), 279–301.

[5] James P. Crutchfield and James E. Hanson. Computational mechanics of cellular automata: an example. *Phys. D* **103**(1–4) (1997), 169–189. Lattice dynamics (Paris, 1995).

[6] James P. Crutchfield, Wim Hordijk and Cosma Rohilla Shalizi. Upper bound on the products of particle interactions in cellular automata. *Phys. D* **154**(3–4) (2001), 240–258.

[7] D.S. Ornstein and B. Weiss. The Shannon-McMillan-Breiman Theorem for a class of amenable groups. *Israel J. Math.* **44**(1) (1983), 53–60.

[8] Kari Eloranta. Partially permutive cellular automata. *Nonlinearity*, **6**(6) (1993), 1009–1023.

[9] Kari Eloranta. Random walks in cellular automata. *Nonlinearity*, **6**(6) (1993), 1025–1036.

[10] Kari Eloranta. The dynamics of defect ensembles in one-dimensional cellular automata. *J. Statist. Phys.* **76**(5–6) (1994), 1377–1398.

[11] P. Grassberger. Chaos and diffusion in deterministic cellular automata. *Phys. D* **10**(1–2) (1984), 52–58. Cellular automata (Los Alamos, N.M., 1983).

[12] Bruce Kitchens. *Symbolic Dynamics: One-sided, Two-sided, and Countable State Markov Shifts*. Springer-Verlag, Berlin 1998.

[13] Petr Kůrka. Cellular automata with vanishing particles. *Fund. Inform.*, **58**(3–4) (2003), 203–221.

[14] Petr Kůrka. On the measure attractor of a cellular automaton. *Discr. & Cont. Dyn. Syst.* (suppl.) (2005), 524–535.

[15] Petr Kůrka and Alejandro Maass. Limit sets of cellular automata associated to probability measures. *J. Statist. Phys.* **100**(5–6) (2000), 1031–1047.

[16] Petr Kůrka and Alejandro Maass. Stability of subshifts in cellular automata. *Fund. Inform.* **52**(1–3) (2002), 143–155. Special issue on cellular automata.

[17] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge UP, New York, 1995.

[18] M. Pivato. Algebraic invariants for crystallographic defects in cellular automata. *Ergodic Theory & Dynamical Systems* **27**(1) (February 2007), 199–240.

[19] M. Pivato. Defect particle kinematics in one-dimensional cellular automata. *Theoretical Computer Science* **377**(1–3) (May 2007), 205–228.

[20] M. Pivato. Spectral domain boundaries cellular automata. *Fundamenta Informaticae* **78**(3) (2007), 417–447.