The one-dimensional *Exactly 1* cellular automaton: replication, periodicity, and chaos from finite seeds

JANKO GRAVNER
Mathematics Department
University of California
Davis, CA 95616
e-mail: `gravner@math.ucdavis.edu`

DAVID GRIFFEATH
Department of Mathematics
University of Wisconsin
Madison, WI 53706
e-mail: `griffeat@math.wisc.edu`

(Final revision, December 2010)

Abstract. In the *Exactly 1* cellular automaton (also known as *Rule 22*), every site of the one-dimensional lattice is either in state 0 or in state 1, and a synchronous update rule dictates that a site is in state 1 next time if and only if it sees a single 1 in its three-site neighborhood at the current time. We analyze this rule started from finite seeds, i.e., those initial configurations that have only finitely many 1's. Three qualitatively different types of evolution are observed: replication, periodicity, and chaos. We focus on rigorous results, assisted by algorithmic searches, for the first two behaviors. In particular, we explain why replication is observed so frequently and present a method for collecting the smallest periodic seeds. Some empirical observations about chaotic seeds are also presented.

1   Introduction

The one-dimensional *Exactly 1* cellular automaton, also known as *Rule 22* [**Wol1**], is a dynamical system ξ_t whose state at each time t is an assignment of 0's and 1's to sites of the integer lattice $\mathbb{Z}$, i. e., $\xi_t \in \{0,1\}^{\mathbb{Z}}$. The update rule specifies that there is a 1 at site x at time $t+1$ if and only if there is a 1 at exactly one of the sites $x-1$, x, $x+1$ at time t. This is *not* an additive rule like the much more mathematically tractable automata for which the state next time is obtained by addition modulo 2 over some neighborhood. Among additive rules, which are often called *linear*, the prototype is *Rule 90*, also known as *Pascal's triangle modulo 2*, and defined in Section 2 as ℓ_t. Another instance is *Rule 150*, in which the addition is over the same three-site neighborhood that *Exactly 1* uses, and many others have been studied [**Wol1**, **Wil1**, **Wil2**, **FLM**].

The literature on *Exactly 1* is almost entirely empirical (e.g., [**Gra1**, **Gra2**, **Zab**, **Eis**, **MM**]), but a few rigorous results such as lower and upper bounds for the rule's spatial entropy may be found in [**Wol2**]. Even among numerical and simulation studies, surprisingly little attention has been devoted to the behavior of *Exactly 1* as a growth model starting from *seeds*, i.e., from initial configurations with finitely many 1's. Befitting the growth interpretation [**GG1**], we often refer to 1's as *occupied* sites and 0's as *empty* sites. To date, the evolution of the occupied set has been documented by space-time snapshots of the first few hundred updates starting from various simple initializations, with accompanying informal descriptions (e.g., [**Wol3**]). Our goal in this paper is to offer a more substantive mathematical investigation. Section 2 will lay the formal groundwork for our study.
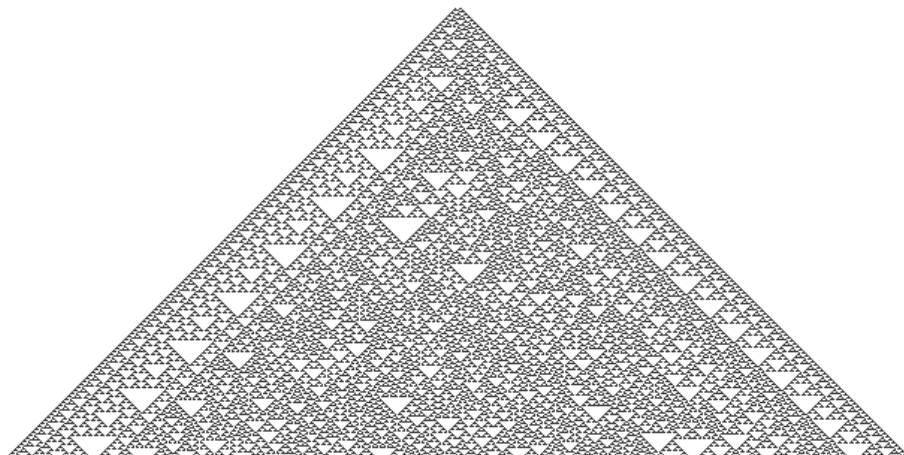
Let us begin by examining the space-time evolution of *Exactly 1* from judiciously chosen small seeds. These are discussed in some detail throughout the paper, and [**Gri2**] contains additional information.
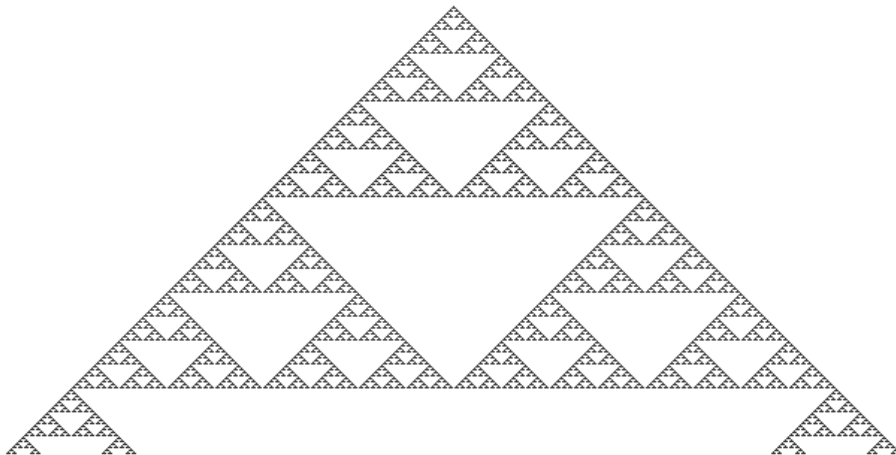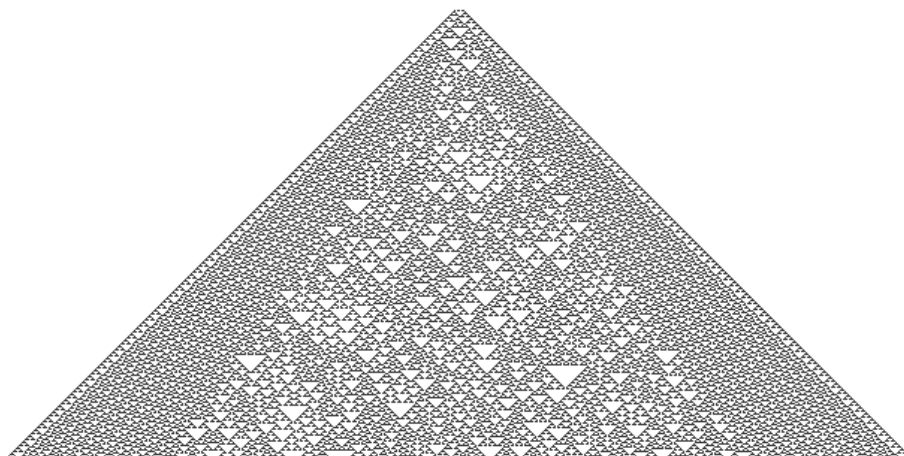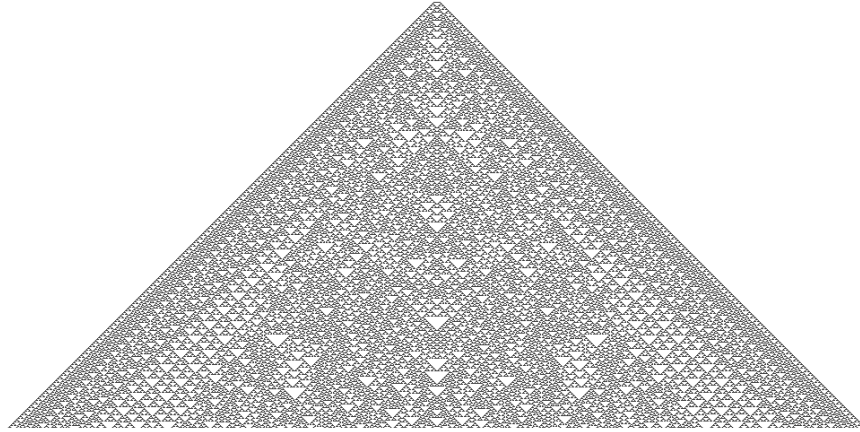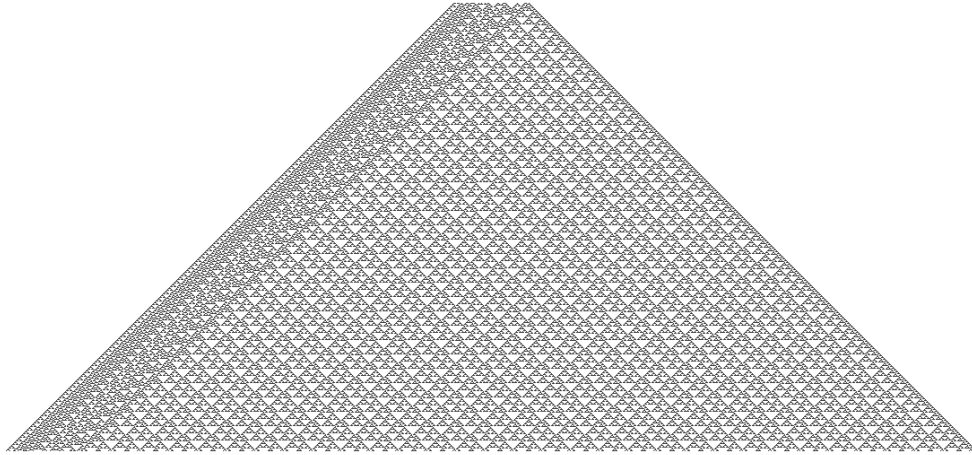

$1a.$
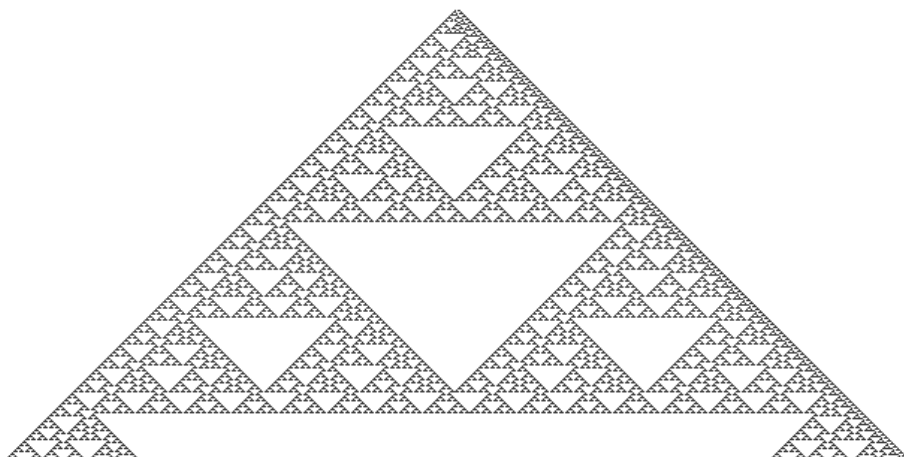
1b.

1c.

1d.

1e.

1f.

Fig. 1a–f**.** Six types of *Exactly 1* evolution. As is customary, the time axis is oriented *downward*.

Example 1. Fig. 1 displays the trajectories, up to time 299 in 1a–d and up to time 499 in 1e–f, from the following seeds, in order: (1a) 1, a maternal replicator; (1b) 1011, a fraternal replicator; (1c) 10011, a chaotic seed; (1d) 1011011, a chaotic seed with periodic wedges; and (1e) 11111111, another chaotic seed that develops substantial bands near the left and right boundaries with a well-defined periodic state. These bands do not continue to spread; however, there is a periodic seed (1f) of length 85, given below, with the same periodic state.

Seeds such as (1a) and (1b), which produce self-similar structure reminiscent of the additive case [**Wil1**, **Wil2**], are known as *replicators* [**Eva**, **Epp**, **BP**, **GG2**]. The replicating elements are identical in the former case, but slightly different in the latter (see Section 2 for details). To distinguish the two types, we use the twin analogy and call them *maternal* and *fraternal*, respectively. Note that our terminology identifies the seed with the attractor, i.e., we call *any* finite seed that gives rise to replicating behavior a replicator, whether or not the seed itself is reproduced by the dynamics. Similarly, in the sequel we call a seed periodic if it is in the domain of attraction of a periodic state, and the same convention applies to chaotic seeds.

In addition to Example $1a$–b, solid blocks of 1's of length up to 7 and many other small seeds are replicators; p. 263 in [**Wol3**] shows three additional examples, all maternal. Despite the essential nonlinearity of *Exactly 1*, replicators evidently occur more frequently than one might expect, especially among small seeds. For instance, we will present data indicating that proportion $1277/2048 \approx 0.62$ of seeds with length $\nu \leq 12$ are replicators. Section 4 will illuminate this abundance, and an apparent scarcity among seeds of longer length, by analyzing various structural characteristics of replicators in some detail.

Fig. $1c$ shows the evolution of a *chaotic* seed. Periodic patterns arise along the edges of the occupied region — this property of *Exactly 1 edge dynamics* will play a key role in our analysis (see also the proof of Theorem 3 in [**GG2**]) — but chaotic space-time structure fills out a central wedge. See p. 951 of [**Wol3**] for some empirical observations and Section 2 for a precise definition of a chaotic evolution. Experimental evidence suggests a characteristic density of 1's around 0.351 and other statistical properties in good agreement with the chaos generated by *Exactly 1* starting from a density p product measure, i.e., a random configuration in which the state at every site is 1 independently with probability $p \in (0, 1)$. Such infinite starting states were studied in [**Gra2**, **Zab**]. Despite the prevalence of replicators among seeds of short length ν, our study suggests that chaos predominates once ν is large.

A more complicated evolution arises from seed $(1d)$. Again, there is a central disordered wedge, but to each side is an additional wedge, spreading linearly in time, within which a space-time periodic pattern emerges. Both periodic patterns have temporal period $\tau = 52$, spatial period $\sigma = 104$, and density of 1's $\rho = 11/26 \approx 0.423$. It turns out that this $11/26$ pattern is very common starting from small seeds. For instance, among all cases with $\nu \leq 12$, the proportion with this feature exhibited (either on one side, or like $(1d)$ on both sides) by time $10,000$ is $512/2048 = 0.25$. This and other much less common periodic structures for *Exactly 1* will be another focus of our study.

Another variety of asymptotic growth, quite rare and particularly remarkable, arises by careful examination of the evolution starting from seed $(1e)$, a solid block of length 8. Fig. $1e$ shows its first 499 updates. As in Fig. $1d$, the central wedge is chaotic, but a periodic pattern emerges at the lower left and lower right. What is the edge configuration that is not part of this pattern, but is clearly instrumental in its maintenance? Careful inspection reveals that, starting at the left edge of the occupied region at time 220, the following spatial sequence of length 85 repeats every 16 updates:

$(1f)$ 1110111001000011100111111000010010001001100000000011000011110000000000010000000111001

The last evolution of Fig. 1 starts from this seed $(1f)$. Surprisingly, the same periodic pattern appears, with $\sigma = \tau = 32$ and $\rho = 19/64 \approx 0.297$, but now within the *entire* occupied region except for cells a uniformly bounded distance from the boundary. This *periodic* seed was discovered by Dean Hickerson [**Hic1**, **GG1**] using a random search algorithm based on the label tree we will describe in Section 5. He also noted that the seed has a length 83 predecessor:

10001000011101001100100111100001110000101111111010000001001111111111101011111010011

At present, this has the shortest length of any known periodic seed for *Exactly 1*.

In the context of dynamical systems, *Exactly 1* is thus a cellular automaton (CA) with many features of nonlinear ordinary differential equations: relatively simple linear-type attractors (replicators), periodic attractors, and a chaotic attractor (again, see Section 2 for a more detailed discussion). Significantly, all these are accessible from finite seeds without the need for infinite precision or another infinite amount of information. Thus entropy estimation is meaningful in assessing the size of the respective domains of attraction and we will devote much of the paper to this task (Sections 4, 6, and 7). Moreover, identification of periodic orbits has long been considered fundamental in understanding the evolution of a dynamical system (cf. [**ACEGP**]), so this is the other focus of our attention (Section 5). The largely empirical Section 8 is devoted to chaotic seeds. Throughout, our emphasis is on the (replicating, periodic, or chaotic) behavior of single trajectories rather than on the collective behavior from an ensemble of initial states (as is done for additive rules in, say, [**FLM**]). This we consider the proper perspective for investigation of growth phenomena. We also note that CA capable of evolution of all three types do not abound – apparently *Exactly 1* is by far the simplest instance, and the only other natural example we know is described in [**GG2**]. See [**Gra1**] for more about the special properties of *Exactly 1*.

There is a large literature about diffusively coupled maps on the one-dimensional lattice that addresses analytical theory and wide-ranging applications; here we simply mention the readable account in [**KT**]. *Exactly 1* fits well into this context since it can be represented as coupled iteration of the map $\phi : [0, 1] \to [0, 1]$, chosen to be any characteristic function of an interval that contains $1/3$ and is included in $(0, 2/3)$. The variables $y_n(x)$, $x \in \mathbb{Z}$, are updated to

$$y_{t+1}(x) = \frac{1}{3} \left(\phi(y_t(x-1)) + \phi(y_t(x)) + \phi(y_t(x+1)) \right).$$

The map ϕ is simple to the point of caricature and the uncoupled system reaches the quiescent state $y \equiv 0$ in two steps. By contrast, the diffusively coupled maps are evidently capable of creating quite a rich structure from bounded perturbations of the quiescent equilibrium: transmission and replication of information, periodic synchronization, and chaos. Whether still more complicated behavior is possible, such as solitons moving on a synchronized background, is perhaps the most intriguing open question about this rule. Additional open problems are interspersed throughout our study and we conclude the paper with several more in Section 9.

Many of our results for *Exactly 1* were first conjectured through computer experimentation, using *MCell* [**Woj**] and *Golly* [**TR**]. We will maintain a collection of experiments in a format recognized by those two programs at [**Gri2**], so that the reader can experience this fascinating rule in action.

2 Preliminaries

A *configuration* is a finite, semi-infinite (extending to infinity only rightward), or doubly infinite sequence of 0's and 1's. In many cases it is immaterial where a configuration is placed on the lattice $\mathbb{Z}$. However, if important we will call the (finite or a semi-infinite) configuration's left endpoint its *placement*. This is particularly significant when a new configuration is built by

a few appropriately placed finite configuration. If unspecified, the placement is at the origin. Another convention is that the state of a site, when not given, is 0. As is customary we put a configuration η in a superscript to indicate that it is used as the initial configuration, e.g., ξ_t^η indicates that $\xi_0 = \eta$. The term *seed* will be reserved for *finite* initial configurations.

We will make extensive use of *Rule 90*, the additive dynamics ℓ_t on $\{0,1\}^{\mathbb{Z}}$ given for $t \geq 1$ by

$$\ell_t(x) = (\ell_{t-1}(x-1) + \ell_{t-1}(x+1)) \bmod 2 \ .$$

A *replicator rule* is a finite set $\mathcal{K}$ of finite configurations, called *replicating elements*, together with functions $\texttt{left}, \texttt{right} : \mathcal{K} \to \mathcal{K}$. A finite initial configuration ξ_0 is a *replicator* if, after a proper placement, there exist a replicator rule $(\mathcal{K}, \texttt{left}, \texttt{right})$, and $t_0 \geq 0, n_0 \geq 0$, so that the configurations ξ_t at times $t = t_0 + 2^{n_0}(k-1)$, $k = 1, 2, \ldots$, satisfy the following:

- for every k and x such that $\ell_k^{\{0\}}(x) = 1$ there is a replicating element, that is, a configuration $K_{k,x} \in \mathcal{K}$, placed at $2^{n_0}x$;

- all other states are 0's; and

- $K_{k,x} = \texttt{right}(K_{k-1,x-1})$ if $\ell_{k-1}^{\{0\}}(x-1) = 1$ and $\ell_{k-1}^{\{0\}}(x+1) = 0$, and $K_{k,x} = \texttt{left}(K_{k-1,x+1})$ if $\ell_{k-1}^{\{0\}}(x-1) = 0$ and $\ell_{k-1}^{\{0\}}(x+1) = 1$.

Thus all $K_{k,x}$ are determined by the initial pair $(K_{1,-1}, K_{1,1})$ and successive applications of $\texttt{left}$ and $\texttt{right}$. Heuristically, t_0 is the onset time, when the initial pair of replicating elements appears; by a suitable translation these two can be placed at $\pm 2^{n_0}$. Starting at t_0, *Exactly 1* emulates *Rule 90* in time steps of length 2^{n_0}: any replicating element branches into two successors, one or both of which may be annihilated by contact with other successors. Fig. 2 depicts an illustrative example. We remark that $\mathcal{K}$ is not unique: all conditions are satisfied if each replicating element is replaced by its *Exactly 1* successor, and t_0 and n_0 suitably enlarged.
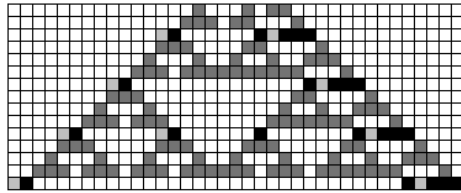
Fig. 2. Evolution of the replicator 10001011 up to time 14. Highlighted in black are 1's at the onset time $t_0 = 2$ and thereafter at three multiples of the replication step $2^{n_0} = 4$. At all these times, 0's which are part of a replicating element are light grey (and at other times 1's are dark grey). The set of replicating elements is $\mathcal{K} = \{1, 01, 10111\}$; the initial pair is $(01, 10111)$, placed at $\pm 2^{n_0} = \pm 4$; and $\texttt{left}$ and $\texttt{right}$ are identity functions except that $\texttt{left}(10111) = 1$.

We call ξ_0 a *maternal* replicator if there exists a configuration K so that any replicating element equals K, possibly with 0's appended at either end. Any replicator which is not maternal is *fraternal*. The distinction is important as fraternal replicators are a non-additive phenomenon — additive and quasiadditive rules discussed in Section 3 only create maternal ones.

For instance, a singleton (Example 1a) has $\mathcal{K} = \{1\}$, and $t_0 = 2$, $n_0 = 1$. The seed 11 is a maternal replicator with $\mathcal{K} = \{01, 1\}$, both `left` and `right` are identity functions, and $t_0 = n_0 = 1$. The simplest fraternal replicator is 10111 and has $\mathcal{K} = \{10111, 1\}$, with `left` and `right` identity functions, except that $\texttt{left}(10111) = 1$, and with $t_0 = 4$, $n_0 = 2$. (This example will be used as a building block for exponentially many fraternal replicators later.) The seed 1011 of Example 1b has three replicating elements:

$$K_1 = 010001000100010001000100000001, \quad K_2 = 1000100010001000100010000000010111,$$

and K_1' that equals K_1 without the leading 0. The only non-identity assignment of `left` and `right` is $\texttt{left}(K_2) = K_1'$. The rather long replicating elements necessitate a large time $t_0 = 46$ and $n_0 = 5$.

As in [**GG2**], one could formulate a condition to verify that an initial state is a replicator after only finitely many replications, but we will not do so here.

Assume that η is a (doubly infinite) configuration with *spatial period* σ, i.e., $\eta(x+\sigma) = \eta(x)$ for every x. Assume also that it is periodic with *temporal period* τ for *Exactly 1*, i.e., $\xi_{t+\tau}^{\eta} = \xi_t^{\eta}$ for all $t \geq 0$. Then we call η a *periodic solution*; we will assume that both periods σ and τ are minimal.

Further, we define a space-time *wedge* $W = W_{\alpha,\beta} = \{(x,t) \in \mathbb{Z}^2 : \alpha t \leq x \leq \beta t\}$. Here $-1 \leq \alpha < \beta \leq 1$. We call ξ_0 a *periodic seed*, attracted to a periodic solution η within the wedge $W_{\alpha,\beta}$, if for every $\epsilon > 0$ there exists a large enough time $T = T(\epsilon)$ so that whenever $t \geq T$ and $(\alpha + \epsilon)t \leq x \leq (\beta - \epsilon)t$,

$$\xi_t^{\xi_0}(x) = \xi_t^{\eta}(x).$$

We call such a periodic solution η a *periodic attractor*, and of course consider any spatial translation of η, or any of its temporal iterations ξ_t^{η}, to be the same attractor. When the seed is simply called *periodic* without specifying the wedge, we presuppose some wedge $W_{\alpha,\beta}$ with $\alpha < 0 < \beta$. For all known instances the wedge will in fact be maximal: $W_{-1,1}$ (cf. Problem 5 of Section 9).

Next, we give the precise meaning of a *chaotic* seed. We feel a rigorous definition is called for (see [**FG**]) although we are not able to prove that any seed is chaotic (but see Section 8). Ultimately, this issue boils down to what is meant for a single deterministic configuration $\Xi \in \{0,1\}^{\mathbb{Z}^2}$ to be chaotic within a wedge. For a finite set $A \subset \mathbb{Z}^2$ and wedge $W = W_{\alpha,\beta}$, define

$$\rho_A^n = \frac{\#\{w \in \mathbb{Z}^2 : w + A \subset W \cap \{t \leq n\} \text{ and } \Xi \equiv 1 \text{ on } w + A\}}{((2\beta)^{-1} - (2\alpha)^{-1}) \cdot n^2}.$$

(Here $\#$ denotes cardinality.) Note that the denominator is asymptotically the number of sites in $W \cap \{t \leq n\}$ and that, if A is a single site, the numerator simply counts the number of occupied sites Ξ has in $W \cap \{t \leq n\}$. We call Ξ *chaotic within* W if

- $\rho_A = \lim_{n\to\infty} \rho_A^n$ exists for every A;

- this limit is not 0 or 1 when A is a singleton; and

- correlations decay, i.e., there exists a constant $\gamma > 0$ such that, for every pair of finite sets A and B,

$$|\rho_{A \cup (z+B)} - \rho_A \cdot \rho_B| \leq ||z||^{-\gamma}$$

  for large $||z||$.

Finally, the initial seed ξ_0 is *chaotic within* $W_{\alpha,\beta}$ if the same holds for the induced space-time configuration given by $\Xi(x,t) = \xi_t(x)$. As in the periodic case, we will simply call a seed *chaotic* if this is true for some $\alpha < 0 < \beta$, but in contrast with the periodic case it appears that such a wedge is often not maximal (see Section 8). By inclusion-exclusion the existence of a density and correlation decay can be extended to any configuration of 0's and 1's on sets A and B. Another straightforward exercise is to show that neither a replicator nor a periodic seed is chaotic: the former has $\rho_{\{0\}} = 0$ and the latter lacks correlation decay. An experimental method that looks for a signature property of chaos starting from product measure is presented in [**MM**].

Call two chaotic configurations *equivalent* if their respective densities ρ_A agree for all finite sets A. It is an open question whether all chaotic configurations for *Exactly 1* are equivalent (in which case *Exactly 1* could be said to have a unique chaotic attractor); we see no evidence to the contrary in our empirical data.

The power γ measures the degree to which a configuration is chaotic. Empirical results [**Gra2**] strongly suggest that γ can be chosen to be arbitrarily large, i.e., that correlations decay faster than algebraically. It is conjectured in [**Gra2**] that this decay is slower than exponential, but due to problems approximating correlations for large $||z||$ we consider the evidence inconclusive (see also [**Eis**]).

Conceptually, a very convenient tool for our analysis of *Exactly 1* is the (*left*) *edge dynamics*. This amounts to a shift in perspective, whereby one views the evolving configuration from its left edge, which is thus fixed at the origin. More precisely, the edge dynamics is the *Exactly 1* CA with the neighborhood of x changed to $\{x-2, x-1, x\}$ and the additional convention that the origin is initially the leftmost occupied site. Clearly the edge dynamics preserves the latter property for all time. Furthermore, for any $k \geq 1$, the edge dynamics on its leftmost k sites is closed, i.e., unaffected by other sites, and is therefore eventually periodic. We made modest use of these periodic orbits in [**GG2**], and will devote much more attention to them in Section 5.

3 Quasiadditive initial states and replication

In this section we prove that, for a class of initial states, *Exactly 1* emulates Rule 146 of E. Jen [**Jen2**, **Jen3**]. It follows that such states, and all seeds in their basin of attraction, are maternal replicators. Most of the arguments here are very similar to those in [**Jen1**], [**Jen2**] and in Section 2 of [**GG2**], but we give some details to adapt the relevant notions to the present context, and to correct and clarify a few minor misstatements in [**Jen2**].

Rule 146 is a one dimensional CA with the same neighborhood of x as *Exactly 1*, and evolution b_t such that $b_t(x) = 1$ iff either

- $b_{t-1}(x-1) = b_{t-1}(x) = b_{t-1}(x+1) = 1$, or

- $b_{t-1}(x) = 0$ and exactly one of $b_{t-1}(x-1)$, $b_{t-1}(x+1)$ is 1.

We begin by dividing a one-dimensional configuration in $\{0,1\}^{\mathbb{Z}}$ into maximal contiguous 1-*blocks* and 0-*blocks* of single type, 1 and 0 respectively. *Odd* and *even* blocks consist of odd and even numbers of sites, respectively, and we also call any infinite 0-block odd.

An initial state ξ_0 is *quasiadditive* if each 1 is isolated (i.e., each 1-block has length 1) and each 0-block is odd. For such a state we will, without loss of generality, assume that the isolated 1's are at even sites of $\mathbb{Z}$ and define its *reduction* to have 1 at x iff ξ_0 has a 1 at $2x$.

Lemma 3.1. *Assume that the initial state ξ_0 for Exactly 1 is quasiadditive and let the initial state b_0 for Rule 146 be its reduction. Then ξ_t is quasiadditive at all even times and*

$$\xi_{2t}(2x) = b_t(x).$$

Proof. A direct verification, by checking all cases at time 2, and then by induction. □

The next four lemmas establish helpful properties of Jen's *Rule 146*. Note that no reduction is used in any of them.

Lemma 3.2. *If $b_0 = \ell_0$ is quasiadditive, then $b_t = \ell_t$, and is quasiadditive, for all t.*

Proof. Direct verification. □

Following [**Jen2**], we call any 0-block of even length and any 1-block of length at least 2 *irregular*.

Lemma 3.3. *If the number of irregular blocks is finite, then it is nonincreasing in time.*

Proof. The irregular blocks can be tracked through time so that each has a well-defined successor unless it disappears on its own or gets annihilated by another irregular block. The rule of succession is as follows. (See also Section 4 of [**Jen2**].)

An irregular block of either type of size at least 3 shrinks by 1 from each of its ends and the result is its successor. A 0-block of size 2 turns into a 1-block of size 2, its successor. A 1-block of size 2 changes to a pair of 0's, part of some block of 0's which, if irregular, is its successor. To find the right edge of the block of 0's, move rightward and find the first 1 with either two 0's or two 1's on its right. The position of this 1 is the right edge, and the left edge is defined analogously. It is easy to check that the resulting 0-block is even iff the number of 11 blocks it succeeds is odd — in this sense, the 11 blocks annihilate in pairs and constitute the only mechanism by which the number of *even* irregular blocks decreases. An odd 1-block of length 3 or more dies by shrinking to a singleton. □

*Jen's conjugacy map $J(\Lambda)$ on a configuration $\Lambda \in \{0,1\}^{\mathbb{Z}}$ inserts a 0 into any even 0-block, and replaces any 1-block of size $k \geq 2$ with a configuration of length $k+1$ consisting of a 0-block of size $k-1$ flanked by two 1's.

Lemma 3.4. *Assume that the number of irregular blocks is finite and constant on some time interval $[t_1, t_2]$. Let $\Lambda = b_{t_1}$ and restart the dynamics at t_1 (i.e., translate time so that $t_1 = 0$). Then*

$$J(b_t^\Lambda) = b_t^{J(\Lambda)} = \ell_t^{J(\Lambda)}, \qquad t \in [t_1, t_2].$$

Proof. See Theorem 1 in [**Jen2**]. $\square$

Lemma 3.5. *Assume that Λ is a finite seed for b_t. Then there exists a finite set $\Lambda_0 = \Lambda_0(\Lambda)$ comprised of only regular blocks, and a $t_0 = t_0(\Lambda) \geq 0$ and $n_0 = n_0(\Lambda) \geq 0$, so that, after a proper placement of Λ, the following holds.*

At time t_0, the configuration $b_{t_0}^\Lambda$ consists of two disjoint copies of Λ_0 placed at sites at a distance D_0. The number D_0 is either 2^{n_0} or $2^{n_0} - 1$, depending on whether the number of even blocks in Λ is, respectively, even or odd.

When $D_0 = 2^{n_0}$, the configuration b_t, at times $t = t_0 + 2^{n_0}(k-1)$, $k = 1, 2, \ldots$, consists of disjoint copies of Λ_0 placed at each $2^{n_0}x$ such that $\ell_k^{\{0\}}(x) = 1$. When $D_0 = 2^{n_0} - 1$, the only change is that for $x > 0$ the copies are shifted one unit to the left, i.e., placed at $2^{n_0}x - 1$.

Proof. From some time t_1 on, the number of irregular blocks must be constant — in particular, there is no 1-block of length 3 or more. Without loss of generality, we may assume that this is already true initially: we use the seed Λ_1, given by $b_{t_1}^\Lambda$, as the new initial seed. Thus, Λ_0, t_0 and n_0 will depend on Λ_1. Also, Λ is placed so that the leftmost 1 of Λ_1 is at the origin.

Fix a 11 block with its placement and time given by a pair (x_0, t_0), and track the placement of its successor 11 blocks, say $(x_1, t_1), (x_2, t_2), \ldots$. It is not hard to check (cf. the proof of Lemma 2.4 in [**GG2**]) that for $k \geq 0$,

(3.1) $-(t_{k+1} - t_k) < x_{k+1} - x_k < t_{k+1} - t_k.$

Now, replication properties of ℓ_t guarantee that at sufficiently large times of the form $t = 2^n$, $\ell_t^{\Lambda_1}$ creates two copies of Λ_1, separated by a long interval of 0. At these times, by Lemma 3.4, $b_t^{\Lambda_1}$ consists of two finite sets, each at most twice the length of Λ_1, which are separated by a long interval of 0's. Moreover, the rightmost 1 in $b_t^{\Lambda_1}$ moves at the *speed of light*, i.e., one step to the right at each time unit, and so by (3.1) any 11 block must have the described long middle interval as its successor. Therefore, this interval is the sole remaining irregular block in the event that any remains.

As the number of irregular blocks was assumed constant, it can only be 0 or 1 initially (that is, at the original time t_1). Therefore, by the previous paragraph and Lemma 3.4, there exists an n_0 so that at times $t = 2^n$, $n \geq n_0$, $b_t^{\Lambda_1}$ consists of two identical configurations, separated by an interval of 0's. These configurations are placed at these times either at distance 2^n or at distance $2^n - 1$, and are either copies of Λ_1 or copies of $J(\Lambda_1)$. Accordingly, Λ_0 is defined to be either Λ_1 or $J(\Lambda_1)$, and $t_0 = t_1 + 2^{n_0}$. Note, for use in the next paragraph, that from time 2^{n_0} to time $2 \cdot 2^{n_0}$, the two copies of Λ_0 reproduce at the two ends, but annihilate in the middle, creating a longer interval of 0's.

If Λ_1 has no irregular blocks, then the dynamics is ℓ_t thereafter and additivity easily finishes the proof. Otherwise, we look at the dynamics $b_t^{\Lambda_1}$ between times 2^n and 2^{n+1}, $n \geq n_0$. At time 2^n, we have two copies of Λ_0, which evolve independently (by the speed of light) and additively (as b_t, started from Λ_0, acts as ℓ_t) up until time $2^{n+1} - 2^{n_0}$. From that time to time 2^{n+1}, only the two copies of Λ_0 closest to the middle (and placed at distance $2^{n_0}-1$) interact, and annihilate each other. The remaining copies of Λ_0 interact additively in this time interval, therefore the configuration at time 2^{n+1} again consists of two copies of Λ_0. $\qquad\square$

More precise results on the dynamics of irregular blocks remain elusive, although [**KN**] proves that a single irregular block embedded in a suitably chosen initial product measure performs an unbiased random walk. The main result of this section now follows immediately.

Theorem 1. *Every quasiadditive seed is a maternal replicator, with at most two replicating elements.*

We will denote the set of all quasiadditive seeds by $\mathcal{S}_q$. The *additive* set $\mathcal{S}_a \subset \mathcal{S}_q$ consists of all seeds with isolated 1's and 0-blocks of length $3 \bmod 4$. By Lemmas 3.1 and 3.2, *Exactly 1* started from a seed in $\mathcal{S}_a$ has its state in $\mathcal{S}_a$ at all even times and, restricted to these times, evolves equivalently to the additive dynamics ℓ_t. We also define $\mathcal{S}_7 \subset \mathcal{S}_a$ to be the set of configurations satisfying the additional restriction that 0-blocks have length $7 \bmod 8$. This last set will be used in Sections 6 and 7.

4 Entropy of replicators

Let $r_m(n)$ and $r_f(n)$ be the number of configurations on the interval $[0, 1, \ldots, n+1]$ that begin and end with 1 (i.e., with length $\nu = n+2$), and are maternal and fraternal replicators, respectively. For completeness, we also introduce the total replicator counts $r(n) = r_m(n) + r_f(n)$.

A principal objective in this section is to obtain lower bounds on various *entropies* that indicate the rate of decrease in the proportion of replicators among all seeds of a given length. As we shall see, parity plays a significant role, so we introduce maternal entropies

$$h_m^o = \liminf_{\text{odd } n} n^{-1} \log_2 r_m(n),$$
$$h_m^e = \liminf_{\text{even } n} n^{-1} \log_2 r_m(n),$$

and define corresponding fraternal entropies h_f^o, h_f^e, and total entropies h^o, h^e analogously.

Spatio-temporal entropies were central to early CA research (e.g., [**Wol1**]), but here the term has a more straightforward meaning, as the measure of the amount of choice, per bit, one has in constructing a long initial set with a desired property. To get a feeling for the above replication entropies, we have collected data, by both exhaustive search and simulation, on seeds with $n \leq 41$. Up to $n = 13$, we tally all seeds that exhibit the "mass extinction" signature of replication by a time cutoff: the evolution runs to time $t = 500\,n$, and once $t > 100\,n$ we check whether the density between the extreme 1's changes by at least a factor of 10, from above 0.6 to

below 0.06. (In all replicators we have seen, configurations that precede two replicating elements separated by 0's are close to $\ldots 011101110111\ldots$, thus with density about 0.75; so 0.6 is chosen somewhat smaller.) This provides a lower bound on the total number of replicators, but we cannot rule out a first replication later on. Counts for $n \geq 14$ use Monte Carlo: replicators are searched at random until 200 ($n \leq 23$ or $n = 29$) or 5 (remaining $n \leq 41$) are found. The results are shown in Table 1.

n	$r(n)$	$r_m(n)$	$r_f(n)$
1	2	2	0
2	4	1	3
3	6	4	2
4	8	3	5
5	18	12	6
6	17	6	11
7	47	31	16
8	39	20	19
9	150	87	63
10	118	51	67
11	517	252	265
12	349	179	170
13	1484	733	751

n	no. trials	no. repl.	no. mat.	no. frat.
14	2846	200	76	124
15	1713	200	99	101
16	4198	200	94	106
17	2151	200	85	115
18	5311	200	72	128
19	2395	200	101	99
20	6773	200	84	116
21	3405	200	82	118
22	7851	200	69	131
23	4295	200	83	117
24	223	5	3	2
25	223	5	4	1
26	333	5	0	5
27	100	5	2	3
28	440	5	3	2
29	9195	200	81	119
30	435	5	1	4
31	291	5	2	3
32	777	5	0	5
33	160	5	3	2
34	1137	5	2	3
35	440	5	3	2
36	1183	5	2	3
37	280	5	2	3
38	1753	5	1	4
39	540	5	2	3
40	1443	5	2	3
41	1641	5	2	3

Table 1. Replicator counts for small seeds.

While replicators are well represented in seeds with small length n, their proportion decreases rapidly after $n = 30$. Finding even one replicator at random past $n = 50$ would seem to require a great many tries. From admittedly very small values of n, all the h-values seem fairly close to 0.8. Both the data and our rigorous lower bounds suggest asymptotically lower entropies for even n. By contrast, it appears possible that maternal and fraternal entropies are equal, that is, that $h_m^o = h_f^o$ and $h_m^e = h_f^e$. Rigorous methods to investigate such questions remain elusive.

Let us now turn to some rigorous results. In light of Theorem 1, configurations in $\mathcal{S}_q$, which can be counted easily, give $h_m^o \geq \frac{1}{2}$. By identifying the immediate predecessors of such states,

this bound is improved to $\log_2 \phi$, where ϕ denotes the golden ratio, so $h_m^o \geq 0.694$. Recursive estimation of the frequency of seeds that lead to $\mathcal{S}_q$ after j steps, up to $j = 8$, yields the following improved lower bound on the entropy for odd maternal replicators. A judicious modification at one edge of the seeds attracted to $\mathcal{S}_q$ then yields the same result for odd fraternal replicators. Trivially, the common value also bounds the total replicator entropy for seeds of odd length.

Theorem 2. *The entropies h_m^o, h_f^o, and h^o are all bounded below by* 0.7555.

Proof. (*The odd maternal case*) We count predecessors of quasiadditive configurations, i.e., configurations that lead to $\mathcal{S}_q$ after one update, and in general those that accomplish this in j steps. For this purpose we use *j-triples*, which are introduced next.

The j-triples are defined recursively: a $(j+1)$-triple is a scheme $\begin{smallmatrix} a & b \\ & c \end{smallmatrix}$, where $a, b \in \{0, 1\}$ and c is a j-triple. Here b is the *type* of the triple. Additionally, a triple cannot have $a = b = 1$ and c of type 1. For $j = 0$, the triples are misnomers – there are three of them, ordered as given here: two 0-types, which we denote by 0_o and 0_e, and a 1-type, which we denote simply as 1.

The 0-triples are states needed to count quasiadditive configurations. Namely, we need to count the number of paths from 1 to 1 in a directed graph with oriented connections $1 \rightarrow 0_o$, $0_o \rightarrow 0_e$, $0_o \rightarrow 1$, $0_e \rightarrow 0_o$. (Here, 0_o and 0_e represents 0 sites which are at odd and even positions within a 0-block, respectively.)

Since we wish to represent the j-step predecessors as paths through j-triples, we are forced to recursively list transitions for $(j+1)$-configurations as follows:

$$
\begin{array}{cc} 0 & 0 \\ & \underline{0} \end{array} \rightarrow \begin{array}{cc} 0 & 0 \\ & c \end{array}
$$

$$
\begin{array}{cc} 0 & 0 \\ & \underline{1} \end{array} \rightarrow \begin{array}{cc} 0 & 1 \\ & c \end{array}
$$

$$
\begin{array}{cc} 1 & 0 \\ & \underline{0} \end{array} \rightarrow \begin{array}{cc} 0 & 1 \\ & c \end{array}
$$

$$
\begin{array}{cc} 1 & 0 \\ & \underline{1} \end{array} \rightarrow \begin{array}{cc} 0 & 0 \\ & c \end{array}
$$

$$
\begin{array}{cc} 0 & 1 \\ & \underline{0} \end{array} \rightarrow \begin{array}{cc} 1 & 1 \\ & \underline{0} \end{array}
$$

$$
\begin{array}{cc} 0 & 1 \\ & \underline{1} \end{array} \rightarrow \begin{array}{cc} 1 & 0 \\ & c \end{array}
$$

$$
\begin{array}{cc} 1 & 1 \\ & \underline{0} \end{array} \rightarrow \begin{array}{cc} 1 & 1 \\ & \underline{0} \end{array} \qquad \begin{array}{cc} 1 & 1 \\ & \underline{0} \end{array} \rightarrow \begin{array}{cc} 1 & 0 \\ & c \end{array}
$$

The underlined sites represent any j-triple of the corresponding type, possibly different on the left and right side of the transition, and c is an arbitrary j-triple. All transitions are given by the possible j-triple transitions (satisfying the restriction on type). We order $(j+1)$-triples as in the above transition list, assuming that j-triples are already ordered.

The above rules induces a recursive construction of the transition matrix. Let n_0 and n_1 be the number of j-triples of the types 0 and 1, and n_0', n_1' the two numbers for $(j+1)$-triples. Then $n_0' = 2n_0 + 2n_1$ and $n_1' = 2n_0 + n_1$, with $n_0 = 2$, $n_1 = 1$ for $j = 0$. Assume the j-triples transition matrix is given by blocks, with submatrix dimensions in the superscript, and subscripts reflecting the types,

$$M = \begin{bmatrix} M_{00}^{n_0 \times n_0} & M_{01}^{n_0 \times n_1} \\ M_{10}^{n_1 \times n_0} & M_{11}^{n_1 \times n_1} \end{bmatrix}$$

Then the $(j+1)$-tuples matrix M' is given by the same scheme and the submatrices

$$M_{00}' = \begin{bmatrix} M_{00}^{n_0 \times n_0} & M_{01}^{n_0 \times n_1} & 0^{n_0 \times n_0} & 0^{n_0 \times n_1} \\ 0^{n_1 \times n_0} & 0^{n_1 \times n_1} & 0^{n_1 \times n_0} & 0^{n_1 \times n_1} \\ 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & 0^{n_0 \times n_0} & 0^{n_0 \times n_1} \\ M_{10}^{n_1 \times n_0} & M_{11}^{n_1 \times n_1} & 0^{n_1 \times n_0} & 0^{n_1 \times n_1} \end{bmatrix} \quad M_{01}' = \begin{bmatrix} 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & 0^{n_0 \times n_0} \\ M_{10}^{n_1 \times n_0} & M_{11}^{n_1 \times n_1} & 0^{n_1 \times n_0} \\ M_{00}^{n_0 \times n_0} & M_{01}^{n_0 \times n_1} & 0^{n_0 \times n_0} \\ 0^{n_1 \times n_0} & 0^{n_1 \times n_1} & 0^{n_1 \times n_0} \end{bmatrix}$$

$$M_{10}' = \begin{bmatrix} 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & 0^{n_0 \times n_0} & 0^{n_0 \times n_1} \\ 0^{n_1 \times n_0} & 0^{n_1 \times n_1} & M_{10}^{n_1 \times n_0} & M_{11}^{n_1 \times n_1} \\ 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & M_{00}^{n_0 \times n_0} & M_{01}^{n_0 \times n_1} \end{bmatrix} \quad M_{11}' = \begin{bmatrix} 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & M_{00}^{n_0 \times n_0} \\ 0^{n_1 \times n_0} & 0^{n_1 \times n_1} & 0^{n_1 \times n_0} \\ 0^{n_0 \times n_0} & 0^{n_0 \times n_1} & M_{00}^{n_0 \times n_0} \end{bmatrix}$$

For $j = 0$,

$$M_{00} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad M_{01} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$M_{10} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad M_{11} = \begin{bmatrix} 0 \end{bmatrix}$$

This algorithm results in the following estimates, computed by MATLAB's sparse matrix routines, with $\lambda_{\max}$ the maximum eigenvalue of M and the last column rounded down. The final number is featured in the statement of the theorem.

j	$n_0 + n_1$	$\lambda_{\max}$	$\log_2 \lambda_{\max}$
0	3	$\sqrt{2}$	0.5
1	11	$(1 + \sqrt{5})/2$	0.6942
2	39	1.6226	0.6983
3	139	1.6295	0.7044
4	495	1.6598	0.7309
5	1763	1.6562	0.7278
6	6279	1.6746	0.7438
7	22363	1.6717	0.7413
8	79647	1.6883	0.7555

Table 2. Entropy estimates for $\mathcal{S}_q$ based on j-triples.

The matrices M are periodic with period 2, so they have eigenvalues $\pm\lambda_{\max}$. Thus it is crucial to note that the path between extreme 1's takes an even number of steps.

Another important remark concerns nonmonotonicity of the estimates in Table 1. This is the result of non-surjectivity of *Exactly 1*. For instance, it is easy to check that any configuration containing the string 10101001 cannot have a predecessor, so it is known as an *orphan* (or *garden*

of Eden) in the CA literature. Therefore the proportion of non-orphan configurations converges to 0 exponentially fast. Evidently, among the $j = 4$ and $j = 6$ predecessors there are many orphans, balancing the branching effect due to increased j. For some experimental results on the prevalence on orphans, and more on general methods for computing the set of predecessors of a configuration, see [**DG**]. On the other hand, *Exactly 1* preserves $\mathcal{S}_q$ after two time steps, so the table entries must increase in steps of size 2. Finally, we remark that our j-triple method is a sort of reverse counterpart to the technique in [**Wol1**] for computing spatial entropy.   □

Proof. (*The odd fraternal case*) We claim that any seed of the form

$$11101[7 \text{ 0's}]1[19 \text{ 0's}]M,$$

where M is any maternal replicator that leads to S_q after 8 updates, is a fraternal replicator. Consult Fig. 4a for a representative example of the construction.

We begin with the following observation, which will also be applied later in constructing exponential families of periodic and chaotic seeds.

Lemma 4.1. *Let R be a finite configuration, with its leftmost 1 placed at 0. Append a configuration A on its left to get AR. Suppose the Exactly 1 evolutions from R and AR are identical at all space-time sites (x, t) with $x + t \geq d$. Now let R' be any configuration that agrees with R at sites x such that $x \leq d + 1$. Then the evolution of AR' agrees with that of AR for $x + t \leq d + 1$, and is identical to that of R' for $x + t \geq d$.*

Proof. Since $R = R'$ for $x \leq d + 1$, by the speed of light both equalities $\xi_t^{AR}(x) = \xi_t^{AR'}(x)$ and $\xi_t^R(x) = \xi_t^{R'}(x)$ hold for $x + t \leq d + 1$. As we also assume that $\xi_t^{AR}(x) = \xi_t^R(x)$ for $x + t \geq d$, it follows that $\xi_t^{AR'}(x) = \xi_t^{R'}(x)$ when $x + t \in \{d, d+1\}$. To finish the proof, we show by induction that this last equation in fact holds whenever $x + t \geq d$. But for any site (x, t) with $x + t \geq d + 2$, $\xi_t(x)$ is decided by the states of sites $(x, t-1)$ with $x \geq d + 2 - t - 1$, or equivalently, $x + t - 1 \geq d$. And for these, $\xi_{t-1}^{AR'}(x) = \xi_{t-1}^{R'}(x)$ by the inductive hypothesis.   □

Apply the lemma with R a singleton and A a block of 3 1's followed by a single 0. Thus, AR is the seed 11101 with rightmost 1 at the origin. We claim the hypothesis is satisfied with $d = 4$. This is evident from Fig. 3, which shows the edge dynamics for R and AR. Note that the two are identical in the three rightmost columns shown and hence at all greater distances from the edge. This translates to equality for $x + t \geq 4$ for the *Exactly 1*.
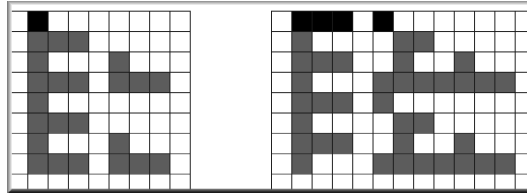

Fig. 3. The edge dynamics started from 1 (*left*) and 11101 (*right*) on, respectively, the 7 and 11 leftmost sites. Both have settled into a periodic cycle by time $t = 7$, and states of the rightmost 3 sites are identical up to this time, hence for all time.

Now consider R' of the form

$$1[7 \text{ 0's}]1[19 \text{ 0's}]M,$$

with M as in the claim. By inspection, R' agrees with R for $x \leq 5$. The configuration after 8 updates has the form

$$1[7 \text{ 0's}]1[7 \text{ 0's}]1[7 \text{ 0's}]1[3 \text{ 0's}]M',$$

where $M' \in S_q$. Note that this entire configuration belongs to S_q, so it is a maternal replicator. By Lemma 4.1, the evolution of seed AR' of the claim agrees with that of R' at all sites 6 or more cells from the left edge, whereas sites at most 5 cells from the edge evolve as in Fig. 3. It follows that the left replicating configuration of AR' is obtained from the left maternal pattern of R' by attaching 1110, and is therefore a fraternal replicator. For large n there are, for some constant $c > 0$, at least $c \cdot 2^{0.7555\,n}$ distinct choices of M with length $n + 2 - 30$ which thus make AR' of length $n + 2$. Therefore $h_f^o \geq 0.7555$. $\qquad\square$
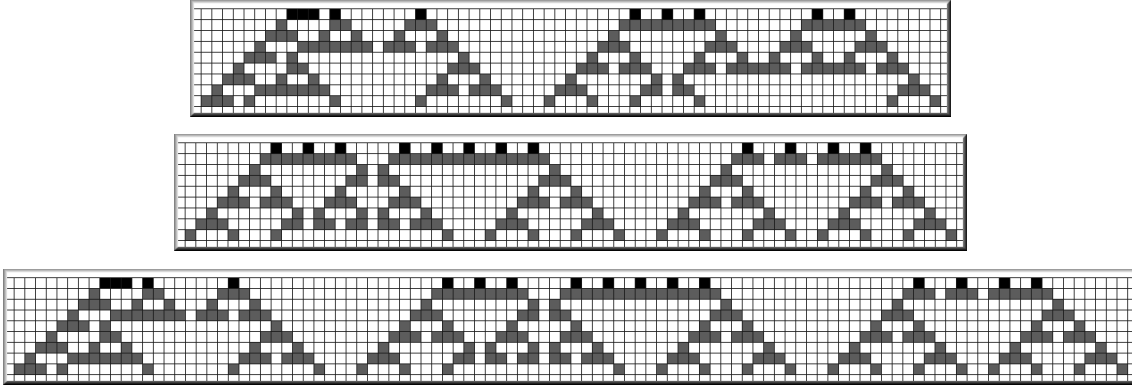

Fig. 4*a–c.* Examples of the construction of exponentially many, from top to bottom, (*a*) odd fraternal, (*b*) even maternal, and (*c*) even fraternal replicators.

For small values of n there are decidedly fewer even replicators in Table 1 than there are odd ones with one cell less. This reflects the key role of $\mathcal{S}_a$ and $\mathcal{S}_q$, both consisting entirely of odd seeds. One source of even replicators is *symmetry*: any symmetric seed with 1001 in the middle and quasiadditive structure elsewhere (isolated 1's and odd intervals of 0's) is also a maternal replicator. We omit the proof, as this approach cannot give us a better lower bound for h_m^e than $h^o/2 \leq 0.5$. Instead, we now describe another class of even replicators with substantially better entropy bounds.

Theorem 3. *The entropies h_m^e, h_f^e, and h^e are all bounded below by* 0.6730.

Proof. (*The even maternal case*) The key observation is the following simple conjugacy property.

Lemma 4.2. *Assume that the initial state $\xi_0 = A$ is additive, i.e., $A \in \mathcal{S}_a$. Replace a 1 in A with 1001. The so obtained seed is an even maternal replicator. In fact, its evolution is conjugate to the one obtained by replacing the same 1 with 10001, with insertion of an extra 0 being the conjugacy map.*

Proof. Let A' and A'' be the described perturbations, by substitution $1 \to 1001$ and $1 \to 10001$, respectively. We know from Section 3 that at even times $\xi_t^{A''}$ emulates additive dynamics at even times, and we assume t is even for the remainder of the proof. For $\xi_t^{A'}$, track the successor 0-block of the 00 pair, as in the proof of Lemma 3.3: at time 2 it is a 0-block, of length $2 \bmod 4$ and at least 6. This block proceeds to shrink by 4 at each successive even time until it reaches size 2 again, and the procedure repeats. By similarly tracking successors of the 000 block in $\xi_t^{A''}$ it is clear that the two agree except for an extra 0. □

We remark that the lemma does not generalize to other even perturbations such as a 100001 substitution or three 1001 substitutions.

Now consider any seed of the form

$$M^e = M[19 \text{ 0's}]100010001001,$$

where M is a maternal replicator that leads to $M' \in S_a$ after 8 updates. See Fig. 4b for an example. Then after the same 8 updates M^e leads to

$$M'00010001000100010010001000100010001.$$

An easy application of Lemma 4.2 shows that this is an even maternal replicator, so M^e is also. As in the odd maternal case, it therefore suffices to compute the entropy of the specified configurations M. Applying the recursive j-triple technique to $\mathcal{S}_a$, we obtain the bounds shown in Table 3, and the desired result for $j = 8$.

j	$n_0 + n_1$	$\lambda_{\max}$	$\log_2 \lambda_{\max}$
0	5	$2^{1/4}$	0.25
1	19	1.5127	0.5971
2	67	1.5361	0.6193
3	239	1.5341	0.6174
4	851	1.5640	0.6452
5	3031	1.5560	0.6378
6	3031	1.5560	0.6587
7	38447	1.5757	0.6560
8	136931	1.5944	0.6730

Table 3. Entropy estimates for $\mathcal{S}_a$ based of j-triples. □

Proof. (*The even fraternal case*) This construction combines those for the odd fraternal and even maternal cases. Given an even maternal replicator M^e of the form just described, append the same configuration on the left as in the proof of Theorem 2 to get

$$11101[7 \text{ 0's}]1[19 \text{ 0's}]M^e.$$

See Fig. 4c. Applying Lemma 4.1 as before, this is an even fraternal replicator. So we obtain the same lower bound for h_f^e as for h_m^e. $\square$

5   Periodic attractors

Let L be a finite configuration of length σ. Take another finite configuration H, and assume it has length h, with a 1 at the left end. Form the configuration HL^∞ by placing infinitely many copies of L to the right of H, next to each other. Starting from HL^∞, run the edge dynamics until time π. Assume that the starting configuration lies on a (semi-infinite) orbit, i.e., that there exist configurations $H_0, H_1, \ldots H_{\pi-1}$ of length h, and $L_0, L_1, \ldots L_{\pi-1}$ of length σ, so that $H_0 = H$, $L_0 = L$ and at time t the configuration generated by the edge dynamics is $H_{t \bmod \pi} L_{t \bmod \pi}^\infty$. Then we call H a *handle*, L a *link*, and the handle-link pair, denoted by $H + L$, has *temporal period* π and *spatial period* σ.

One could define a handle-link pair so that it merely repeats after π time steps. With suitable enlargement of the handle, and the same link, the condition from the previous paragraph is then satisfied.

We will fix the temporal period π in most of the discussion. Assume $H + L$ is such a handle-link pair, and start from HL^∞. For $k = 0, 1, \ldots$, and $m = 0, \ldots, \pi - 1$, let $\lambda_k[m] \in \{0, 1, 2, 3\}$ be the number whose binary representation is the state at sites $k - 1, k$ at time m. The *label at generation k* is the vector $\lambda_k = (\lambda_k[m]; m = 0, \ldots, \pi - 1)$. Note that here, and in the sequel, square brackets will be used to denote coordinates of labels. Also observe that the state at site k at time m is given by $\lambda_k[m] \bmod 2$.

Clearly, λ_k and the state at $k + 1$ at time 0, or any other time, together determine λ_{k+1}. Moreover, if the label contains at least one 3 (i.e., 11), say at position m, then the state at $k + 1$ at time $m + 1$ is 0. Therefore, for such λ_k, λ_{k+1} is uniquely determined. Moreover we may, and will, assume that $\lambda_h = \lambda_{h+\sigma}$. Indeed, if necessary we can make the handle longer by 2 sites and accordingly redefine L, without changing HL^∞ or σ. We again emphasize that spatial and temporal periods are assumed minimal, i.e., the link L cannot be divided into two or more identical pieces and HL^∞ does not repeat at some time which is a proper divisor of π.

The handle-link pair is *robust* if each label λ_k; $k = h, \ldots, h + \sigma - 1$, contains at least one 3. The reason for this terminology is our first result, which states that the handle part of a robust handle-link pair is attracted to the corresponding semi-infinite periodic configuration regardless of the states at sites to its right. For the edge dynamics, a robust handle-link pair is a kind of stable periodic attractor.

Proposition 5.1. *Assume that $H + L$ is a robust handle-link pair, and let A_0 be any seed that agrees with H on $(-\infty, h - 1]$. Then the configuration of the edge dynamics from A_0 at time t agrees with that started from HL^∞ on $(-\infty, h - 1 + \lfloor t/\pi \rfloor]$. Consequently, A_0 is periodic for Exactly 1 within the wedge $W_{-1, 1/\pi - 1}$.*

Proof. Assume the two states agree up to position k at time t. Then the two states agree at position $k + 1$ by time $t + \pi$ at the latest — the exact time of agreement depends on the distribution of 3's in λ_k. □

Note that while the spatial period σ is the same for both *Exactly 1* and its edge dynamics, the temporal periods are different. For instance, the periodic attractor in Example 1d has $\pi = 8$ and $\tau = 52$.

We call two handle-link pairs *matching* if their links are mirror images. The next proposition explains how a matching handle-link pair generates a periodic seed for *Exactly 1*. The two handles are unrelated and often very different. For any configuration A, we denote its mirror image by $\bar{A}$.

Proposition 5.2. *(a) For a matching handle-link pair, the spatial period σ divides 2π.*

(b) Assume that $\sigma = 2\pi$. Let $H + L$ and $H' + \bar{L}$ be matching handle-link pairs, and start Exactly 1 from $HL\bar{H}'$. At times $n\pi$, the resulting configuration is $HL^n\bar{H}'$ translated by $-n\pi$, so this seed is periodic.

We should emphasize that (a)–(b) are the only results in this section dealing with the *Exactly 1* rule directly; all the others are essentially about its edge dynamics. It follows from (b) that for the periodic seeds we will present, all with $\sigma = 2\pi$, the corresponding temporal period τ under *Exactly 1* is 2π. These and several other examples will be discussed from both the original and edge perspectives, as convenient.

Proof. To prove (a), note that the state of the link L at $(x + \pi) \bmod \sigma$ must equal that at $(x - \pi) \bmod \sigma$. So the state at x and agrees with those at $(x + 2\pi) \bmod \sigma$, and consequently also with those at $x + \gcd(\sigma, 2\pi) \bmod \sigma$.

For (b), let the handle lengths of H and H' be h and h', respectively. In time π, *Exactly 1* started from configuration HL^n translates this configuration by $-\pi$ onto $[-\pi, -\pi + h + n\pi]$. Also in time π, $H'\bar{L}^n$ is translated by $-\pi$ onto $[-\pi, -\pi + h' + n\pi]$. Therefore, starting from $HL^n\bar{H}'$ and using (a), at time π *Exactly 1* creates the configuration $HL^{n+1}\bar{H}'$ translated by $-\pi$. The proof is completed by induction. □

The rest of this section is devoted to search strategies for robust and matching handle-link pairs. Our key tool is the following *label dynamics*. We begin by listing a few facts that are easy to check. As we already know, any label $\lambda \in \{0, 1, 2, 3\}^\pi$ in generation k determines the label in the next generation if one of its coordinates is 3. If none of its coordinates is 3, and the combined number of 1's and 2's is even, then there are two possibilities for the next-generation label: λ^0 and λ^1. The former, called the 0-successor, is generated by state 0 at position 0, and

the latter, called the 1-successor, by state 1 at the same position. Finally, if there is no 3 and the combined number of 1's and 2's is odd, then there is no next label possible with temporal period π. However, if we embed period π labels in period 2π ones in the natural way, then we do get two successor labels. In particular, the following proposition holds. The proof repeats observations from this paragraph, but we feel it is worth doing so since they form the basis of our arguments in the sequel. Moreover, the simple proof uses no specific property of *Exactly 1*, so the conclusion holds for any edge dynamics with 2 states.

Proposition 5.3. *Temporal periods of handle-link pairs are powers of 2.*

Proof. Assume that the boundary dynamics repeats a configuration A after π steps, but not earlier. Append a 0 to the right end of this configuration. Write $0 \to 0$ if this 0 returns to 0 after π steps, and similarly for the other three possibilities. If $0 \to 1$ and $1 \to 1$, then the only appended state that yields a periodic configuration is 1, and the period remains π. Similarly if $1 \to 0$ and $0 \to 0$, except that the sole appended state is now 0. If $1 \to 1$ and $0 \to 0$, we can append both states and the period remains π. The final possibility is $1 \to 0$ and $0 \to 1$, when either appended state repeats after 2π time steps, but not earlier. $\square$

A natural object is thus the *label tree* associated with a given period π, which we will assume from now on is a power of 2. In this tree, every node has a label. The root label is the vector 1^π with all coordinates 1, and a node with label λ has a 0, 1, or 2 of successors, with their labels given by the rules in the paragraph before Proposition 5.3. Clearly, if the label tree is finite, then no handle-link pair exists for that π. Such is the case for $\pi = 1$ (the root has no successor), $\pi = 2$ (with a total of 7 nodes and 3 generations), and $\pi = 4$ (51 nodes, 11 generations). If a label at a node equals, up to rotation, a label at a node in its ancestral line, then the tree is infinite. This indeed happens for $\pi = 8$, and consequently for all larger π. Here, the label λ' is *a rotation* by $k \in \{0, \ldots, \pi - 1\}$ of a label λ if $\lambda'[i] = \lambda[(i - k) \bmod \pi]$ for all $i \in \{0, \ldots, \pi - 1\}$.

Instead of dealing with the complete label tree, an infinite object for $\pi \geq 8$, we construct the *Stage I tree* as follows. Order nodes first so that previous generations are ahead of later generations, and within the same generation so that 0-successors (resp. their progeny) are ahead of 1-successors (resp. their progeny). A node is without successors, i.e., a leaf, if it either

- dies by period doubling: its label has no 3 and an odd combined number of 1's and 2's; or

- dies by repetition: its label equals up to rotation the label of a node ahead of it in the ordering.

Clearly, a Stage I tree is always finite, but may be very large. For $\pi = 8$, this tree has 90 nodes in 49 generations; for $\pi = 16$ it has 1050 nodes in 147 generations; and for $\pi = 32$ it has 534834 nodes in 4910 generations. For $\pi = 64$ the number of nodes is in at least the tens of millions and computer memory limitations have prevented us from completing the Stage I construction. By contrast, time is a much less significant constraint provided one uses binary search trees for labels that are already present in the tree. When Stage I is complete, the labels present constitute all possible *handle ends*, i.e., all possible labels given by the the last pair of sites in the orbit of the handle part of a handle-link pair.
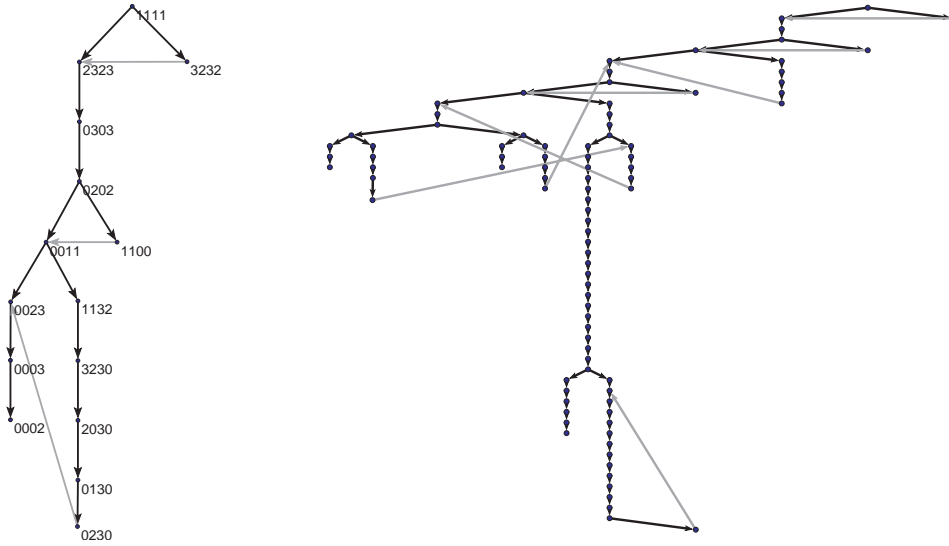
Fig. 5. The Stage I tree for $\pi = 4$ with labels, and for $\pi = 8$ without labels. Backward (gray) edges that connect a label with its rotation are also included. The bottommost backward edge in the $\pi = 8$ tree points from a leaf to a node 13 generations back, whose label it rotates by 3. This backward edge is responsible for the robust pair with $\sigma = 13 \cdot 8/\gcd(8, 3) = 104$.

Once the Stage I tree has been generated for a given π, we run the following algorithm to find all possible links in a robust handle-link pair. Clearly, the link is the more important part of a pair because it determines the periodic attractor to which the finite seed converges, whereas the different possible handles for a given link relate to the size of its attractor basin.

The R-algorithm.

A leaf in the Stage I tree with a label that includes a 3 will be called a *3-leaf*; note that such nodes die by repetition. For each 3-leaf, carry out the following procedure:

(R1) Generate the successors of the leaf, adding them to the Stage I tree, until one of them is an R-repetition or has a label with no 3.

The criterion for node x to be an *R-repetition* is as follows. Trace back along the chain of proper (i.e., different from x) ancestors of x with labels that include at least one 3. Call the resulting set of labels (all of which include at least one 3) the *R-set* of x. Then x is an R-repetition if its label is a rotation of a label in its R-set.

Every 3-leaf x for which (R1) ends in an R-repetition gives a robust handle-link pair, by the following formulas. These may look confusing, but we merely read off the handle and link configurations from the relevant labels. Waiting for an exact label replica rather than allowing for a rotation simplifies the formulas (since $r = 0$ in that case) but increases the computation time.

(R2) Denote the label of node x by $\lambda(x)$. Consider the lineage starting from the root x_0, and continuing as $x_1, \ldots, x_{i_0}, \ldots, x_g = x$, where $\lambda(x_{i_0})$ is in the R-set of x_g, and $\lambda(x_g)$ is the rotation by r of $\lambda(x_{i_0})$. Let $n_2 = \pi/\gcd(\pi, r)$, with the convention that $n_2 = 1$ if $r = 0$. Then $h = i_0 + 1$, $\sigma = (g - i_0) \cdot n_2$, $H_i = x_i \bmod 2$, $i = 0, \ldots, i_0$, and L the vector with σ entries

$$\lambda(x_{i_0+1})[(jr) \bmod \pi] \bmod 2 , \ \ldots \ , \ \lambda(x_g)[(jr) \bmod \pi] \bmod 2 ; j = 0, \ldots n_2 - 1,$$

specify a robust handle-link pair.

We should add that the spatial and temporal periods in (R2) may not be minimal and so may need to be shortened. In any case, the initial condition consisting of handle alone generates the periodic links, as guaranteed by Proposition 3.1.

After the procedures in (R1) and (R2) are concluded, additions to the stage I tree are erased and the algorithm proceeds to the next 3-leaf.

Proposition 5.4. *Up to rotation, all possible links from robust handle-link pairs are obtained by the R-algorithm.*

Proof. To the Stage I tree, whose oriented edges we call *forward*, add *backward* edges from every node that dies by repetition to the node *which is not a leaf in the tree* that it equals up to rotation. Thus at most one backward edge can originate at a node, but many can end at it. See Fig. 5 for two examples of such augmented trees.

Any link of a robust handle-link pair has a representation in the Stage I tree as follows: $y_0 \downarrow y_1 \uparrow y_2 \downarrow y_3 \ldots y_k \uparrow y_0$, where $\downarrow$ signifies a connection through a sequence of forward edges to a leaf and $\uparrow$ a single backward arrow from a leaf. Moreover, all $\lambda(y_i)$ contain a 3. Now clearly y_1 is a 3-leaf detected by the R-algorithm. $\qquad\square$

We ran the R-algorithm for $\pi = 8, 16$, and 32. Remarkably, there is a robust handle-link pair for $\pi = 8$ with $\sigma = 104$ and shortest handle of 35 sites, as featured in Example 1d and most likely known by CA experts. There are no new examples for $\pi = 16$. The $\pi = 32$ case, however, yields new examples having 26 different links: four with $\sigma = 62$, four with $\sigma = 64$, eight with $\sigma = 960$, eight with $\sigma = 992$, and one each with $\sigma = 2944$ and $\sigma = 11440$.

Example 2. The following seed generates a robust handle-link pair with $\pi = 64$:

$$1[37 \text{ 0's}]1[28 \text{ 0's}]1[12 \text{ 0's}]1[13 \text{ 0's}]1$$

However, the link on its own is the same as in Example 1d. With $\pi = 8$, the Stage I tree is unable to find this robust pair despite the fact that the periodic attractor has temporal period 8, since extra-dimensional branching is needed to connect the root label to the link. Thus, even though our method gives a systematic way to search for all robust handle-link pairs of a given period (assuming sufficient computing power), it will not necessarily find all possible links of that period.

Example 3. Particularly simple seeds generate several of the 26 robust links for $\pi = 32$. We provide three samples; for the complete collection, see [**Gri2**].

$(3a)$ 1[25 0's]101[6 0's]1

This link has $\sigma = 62$ and $\rho = 123/496 \approx 0.248$. The periodic density is much lower than the chaotic density ≈ 0.351 due to a preponderance of dramatically larger triangles of 0's in the periodic state than are common in the chaotic state.

$(3b)$ 1[27 0's]1[28 0's]1

Here $\sigma = 992$, $\rho = 267/992 \approx 0.269$. Again, large triangles of 0's predominate, though not quite as much as in the previous case.

$(3c)$ 1[23 0's]1[8 0's]1

In this case $\sigma = 11440$, $\rho = 8191/22880 \approx 0.358$. This handle-link pair is unusual in that it gives the impression of emerging entirely by chance from chaotic dynamics. The periodic state it generates is statistically similar to the chaotic state, so its density is quite close to the chaotic density.

Since we are unable to complete the Stage I tree for $\pi = 64$, a reduced search strategy is required in order to find additional examples of periodicity. One approach is to simply follow random paths through the label tree until a repeat is encountered or a cutoff is reached. In this way Hickerson [**Hic2**] recently discovered nine robust pairs with $\pi = 64$: seven with $\sigma = 3968$, and two with $\sigma = 12032$. We now describe an alternate method, motivated especially by the search for periodic seeds, which works surprisingly well.

The S-algorithm.

A *special* label is by definition nonzero and has only intervals of 0's and 1's of even size, starting with an interval of 0's. By Lemma 5.1 below, any such label can be connected back to 1^π, so assume such a connection and start with a special label. Now generate a single path by taking the first successor to be the 1-successor, followed by 0-successors thereafter, for 5000 steps, or until a label repeats up to rotation on this path. In the latter case, generate a handle-link pair as in (R2), with reduction of space and time periods if necessary, and check whether it is robust.

Lemma 5.1. *Any special label is in the (complete) label tree.*

The reader may want to quickly check that the only special label for $\pi = 4$ appears in the Stage I tree of Fig. 5.

Proof. All special labels can be reached from $\mathcal{S}_a$. In the edge dynamics for such seeds, 100 at even times is followed immediately below by 111. Also, provided the origin is initially occupied, all odd sites (and more) are automatically 0 at even times. So we only need to consider even times, even locations, and an appropriate version of the additive dynamics. Let us introduce the variant ℓ'_t given by

$$\ell'_t(x) = (\ell'_{t-1}(x-1) + \ell'_{t-1}(x)) \bmod 2\,,$$

and consider the label dynamics for ℓ'_t on vectors of size $\pi' = \pi/2$. Clearly, as the rule has only one neighbor to the left, such labels are in the set $\{0,1\}^{\pi'}$. To complete the proof, we need to show that all such labels with a nonzero number of 1's are in the resulting label tree dynamics.

The crucial observation is that any label λ has a unique predecessor λ'. This predecessor is again given by an additive rule, specifically

$$\lambda'[i] = (\lambda[i] + \lambda[i+1]) \bmod 2 \,,$$

where the coordinate indices in square brackets are taken modulo π'. Iterates of this $\lambda \to \lambda'$ rule, which we call the *reverse label dynamics*, reach all 0's in π' steps or earlier. The reason is simple: since π' is a power of 2, the additive dynamics from a label with a single 1 creates a vector of all 1's after $\pi' - 1$ time steps (as it does not feel the periodic boundary up to this time) and is 0 after that. By additivity, then, any label also creates all 0's at those times. Now a vector of 0's is created by only two vectors in the reverse label dynamics: all 0's, or all 1's. Since we start with a nonzero vector, the first step with all 0's must be preceded by all 1's. $\square$

There are too many (2^{31}) special labels for $\pi = 64$ to check them all, so we further restrict our search to manageable subsets. The set of $\binom{31}{7}$ special labels with 4 intervals of each type is by far the most successful of those we have tried, yielding robust pairs with spatial periods $\sigma = 126, 128, 3904, 3968, 4000, 4032$, and 12032.

Example 4. Here we present two of the robust handle-link pairs discovered by the the S-algorithm with $\pi = 64$; for the complete collection, see [**Gri2**].

($4a$) 1[3 0's]1[11 0's]1[19 0's]1[3 0's]1[11 0's]1[7 0's]1[56 0's]1

This seed generates a robust pair with $\sigma = 126$ and $\rho = 451/2016 \approx 0.224$. The periodic state has an unusually large empty triangle with a row of 63 0's at its top, and hence an especially low density of 1's.

($4b$) 1[3 0's]1[11 0's]1[11 0's]1[3 0's]1[23 0's]1[3 0's]1[35 0's]1[3 0's]1[19 0's]11

In this case $\sigma = 4000$ and $\rho = 69/250 = 0.276$. The periodic state is most notable for its remarkably slow advance against the chaotic state and dramatic fluctuations at the interface, as mentioned at the end of Section 8.

Let us now turn to our quest for periodic seeds via matching handle-link pairs. We have no algorithm that looks for them directly; instead, we rely on Proposition 5.2(a), as explained below.

The M-algorithm.

For every leaf x'_0 in the Stage I tree that dies by repetition, perform the following procedure, in which subsequent nodes are denoted with a prime:

(M1) Generate the label tree rooted at x'_0, and add it to the Stage I tree, until all nodes at generation 2π of the second stage are computed.

A *second stage generation* is simply the generation counted from x_0' (so that x_0' has second stage generation 0). Recall that in a label tree there is no killing by repetition. We also make the obvious remark that (M1) needs to be performed only once for each set of labels that are all equal up to rotations.

Within the tree obtained in (M1), look for labels at second stage generation 2π that are exact repetitions of the label $\lambda(x_0')$.

(M2) Consider the lineage starting from the root x_0, and continuing with $x_1, \ldots, x_{i_0} = x_0'$, $x_1', \ldots, x_\sigma'$, where $\lambda(x_0') = \lambda(x_\sigma')$ and $\sigma = 2\pi$. Then $h = i_0 + 1$, $H_i = x_i \bmod 2$; $i = 0, \ldots, i_0$, and L the vector with σ entries

$$\lambda(x_1')[0] \bmod 2, \ldots, \lambda(x_\sigma')[0] \bmod 2,$$

determine a handle-link pair.

After (M1) and (M2) are concluded, additions to the stage I tree are erased and the algorithm proceeds to the next leaf that dies by repetition.

In the end, we obtain a list of all possible links with $\sigma = 2\pi$, by Proposition 5.5 below. Then the final step of the M-algorithm is to look for matching pairs among them.

Proposition 5.5. *Up to rotation, all links from a handle-link pairs with $\sigma = 2\pi$ are obtained by the M-algorithm.*

Proof. Consider the augmented Stage I tree as in the proof of Proposition 5.4. Take a label λ_0 not on a leaf that dies by repetition. This label then has a unique representative, up to rotation, somewhere in the Stage I tree. The only way λ_0 can be within a link of size 2π (by which we mean the 2π labels generated by a link together with its iterates) is if there is a sequence of nodes $y_0, \ldots, y_k$ so that $\lambda_0 = \lambda(y_0)$ and $y_0 \downarrow y_1 \uparrow y_2 \ldots \uparrow y_k \downarrow y_0$. (Here, the meaning of arrows is the same as in the proof of Proposition 5.4.) This means that $\lambda(y_1)$ generates a link (with a different handle) with the same spatial period as λ_0, with λ_0 contained, up to rotation, within the iterates of the link. Thus $\lambda(y_1)$ gives another solution, equivalent to that given by λ_0, and y_1 is a leaf that dies by repetition. $\square$

Our computer implementation of the M-algorithm concluded that there are no matching handle-link pairs for $\pi = 8$, and that Hickerson's Example 1e is the only matching pair with $\pi = 16$. For $\pi = 32$, the search yielded six more matching pairs. Curiously, all four robust pairs with $\pi = 32$ and $\sigma = 64$ have a match; we are unable to explain this remarkable circumstance. One of the $\pi = 32$ matching pairs (a robust one) was previously discovered by Hickerson [**Hic1**]; the others are new.

Example 5. Seeds with four to eight 1's generate the robust links that may be used to construct $\pi = 32$ periodic seeds. In each case, the periodic seed is obtained by extracting a suitable initial segment from the limit configuration, in the same manner as for Fig. 1e–f of the Introduction.

(5a) 1[3 0's]1[11 0's]1[32 0's]1

This yields a periodic seed of 478 sites with the same link as Hickerson's $\pi = 32$ discovery; his original example had length 2150.

The three new $\pi = 32$ robust examples arise from these initializations:

(5b) 1[39 0's]1[7 0's]1[3 0's]1[8 0's]1

(5c) 1[15 0's]1[7 0's]1[3 0's]1[11 0's]1[7 0's]1[3 0's]1[8 0's]1

(5d) 1[11 0's]1[7 0's]1[15 0's]1[16 0's]1

In each case, an initial segment in S_a leads to a special label followed by a single 1-successor. Possible lengths of periodic seeds are 541, 242, and 849, respectively.

The common structure of the periodic seed generators in Example 5 was the original motivation behind our S-algorithm. For $\pi = 64$, that approach has found 4 additional initializations with the same structure that generate robust handle-link pairs $H + L$. Each such $H + L$ is part of a matching pair, and produces a periodic seed after judicious surgery at the right edge. In each case, a matching handle, much simpler than H, then self-organizes. One of these discoveries is the following.

Example 6.

$$1[11 \text{ 0's}]1[3 \text{ 0's}]1[7 \text{ 0's}]1[3 \text{ 0's}]1[7 \text{ 0's}]1[47 \text{ 0's}]1[7 \text{ 0's}]1[7 \text{ 0's}]1[24 \text{ 0's}]1$$

The periodic cycle of the resulting edge dynamics is shown in Fig. 6; its leftmost 1476 cells in the 21st row comprise a $\pi = 64$ periodic seed. (Similarly, the S-algorithm yielded periodic seeds of length 987, 3345, and 3781 with different links.)
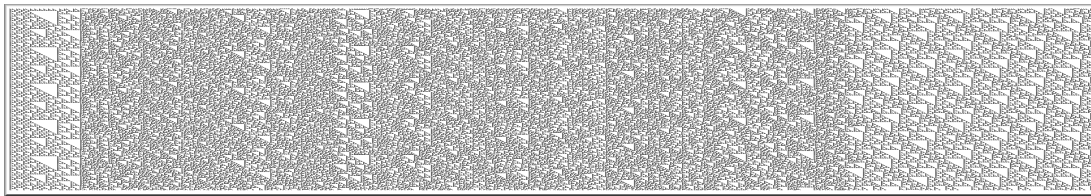

Fig. 6. The $\pi = 64$ cycle of the edge dynamics in Example 6.

In total, we now know periodic seeds for *Exactly 1* with 11 distinct periodic attractors, as summarized in Table 4. We provide the edge period, density, minimal known left and right handle lengths, and also indicate if one of the two handle-link pairs is robust. The first two are Hickerson's [**Hic1**], the next 3 were found with the R-algorithms, the 2 after that with the M-algorithm, and the final 4 with the S-algorithm.

name	π	ρ	l. handle size	r. handle size	robust edge	example no.
H1	16	19/64	93	8	none	1f
H2	32	61/256	481	8	left	5a
R1	32	29/128	544	32	left	5b
R2	32	65/256	247	32	left	5c
R3	32	61/256	860	8	left	5d
M1	32	57/256	151	4	none	—
M2	32	57/256	262	4	none	—
S1	64	109/512	1035	16	left	—
S2	64	67/256	1501	8	left	6
S3	64	217/1024	3346	32	left	—
S4	64	197/1024	3816	12	left	—

Table 4. Summary of known periodic seeds with distinct periodic attractors. In all cases, $\sigma = \tau = 2\pi$.

The design behind matching pairs remains a mystery, so we do not hazard a conjecture about whether their number (by which we mean the number of different periodic attractors) is finite or infinite. By contrast, robust examples are much easier to come by and our experiments suggest that there may well be infinitely many of them, although we do not have a proof.

Let us conclude this section with an experiment showing how a nearly chaotic region may be a feature of a fraternal replicator.
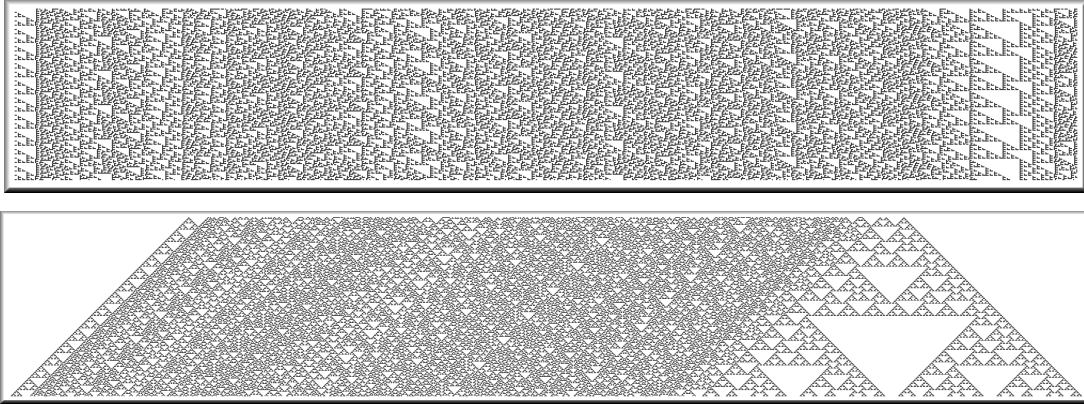

Fig. 7a–b. Construction of a fraternal replicator with nearly chaotic edge dynamics starting from the seed of Example 7.

Example 7.

$$1[24 \text{ 0's}]1[8 \text{ 0's}]101[3 \text{ 0's}]1$$

Starting from this low-density seed, the periodic cycle of its edge dynamics, for the first 1000 cells, is shown in Fig. 7a. Note that 930 cells from the left edge, at the far right of the figure,

the system briefly enters an additive regime, as if this were a space-time slice of the additive evolution from a seed in S_a, before doubling from period 32 to 64. If we truncate at that column, and start *Exactly 1* from one of the length 931 edge configurations in the $\pi = 32$ cycle, then further evolution is additive and we obtain a fraternal replicator, the first 231 updates of which are shown in Fig. 7*b*.

6   Entropy of periodic seeds

Let $p(n)$ be the number of periodic configurations on the interval $[0, 1, \ldots, n+1]$ that begin and end with 1, and define
$$h_p = \liminf_{n} n^{-1} \log_2 p(n).$$

Theorem 4. *The periodic entropy satisfies $h_p \geq 0.5850$.*

Proof. Let A be any configuration that leads to $A' \in S_7$, as defined at the end of Section 3, after 4 updates. Suppose we also have a periodic seed P that after 5 updates leads to a configuration P' with a 1 at the left edge separated from other occupied sites by a 0-block of length at least 5. Then by Lemmas 3.1 and 4.1, if A' has its rightmost 1 at $x = -2$, it follows that $A'P'$ is identical to P' for $x + t \geq 4$. Hence $A'P'$ and AP are also periodic.
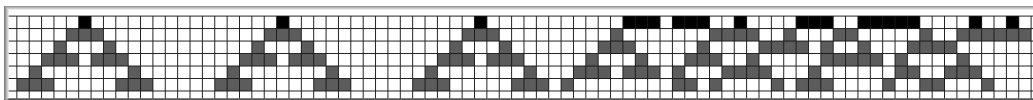

Fig. 8. An example of the construction in the proof of Theorem 4.

For the odd case, Hickerson's length 85 periodic seed P given in the Introduction suffices: after 5 updates it has a 1 at the left edge followed by 8 0's. (Fig. 8 shows a detail near the origin for an example of the construction based on this seed.) For even length, Example 5*c* does the trick. If it is run for 827 steps, then the leftmost 248 cells are extracted, the resulting periodic seed P has the desired property after 5 additional updates: a 1 at the left edge followed by 15 0's.

In order to apply Lemma 3.1 with R a singleton at the origin, A' appended on the left (with a 0 at $x = -1$), and $R' = P'$, it suffices to check that the evolution of seed $A'R$ agrees with that of R for $x + t \geq 4$. We defer this step to the end of the proof.

Clearly, there is a distinct seed AP for each A leading to S_7 in 4 steps. Another application of the j-triples technique, summarized in Table 5, produces the given bound. (For $5 \leq j \leq 7$ the bound is worse and we have been unable to compute the $j = 8$ bound.)

j	$n_0 + n_1$	$\lambda_{\max}$	$\log_2 \lambda_{\max}$
0	9	$2^{1/8}$	0.1250
1	35	1.4831	0.5685
2	123	1.4847	0.5702
3	439	1.4801	0.5656
4	1563	1.5001	0.5850

Table 5. Entropy estimates for $\mathcal{S}_7$ based on j-triples.

All that remains to complete the proof, then, is the following result.

Lemma 6.1. *The evolution of any seed $A'R$, as described above, agrees with that of R for $x + t \geq 4$.*

Proof. We compare the edge dynamics of $A'R$ and R; a representative example is detailed in Fig. 9. Now the edge dynamics from $\mathcal{S}_a$ also emulates an additive dynamics at even times. As in the proof of Lemma 5.1, we find it convenient to use variant ℓ'_t. It is easy to check that an initial seed from $\mathcal{S}_7$ corresponds to starting ℓ'_t from a seed A with only even locations occupied. We also assume that the rightmost 1 of such an A is at the origin. Our first claim is that for $k = 0, 1, 2, \ldots$, either both $\ell'_{2k}(0)$ and $\ell'_{2k+1}(0)$ are 1, or both are 0.

To check this, note that ℓ'_t is the prototypical dynamics with cancellative duality [**Gri1**]: for any initial set A, $\ell'^A_t(x) = 1$ if and only if $\#(\ell'^{\{x\}}_t \cap A)$ is odd. (This can be proved using another equivalence, easily shown by induction: $\ell'^A_t(x) = 1$ if and only if the number of paths from $\{(a, 0) : a \in A\}$ to (x, t) is odd — here paths are allowed to make two directed steps, $(x - 1, s - 1) \to (x, s)$ or $(x, s - 1) \to (x, s)$.) Now, by definition of ℓ'_t, and duality

$$\ell'^A_{2k+1}(0) = (\ell'^A_{2k}(0) + \ell'^A_{2k}(-1)) \bmod 2 \; = \; (\ell'^A_{2k}(0) + \#(\ell'^{\{-1\}}_t \cap A)) \bmod 2 \,.$$

To establish the claim, observe that $\ell'^{\{-1\}}_t$ occupies only odd sites at even times, another property easily proved by induction, so this last intersection is empty.
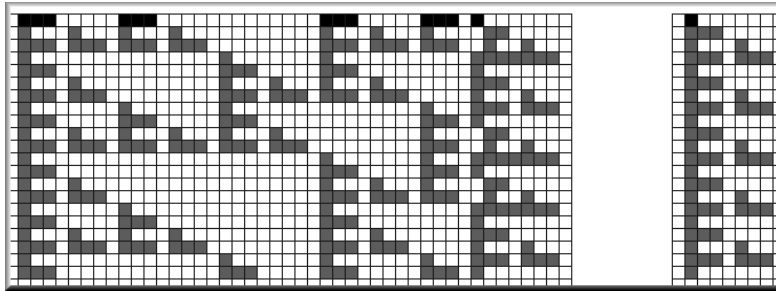

Fig. 9. Edge representations from $A'R$ (left) and R (right) in the construction for Theorem 4. The perturbation due to A' extends only 3 cells beyond the left edge of the occupied set started from R.

Returning to our *Exactly 1* construction, we conclude that the edge dynamics of A' at sites -4 through -1 consists of either *L-blocks*: 1000 followed by 1110, or *E-blocks*: 0000 followed by 0000, and that the same types always appear in successive even-odd pairs. (Only the bottom row of the initial L-block appears, as a 1110 in the seed A'.) It now remains to calculate the interaction of this label structure with the evolution of R, a singleton at the origin. This amounts to somewhat tedious case checking, so we will only describe the key features. Consult Fig. 9 for confirmation in a representative example. First, note that A' produces all 0's at $x = -1$ since that column begins with a 0 and the previous two columns have either 00 or 11 in each row. It follows that the only discrepancies at $x = 0$ due to A', from the column of all 1's created by R, consist of 00 temporal pairs at certain times $2k - 1$ and $2k$, $k \geq 1$, and that each such pair must be preceded by a 1 at time $2k - 2$. Then the only discrepancies possible at $x = 1$, from the alternating pattern of R, are at even times at least 4 apart, where a 1 replaces a 0. Consequently, at $x = 2$ there is complete agreement, and at $x = 3$ the only discrepancies arise one time step later than they did at $x = 1$, again with a 1 in place of a 0. Finally, at both $x = 4$ and $x = 5$ the two evolutions are identical, so the effect of A' has vanished. □

7 Entropy of chaotic seeds

Fig. 1c illustrates the evolution of a seed that is believed to be chaotic in the sense of Section 2. Many additional examples are mentioned in [**Wol3**], p. 951. Although the proportion of seeds with length at most 12 that are chaotic in any wedge is certainly less than 0.38, we conjecture that for large lengths the chaotic seeds have asymptotic density 1. As a modest step in this direction, we let $c(n)$ be the number of chaotic configurations on the interval $[0, 1, \ldots, n + 1]$ that begin and end with 1, define

$$h_c = \liminf_n n^{-1} \log_2 c(n),$$

and determine a lower bound on this chaotic entropy *assuming*

(H) There it at least one chaotic seed of each parity with 110110 at its left edge, say 1101101 (Example 1d) and 11011001.

Theorem 5. *Under assumption H, the chaotic entropy satisfies $h_c \geq 0.5850$.*

Proof. We argue essentially the same way as for the periodic case. Given (H), let C be either of the specified seeds, or any other such, with its leftmost 1 at $x = 5$. As in Theorem 4, let A be any configuration that leads to $\mathcal{S}_7$ after 4 updates. Then we claim that

$$C_* = A[11 \ 0\text{'s}]C$$

is also chaotic. See Fig. 10 for an example. Note that the evolutions of A and C do not interact during the first five updates, and that at $t = 5$ the occupied region has the form

$$A'0C'.$$

Here A' has it's rightmost 1 at $x = -2$ and consists of 1-blocks of length 3 separated by 0-blocks of length $5 \bmod 8$ (the immediate successors of configurations in $\mathcal{S}_7$). C' has its leftmost 1 at the origin, followed by a 0-block of length at least 5.
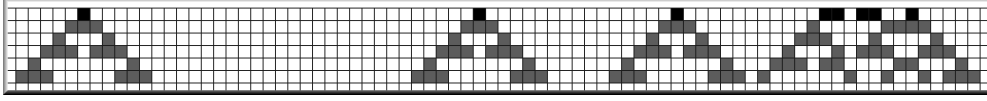

Fig. 10. An example of the construction in the proof of Theorem 5.

By Lemma 6.1, we know that C_* evolves in exactly the same way as the chaotic seed C at all sites more than distance 3 from the left edge. Hence C_* is also chaotic, in the same wedge as C. Finally, we have seen in the proof of Theorem 4 that the number of choices for A with odd length $n + 2$ is at least a constant times $2^{0.5850\,n}$. □

8   Empirical observations on chaotic seeds

To our knowledge, p. 951 of [**Wol3**] contains the only claim in the literature about the evolution of *Exactly 1* from a (finite) chaotic seed: that starting from 101101, the central chaotic wedge advances exactly as fast as the *Lyapunov exponent*, defined as the equilibrium velocity of damage spreading [**Gra2**]. No explanation, evidence, or general condition for validity of this property is offered. Here we will present considerable empirical evidence that Wolfram's statement is false, provided one accepts the natural definition of chaos through thermodynamic limits, as given in Section 2. Instead, the central wedge is distinguished by a different rate of convergence to the thermodynamic quantities. Even this second order effect, although apparently true for the majority of chaotic seeds, does not always hold; in fact it fails with a probability bounded from below for large random seeds. We proceed to illustrate and explain these intriguing phenomena.

Consider first the introductory chaotic case, Example 1c. Our primary focus is on the *density profile*, i.e., the function $\rho(u)$ giving the average occupation on each ray $t = \lfloor ux \rfloor$, for $u \in [-1, 1]$. Given a time t and number N, we obtain an approximation $\hat{\rho}$ to this function as follows. For $-N \le i < N$, with

$$B_i = \left\{ (x, s) : 0 < s \le t, \frac{i}{N} \le \frac{x}{s} < \frac{i+1}{N} \right\},$$

$$n_i = \#B_i,$$

$$n_i^1 = \#\{(x, s) \in B_i : \xi_s(x) = 1\},$$

let $\hat{\rho}(1) = 1$, $\hat{\rho}(i/N) = n_i^1/n_1$, and interpolate linearly. We have computed $\hat{\rho}$ for $t = 10^6$ and $N = 10^4$. At first glance this approximation is very close to a constant function with no interesting features, but a suitably large zoom, displayed in the left frame of Fig. 11, reveals a sharp transition occurring near each edge, at $\pm u_c$ where $u_c \approx 0.77$. This effect is quite subtle and almost invisible before a few hundred thousand iterations. For $|u| < u_c$, the occupied sites appear to exhibit typical properties of a random field with low correlations; we therefore

conclude that the density profile approaches a constant value ρ_c. This value can be estimated by

$$\hat{\rho}_c = \sum_{|i| \leq 0.6\,N} n_i^1 \bigg/ \sum_{|i| \leq 0.6\,N} n_i \ ,$$

giving to five decimals $\rho_c \approx 0.35095$. Note that the estimate is based on a number of sites of order t^2 so, assuming central limit behavior, the error is of order t^{-1}. This assumption is confirmed by the right frame of Fig. 1, which suggests that the average standard deviation

$$\hat{\delta} = \frac{1}{1.2\,N} \sum_{|i| \leq 0.6\,N} |n_i^1 - \hat{\rho}_c \cdot n_i| / \sqrt{n_i}\,,$$

converges, and for large t and N is close to a constant somewhat less than 0.8. Our density computations are without the periodic boundary effects of those in [**Gra2**]; nevertheless, the estimates agree. In the notation of Section 2, $\rho_c = \rho_{\{(0,0)\}}$; we also find, for instance, $\rho_{\{(0,0),(1,0)\}} \approx 0.1848$ and $\rho_{\{(0,0),(2,0)\}} \approx 0.1395$.

What is going on for $|u| > u_c$? It turns out that cells there are trapped in a periodic cycle of the edge dynamics, defining the space-time *quasiperiodic* region. The period of this cycle increases as one moves away from the edge. Presumably, the period approaches infinity near u_c as time gets large, but simulations indicate that this divergence is very slow — e.g., the largest period at time 10^6 is 256. As explained in Section 5, the period is always a power of 2, and can jump only by a factor of 2 unless it jumps to infinity. Amusingly, the dynamics thus appears to exhibit a kind of "period-doubling route to chaos" familiar from iteration of maps.
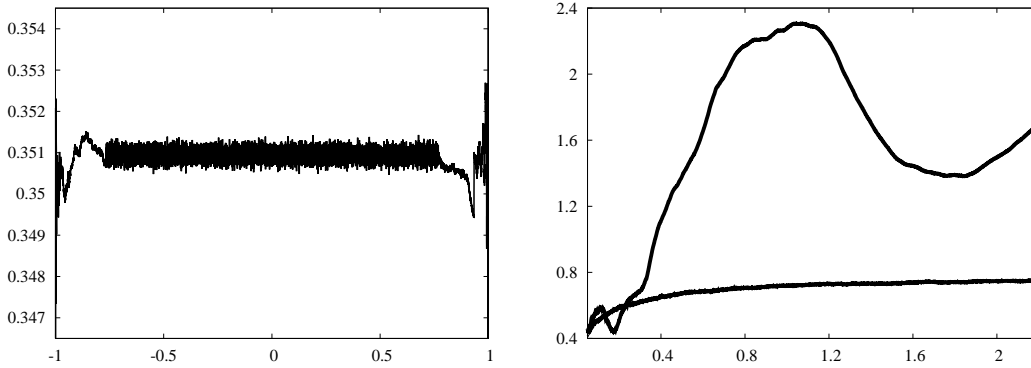

Fig. 11. Example 1c. Left: the approximate density profile $\hat{\rho}(u)$ at time 10^6. Right: standard deviation estimate $\hat{\delta}$ vs. time (in millions); the bottom curve represents central aperiodic region and the top curve the quasiperiodic region. In this figure, and in all subsequent ones, $N = 10,000$ is fixed.

The approximate value of u_c (which we estimate to be between 0.7663 and 0.7664) is strikingly close to the Lyapunov exponent estimated in [**Gra2**]. This seems unlikely to happen by coincidence and there is indeed a compelling, although heuristic, *microscopic* explanation why the two constants should be equal, which we now sketch.

To begin, we formulate damage spreading in the context of edge dynamics. Consider two systems, started from different initial states, that agree for $x < 0$, disagree at 0, and are otherwise in (independent) equilibria. If they become equal at $x = 0$ next time, then equality extends into $x \geq 0$ and the new leftmost point of disagreement is the closest point to the origin at which they disagree. In addition, they are bound to become equal at $x = 0$ if their states at $x = -1$ and $x = -2$ are 11. This is exactly the mechanism by which the periodic region extends at a fixed period π, namely by checking which contingent interval of states beyond the current "trapped" locations of states agrees after π time steps, and again this region will increase by at least 1 if the currently trapped region ends in 11. Computations such as those depicted in Fig. 6 also make it clear that the the periodic sites extend with π fixed for the preponderance of time. To conclude that the two microscopic rates of expansion agree asymptotically, we assume that π is large and argue that states to the right of trapped points are in the nearly independent chaotic equilibria at the beginning and end of the period. Thus the two environments — the first determines the spread of equality in the edge version of damage spreading while the second governs the spread of points trapped in a fixed period π by the edge dynamics — have asymptotically the same statistics.

The discussion so far may suggest that the quasiperiodic state generated by the rule for $|u| > u_c$ is statistically different from the chaotic state in the middle cone $|u| < u_c$. However, we claim that this is not true. First, note that sites (x, t) with $x = tu$ belong to a region of longer and longer period as t increases, as soon as $|u| < 1$. Heuristically, once sites are caught in a really long period, there is no reason to expect the local statistics to be significantly different from those in the aperiodic central cone. To confirm this intuition, we approximated $\rho_{\{(0,0)\}}$, $\rho_{\{(0,0),(1,0)\}}$, and $\rho_{\{(0,0),(2,0)\}}$, for the cone defined by $0.8 \leq |u| \leq 0.9$, at time 10^6. All these quantities differ from their central counterparts by less than 0.0001. Clearly, though, quasiperiodicity imposes significant space and time correlations affecting the fluctuations around this limit. The resulting oscillations are strong enough that we cannot confidently predict, say, the rate of convergence of the density profile to its limit ρ_c, but it seems clear that it is slower than $1/t$ (and may be dependent on the slope u). See Fig. 11, where the right frame also shows the evolution of the counterpart of $\hat{\delta}$ in the quasiperiodic region (defined with the range of summation $0.8\,N \leq |i| \leq 0.9\,N$ and corresponding normalization $0.2\,N$).

Our experiments suggest that the scenario described above holds at every edge which does not find a robust handle-link pair, and that for the majority of large seeds both edges fall in this category. On the other hand, a robust pair occurs with probability at least 2^{-32}. To see this, recall that our shortest handle (from Example 1d) has 35 sites, with a 1-condition required at the 36th site to initiate the link. Since $\pi = 8$, there are at least 8 configurations on the 35 sites following the leftmost 1 that lead to a robust pair. The remaining examples of this section will illustrate what transpires in this case.

For example 1d the approximate density profile at time 10^6 is depicted in Fig. 12. One easily sees the phase transitions between ρ_c and $\rho_1 = 11/26$ at $\pm u_c^1$, with $u_c^1 \approx 0.70$. The periodic region is thus decidedly larger that the quasiperiodic one in the previous example. Needless to say, the periodic profile is constant at 8 beyond u_c^1, so the dynamics transitions directly from this period to chaos.
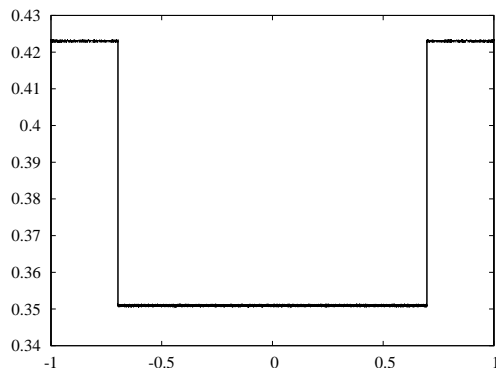
Fig. 12. The approximate density profile $\hat{\rho}(u)$ for Example $1d$.

As a rule of thumb, velocity of growth of a periodic region increases with its density, although, as we will see below, this is not always so. We know of one additional case such that $\rho > \rho_c$ and the periodic cone extends from the left edge to beyond $-u_c$, both barely. That is Example $3c$, with density about 0.358 and the chaotic region's left edge at $-u_c^2 \approx -0.764$.

Next, consider Example $3b$. Its density profile is depicted at time $2.2 \cdot 10^6$ in Fig. 13. The right edge finds no robust pair and therefore behaves like the two edges in Fig. 11. So we concentrate on the left edge. Observe the signature of its $\pi = 32$ robust handle-link pair with density ≈ 0.269 and its transition to chaos at $-u_c^3 \approx -0.88$, decidedly to the left of the Lyapunov exponent $-u_c$. Even though we ran this case for an extended time, the detail at the right of Fig. 13 exhibits no apparent change at $-u_c$.
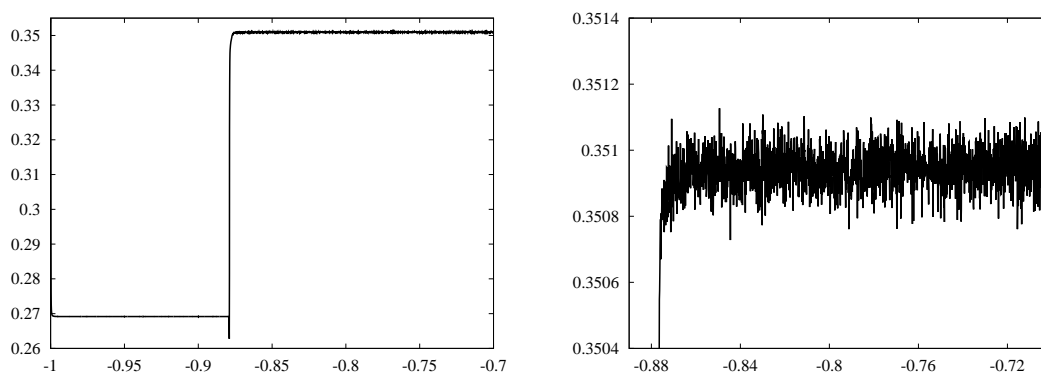

Fig. 13. The approximate density profile $\hat{\rho}(u)$ at the left edge for Example $3b$.

Finally, we examine Example $4b$. Of all the robust seeds we know, this one has the slowest growth of the periodic region, with its edge at $-u_c^4 \approx -0.94$. Convergence to the limiting density profile is also rather slow, as illustrated by Fig. 14, which superimposes the profiles at times 10^5 and 10^6. We remark that the periodic wedge in Example $4a$ grows much faster, with $-u_c^5 \approx -0.83$, despite its smaller density.
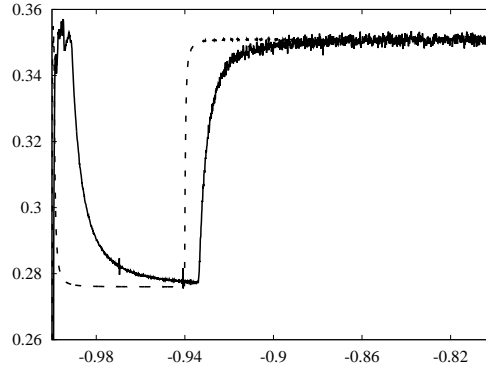
Fig. 14. The approximate density profile $\hat{\rho}(u)$ at times 10^5 and 10^6 (dashed curve), near the left edge, for Example $4b$.

To summarize, the extent of the central chaotic wedge at either side depends on whether a robust handle-link pair appears at the corresponding edge. If it does, the critical slope, which delineates the chaotic wedge from the periodic one, is given by the expansion velocity of the robust periodic state. Therefore, a large, possibly infinite, number of critical slopes are possible. If the edge does not find a robust pair, then the chaotic state extends all the way to the maximum slope (1 or -1) at that edge. This state is divided into two regions with the same asymptotic statistics but the outer, quasiperiodic region feels the effects of the edge dynamics, resulting in a slower rate of convergence. The critical slope between the two is given by a universal constant u_c, the Lyapunov exponent for *Exactly 1*.

9   Additional open problems

(1) Assume seed A is a maternal replicator. Must there be a time t when $\xi_t^A \in \mathcal{S}_q$?

(2) Is any finite sequence of 0's and 1's part of a replicator? The same question arises for periodic seeds.

(3) Can a nontrivial upper bound on the replicating or periodic entropies be proved? Note that negative answers to questions in (2) imply corresponding positive answers to these.

(4) What is the size of the Stage I tree for large π?

(5) Is there a periodic seed attracted to a periodic state within a wedge $W_{\alpha,\beta}$ for some $-1 \leq \alpha < 0 < \beta < 1$, but not within $W_{-1,1}$?

(6) There are a great many robust handle-link pairs with spatial period of the form $\sigma = 31 \cdot 2^k$, e.g., $992 = 31 \cdot 32$ for $\pi = 32$ and $3968 = 31 \cdot 128$ for $\pi = 64$. Why?

(7) Why are all our robust handle-link pairs with $\sigma = 2\pi$ part of a matching pair?

References

[**ACEGP**] D. Auerbach, P. Cvitanović, J.-P. Eckmann, G. Gunaratne, I. Procaccia, *Exploring chaotic motion through periodic orbits*, Phys. Rev. Lett. 58 (1987), 2387–2389.

[**BP**] E. Bilotta, P. Pantano, *Emergent patterning phenomena in 2D cellular automata*, Artificial Life 11 (2005), 339–362.

[**DG**] C. Domain, H. Gutowitz, *The topological skeleton of cellular automaton dynamics*, Physica D 103 (1997), 155–168.

[**Eis**] M. Eisele, *Long-range correlations in chaotic cellular automata*, Physica D 48 (1991), 295–310.

[**Epp**] D. Eppstein, *Cellular Automata: Replicators*.
http://www.ics.uci.edu/ eppstein/ca/replicators/

[**Eva**] K. M. Evans, *Replicators and Larger than Life examples*, in "New Constructions in Cellular Automata," eds. D. Griffeath and C. Moore, Oxford University Press, 2003, 119–159.

[**FG**] P. Flocchini, F. Geurts, *Searching for Chaos in Cellular Automata: New Tools for Classification*, Complexity International 2 (1995). http://www.complexity.org.au/

[**FLM**] P. Favatia, G. Lotti, L. Margara, *Additive one-dimensional cellular automata are chaotic according to Devaney's definition of chaos*, Theoretical Computer Science 174 (1997), 157–170.

[**GG1**] J. Gravner, D. Griffeath, *Cellular automaton growth on $\mathbb{Z}^2$: theorems, examples and problems*, Advances in Applied Mathematics 21 (1998), 241–304.

[**GG2**] J. Gravner, D. Griffeath, *Asymptotic densities for Packard Box rules*, Nonlinearity 22 (2009), 1817–1846.

[**Gra1**] P. Grassberger, *New mechanism for deterministic diffusion*, Phys. Rev. A 28 (1983), 3666–3667.

[**Gra2**] P. Grassberger, *Long-range effects in an elementary cellular automaton*, Journal of Statistical Physics 45 (1986), 27–39.

[**Gri1**] D. Griffeath, "Additive and Cancellative Interacting Particle Systems." Springer, 1979.

[**Gri2**] D. Griffeath, *Primordial Soup Kitchen*.
http://psoup.math.wisc.edu/exactly1/exactly1.html

[**Hic1**] D. Hickerson, private communication (1998).

[**Hic2**] D. Hickerson, private communication (2009).

[**Jen1**] E. Jen, *Aperiodicity in one-dimensional cellular automata*, Physica D 45 (1990), 3–18.

[**Jen2**] E. Jen, *Exact solvability and quasiperiodicity of one-dimensional cellular automata*, Nonlinearity 4 (1991), 251–276.

[**Jen3**] E. Jen, *Transience and dislocations in one-dimensional cellular automata*. In "Cellular automata and cooperative systems (Les Houches, 1992)," NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 396, Kluwer Acad. Publ., 1993, 299–310.

[**KN**] K. Eloranta, E. Nummelin, *The kink of cellular automaton Rule 18 performs a random walk*, Journal of Statistical Physics 69 (1992), 1131-1136.

[**KT**] K. Kaneko, I. Tsuda, "Complex Systems: Chaos and Beyond, A Constructive Approach with Applications in Life Sciences." Springer-Verlag, 2000.

[**MM**] T. G. Mattos, J. G. Moreira, *Universality classes of chaotic cellular automata*, Brazilian Journal of Physics, 34 (2004), 448–451.

[**TR**] A. Trevorrow, T. Rokicki, *Golly*. `http://golly.sourceforge.net/`

[**Wil1**] S. J. Willson, *Cellular automata can generate fractals,* Discrete Appl. Math. 8 (1984), 91–99.

[**Wil2**] S. J. Willson, *Computing fractal dimensions for additive cellular automata,* Physica D 24 (1987), 190–206.

[**Woj**] M. Wójtowicz, *Mirek's Cellebration: a 1D and 2D Cellular Automata explorer*. `http://www.mirwoj.opus.chelm.pl/ca/`

[**Wol1**] S. Wolfram, *Statistical Mechanics of Cellular Automata*, Rev. Mod. Phys. 55 (1983) 601–644.

[**Wol2**] S. Wolfram, *Computational Theory of Cellular Automata*, Communications in Mathematical Physics 96 (1984), 15–57.

[**Wol3**] S. Wolfram, "A New Kind of Science." Wolfram Media, 2002.

[**Zab**] J. G. Zabolitzky, *Critical properties of Rule 22 elementary cellular automata*, Journal of Statistical Physics 50 (1988), 1255–1262.