


# Asymptotic behaviour of the one-dimensional “rock-paper-scissors” cyclic cellular automaton

Benjamin Hellouin de Menibus<sup>1</sup> and Yvan Le Borgne<sup>2</sup>

<sup>1</sup> Laboratoire de Recherche en Informatique,  
Université Paris-Sud - CNRS - CentraleSupélec, Université Paris-Saclay, France

 <https://orcid.org/0000-0001-5194-929X>

<sup>2</sup> Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, F-33400 Talence, France

## Abstract

The one-dimensional three-state cyclic cellular automaton is a simple spatial model with three states in a cyclic “rock-paper-scissors” prey-predator relationship. Starting from a random configuration, similar states gather in increasingly large clusters; asymptotically, any finite region is filled with a uniform state that is, after some time, driven out by its predator, each state taking its turn in dominating the region (heteroclinic cycles).

We consider the situation where each site in the initial configuration is chosen independently at random with a different probability for each state. We prove that the asymptotic probability that a state dominates a finite region corresponds to the initial probability of its prey. The proof methods are based on discrete probability tools, mainly particle systems and random walks.

Keywords: cyclic dominance, heteroclinic cycles, cellular automata, self-organisation, random walk

Cyclic dominance is a general term for phenomena where different states (species, strategies, etc.) are in prey-predator relationships that form a cycle: A preys on B preys on C... preys on A. This phenomenon occurs in many natural or theoretical systems, among which a few examples are:

**Population ecology** male mating strategies in side-blotched lizard [27], antibiotic production and resistance in *E.Coli* [18], parasite-grass-forb interactions [5], oscillations in the population size of pacific salmon [12], etc.

**Game theory** pure or stochastic strategies in rock-paper-scissors type games, iterated prisoner’s dilemma [26, 15], public goods games [13, 24], etc.

**Infection models** The SIRS compartmental model [1] (susceptible / infectious / recovered, when a recovered agent may become susceptible again), forest fire models [2], etc.

Many additional examples can be found in [29] (section 7) and [30].

May and Leonard’s [21] is the first effort to model the evolution of three species with cyclic dominance, using the standard Lotka-Volterra equations; it is a mean-field approximation, that is, it assumes the population is well-mixed. The system exhibits so-called *heteroclinic cycles* where each species in turn dominates almost the whole space before being replaced by its predator. Consequently, cyclic dominance has been proposed as a mechanism to explain the coexistence of various strategies or species [17] (biodiversity), the regular oscillations in population sizes of different species [12], and some counter-intuitive phenomena such as the “survival of the weakest” [11]. In other contexts, heteroclinic cycles appear to coincide with important concepts: for example, social choice among three cyclically dominant choices can lead to an heteroclinic cycle along the so-called *bipartisan set* [19].

Mean-field models do not take into account spatial aspects of the evolution of populations, such as the effect of population structure, mobility, dispersal, local survival, etc. This is why spatial models have been introduced both in ecology [6, 31] and in so-called evolutionary game theory [29]. In both cases agents have a spatial location and can only interact with their neighbours at short range. There is some variety in spatial models:

**Space** a lattice in one, two or more dimensions, or a graph with more structure;

**Updates** discrete or continuous time, synchronous or asynchronous updates;

**Dynamics** usually a predator replaces a prey by a copy of itself (*replicator dynamics*). The model can include empty space, different ranges, threshold effects, invasion probabilities, etc.;

**Boundaries** infinite, periodic or fixed boundary conditions, choice of the initial configuration.

In this article, we consider arguably the simplest spatial model for cyclic dominance: the one-dimensional, 3-state cyclic cellular automaton. Each site on the lattice  $\mathbb{Z}$  is initially associated a state in  $\mathbb{Z}/3\mathbb{Z}$ . At each (discrete) time step, every site is updated synchronously: if any of the two neighbouring sites contains a predator, it becomes the new state for this site. While the restriction to one dimension may not be ecologically realistic (two-dimensional models being the object of more interest [31]), it has two benefits. First, its simple spatial structure makes many questions mathematically tractable, while the two-dimensional models have much more complex dynamics with structured interfaces between regions [7]. Second, its dynamics is similar to a *interacting particle system* with borders progressing at constant speed and annihilating on contact (*ballistic annihilation* - see Figure 1); this is a subject of independent interest [4] and many tools have been developed for it [3].

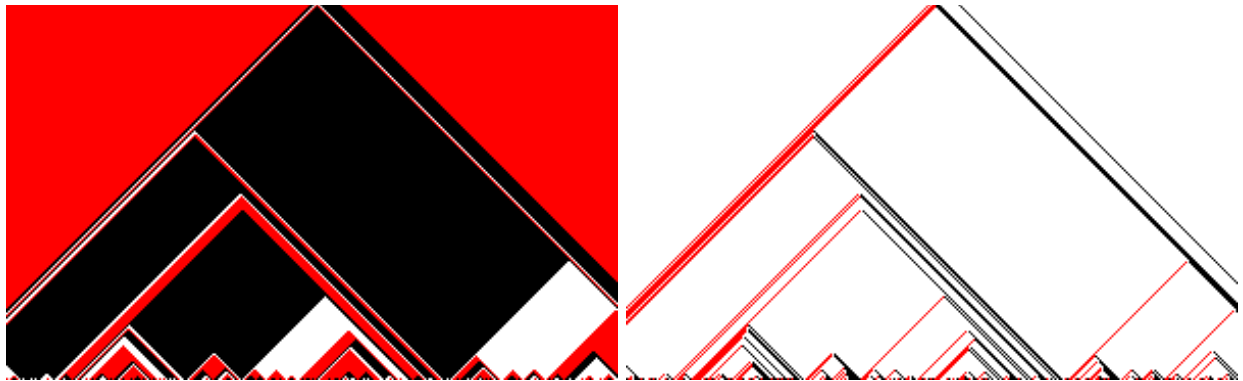


Figure 1: (Left) The 3-state cyclic cellular automaton; (Right) The dynamics of its particles.

**Note** In all space-time diagrams of this article, the initial configuration is drawn horizontally at the bottom and time goes from bottom to top. States are represented by colours following the convention  $0 \mapsto \square$ ,  $1 \mapsto \blacksquare$ ,  $2 \mapsto \blacksquare$ ,  $3 \mapsto \blacksquare$ ,  $4 \mapsto \blacksquare$ .

The seminal work of Fisch [8] focused on the case where each site is independently assigned a random state with uniform probability. He proved a clustering phenomenon: for 3 or 4 states, large monochromatic regions emerge and grow, but each region keeps changing state arbitrarily late (*fluctuation*, the spatial counterpart of heteroclinic cycle); for 5 states or more, the regions reach a limit size then stay unchanged (*fixation*). These behaviours, illustrated in Figure 2, are considered as a prime example of self-organisation in a relatively simple model [25]. These results were later refined in terms of cluster growth rate, number of state changes, etc. [9, 10, 20].

The present article focuses on the asymptotic behaviour of the 3-state cyclic cellular automaton when the initial configuration is chosen independently at random, but with distinct probabilities for each state,

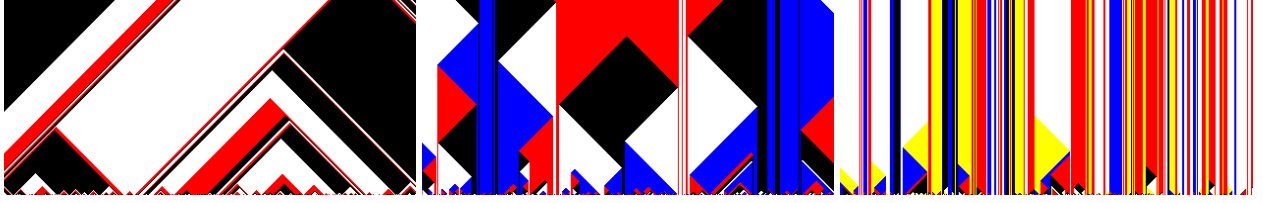


Figure 2: Left to right, the 3-, 4- and 5-state cyclic cellular automata iterated on an initial configuration drawn according to the uniform Bernoulli measure.

breaking the symmetry. It is not hard to see that the same clustering phenomenon as in the uniform case occurs. Our main result (Theorem 6) is that the asymptotic probability for any region to be dominated by a given state corresponds to the initial probability of its prey; this completely determines the limit probability measure. A similar relationship was observed empirically between invasion rates and asymptotic probability in more complex models [32]; see [29], Section 7.7 for a detailed account. However, we could not find a conjecture for this phenomenon in such a simple model, and this is the first formal proof of a similar result to our knowledge.

Our approach is based on a correspondence between the time evolution of the borders and some well-chosen random walk, a method that was already used in the study of one-dimensional cellular automata [3]. Compared to previous work, the random walk is not the standard symmetric walk and the probability of a step up or down depends on the current position.

## 1 Definitions

### 1.1 Symbolic space

For  $\mathcal{A}$  a finite alphabet, define  $\mathcal{A}^* = \bigcup_{n \in \mathbb{N}} \mathcal{A}^n$  the set of finite *patterns* (or words) and  $\mathcal{A}^{\mathbb{Z}}$  the set of (one-dimensional) *configurations*, that is, the set of bi-infinite words over the alphabet  $\mathcal{A}$ . For  $u \in \mathcal{A}^*$ , denote  $|u|$  its length, and for  $i \in \mathbb{Z}$ , define the *cylinder*  $[u]_i = \{x \in \mathcal{A}^{\mathbb{Z}} : x_{i, \dots, i+|u|-1} = u\}$ , with  $[u] = [u]_0$ . Cylinders form a clopen basis of  $\mathcal{A}^{\mathbb{Z}}$  for the product topology. A word  $u \in \mathcal{A}^*$  is a *factor* of a configuration  $x \in \mathcal{A}^{\mathbb{Z}}$  if  $x \in [u]_i$  for some  $i \in \mathbb{Z}$ .

Define the *shift function*  $\sigma : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$  by  $\sigma(x)_i = x_{i-1}$  for any  $i \in \mathbb{Z}$ . From a finite pattern  $w \in \mathcal{A}^*$  define the infinite  $\sigma$ -periodic configuration  ${}^\infty w^\infty$  by  ${}^\infty w^\infty_{0, |w|-1} = w$  and  $\sigma^{|w|}({}^\infty w^\infty) = {}^\infty w^\infty$ .

A *cellular automaton* is a pair  $(\mathcal{A}, F)$  where  $F : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$  is a continuous function that commutes with  $\sigma$  (i.e.  $F \circ \sigma = \sigma \circ F$ ). Alternatively  $F$  is defined by a finite *neighbourhood*  $N \subset \mathbb{Z}$  and a local rule  $f : \mathcal{A}^N \rightarrow \mathcal{A}$  in the sense that  $F(x)_i = f((x_{i+j})_{j \in N})$ .

In the figures, we represent the time evolution of cellular automata starting from an initial configuration  $x \in \mathcal{A}^{\mathbb{Z}}$  by a two-dimensional *space-time diagram*  $(F^t(x)_i)_{t \in \mathbb{N}, i \in \mathbb{Z}}$ .

The *frequency* of a finite word  $u$  in a configuration  $x \in \mathcal{A}^{\mathbb{Z}}$  is defined as:

$$\text{freq}_x(u) = \limsup_{n \rightarrow \infty} \frac{1}{(2n+1)} \text{Card}\{i \in \{-n, \dots, n\} : x \in [u]_i\}.$$

### 1.2 Cyclic cellular automata

**Definition 1** (*n*-state cyclic cellular automaton).  $(\mathbb{Z}/n\mathbb{Z}, C_n)$  is the *n*-state cyclic cellular automaton defined on the neighbourhood  $N = \{-1, 0, 1\}$  by the local rule  $c_n$ :

$$c_n(u_{-1}, u_0, u_1) = \begin{cases} u_0 + 1 & \text{if } u_1 = u_0 + 1 \pmod{n} \quad \text{or} \quad u_{-1} = u_0 + 1 \pmod{n}, \\ u_0 & \text{otherwise.} \end{cases}$$

All operations concerning  $n$ -state cyclic automata are assumed to be modulo  $n$ .

As should be clear from Figure 2, the self-organisation is driven by borders between monochromatic regions behaving as particles. We call particles the factors  $ab$  of length 2 (with  $a \neq b$ ) in a configuration. Each particle moves “from predator to prey”, that is, left if  $a = b + 1$ , right if  $b = a + 1$ , and stays put otherwise. This motivates the following definitions:

**Positive particles**  $p_+ = \{ab : b = a + 1\}$ ;

**Negative particles**  $p_- = \{ab : b = a - 1\}$ ;

**Neutral particles**  $p_0 = \{ab : b \notin \{a - 1, a, a + 1\}\}$ .

and we write  $[p_+]_i$  as a shorthand for  $\bigcup_{ab \in p_+} [ab]_i$ : it means that a positive particle occurs at position  $i$ . Notice that  $p_0 = \emptyset$  for  $n = 3$ . Figure 1 illustrate the particle dynamics for  $n = 3$ .

### 1.3 Probability measures on $\mathcal{A}^{\mathbb{Z}}$

Let  $\mathfrak{B}$  be the Borel sigma-algebra of  $\mathcal{A}^{\mathbb{Z}}$ . Denote by  $\mathcal{M}(\mathcal{A}^{\mathbb{Z}})$  the set of probability measures on  $\mathcal{A}^{\mathbb{Z}}$  defined on the sigma-algebra  $\mathfrak{B}$ . Since the cylinders  $\{[u]_n : u \in \mathcal{A}^*, n \in \mathbb{Z}\}$  form a basis of the product topology on  $\mathcal{A}^{\mathbb{Z}}$ , a measure  $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$  is entirely characterised by the values  $\mu([u]_n)$ .

In this paper, we only consider  $\sigma$ -invariant probability measures, and therefore write  $\mu([u])$  instead of  $\mu([u]_n)$ .

#### Examples.

**Measures supported by a periodic orbit** For a word  $w \in \mathcal{A}^*$ , we define the  $\sigma$ -invariant measure supported by  ${}^\infty w {}^\infty$  by taking the mean of the Dirac measures  $\delta_{\infty w \infty}$  along its  $\sigma$ -orbit:

$$\delta_{\infty w \infty} = \frac{1}{|w|} \sum_{i \in [0, |w| - 1]} \sigma^i(\delta_{\infty w \infty}).$$

When  $w$  is a single letter  $a$ , we obtain the measure supported on the single monochromatic configuration  ${}^\infty a {}^\infty$ .

**Bernoulli measure** Let  $v = (v_a)_{a \in \mathcal{A}}$  be a vector of real numbers such that  $0 \leq v_a \leq 1$  for all  $a \in \mathcal{A}$  and  $\sum_{a \in \mathcal{A}} v_a = 1$ . Let  $\beta_v$  be the discrete probability distribution on  $\mathcal{A}$  such that  $\beta_v(a) = v_a$  for all  $a \in \mathcal{A}$  (a generalisation of the standard Bernoulli law with  $n$  outcomes).

The associated *Bernoulli measure*  $\mathcal{B}er_v$  on  $\mathcal{A}^{\mathbb{Z}}$  is the product measure  $\prod_{i \in \mathbb{Z}} \beta_v$ , that is,

$$\mathcal{B}er_v([u_0 \dots u_n]) = v_{u_0} \cdots v_{u_n} \quad \text{for all } u_0 \dots u_n \in \mathcal{A}^*.$$

In other words, each cell is drawn in an i.i.d. manner according to  $\beta_v$ . We denote  $\mathcal{B}er(\mathcal{A}^{\mathbb{Z}})$  the set of Bernoulli measures on  $\mathcal{A}^{\mathbb{Z}}$  with nonzero parameters  $(v_a)_{a \in \mathcal{A}}$ .

**Uniform measure** In particular, if we take  $v_a = \frac{1}{|\mathcal{A}|}$  for all  $a \in \mathcal{A}$  in the previous definition, we obtain the *uniform (Bernoulli) measure*  $\lambda$ .

The *image measure* of  $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}})$  by a cellular automaton  $(\mathcal{A}^{\mathbb{Z}}, F)$  is defined as  $F\mu(B) = \mu(F^{-1}(B))$  for all  $B \in \mathfrak{B}$ . This defines an action  $F : \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}}) \rightarrow \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}})$ .

We endow  $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}})$  with the *weak\* topology*: for a sequence  $(\mu_n)_{n \in \mathbb{N}} \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}})^{\mathbb{N}}$  and a measure  $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}})$ , we have  $\mu_n \xrightarrow[n \rightarrow \infty]{} \mu$  if, and only if:

$$\forall u \in \mathcal{A}^*, \mu_n([u]) \xrightarrow[n \rightarrow \infty]{} \mu([u]).$$

This topology makes  $F$  continuous and  $\mathcal{M}_\sigma(\mathcal{A}^\mathbb{Z})$  is compact.

A measure  $\mu \in \mathcal{M}_\sigma(\mathcal{A}^\mathbb{Z})$  is  $\sigma$ -ergodic if, for every subset  $S \subset \mathcal{A}^\mathbb{Z}$  such that  $\sigma(S) = S$   $\mu$ -almost everywhere, we have  $\mu(S) = 0$  or  $1$ . The set of  $\sigma$ -ergodic measures is denoted  $\mathcal{M}_\sigma^{erg}(\mathcal{A}^\mathbb{Z})$ . In particular, all examples above are  $\sigma$ -ergodic and the image of a  $\sigma$ -ergodic measure under the action of a cellular automaton is  $\sigma$ -ergodic.

As an example of a non- $\sigma$ -ergodic measure, consider the average of two Dirac measures  $\frac{1}{2}(\delta_{\infty 0 \infty} + \delta_{\infty 1 \infty})$  (the set  $\{\infty 0 \infty\}$  is  $\sigma$ -invariant and has measure  $\frac{1}{2}$ ).

We make use of the following corollary to Birkhoff's theorem:

**Corollary 2.** *Let  $\mu \in \mathcal{M}_\sigma^{erg}(\mathcal{A}^\mathbb{Z})$  and  $u \in \mathcal{A}^*$ . Then:*

$$\forall_\mu x \in \mathcal{A}^\mathbb{Z}, \text{freq}(u, x) = \mu([u])$$

where  $\forall_\mu x$  means for  $\mu$ -almost all  $x$  (that is, for all  $x$  in some set of measure 1).

## 2 Known and new results

The first main result on one-dimensional cyclic cellular automata is the following. It consider the values of the sequence  $(C_n^t(x)_0)_{t \in \mathbb{N}}$  for an arbitrary site (here 0) when iterating  $C_n$  on a uniform random configuration.

**Theorem 3** (Fisch [8], Theorem 1). *Draw an initial configuration  $x$  according to  $\lambda$  be the uniform Bernoulli measure on  $(\mathbb{Z}/n\mathbb{Z})^\mathbb{Z}$ , and consider the sequence  $(C_n^t(x)_0)_{t \in \mathbb{N}}$ . Then:*

- If  $n \leq 4$ , then  $\lambda(x \in \mathcal{A}^\mathbb{Z} : C_n^t(x)_0 \text{ changes infinitely often as } t \rightarrow \infty) = 1$  ( $x_0$  fluctuates);
- If  $n \geq 5$ , then  $\lambda(x \in \mathcal{A}^\mathbb{Z} : C_n^t(x)_0 \text{ changes finitely often as } t \rightarrow \infty) = 1$  ( $x_0$  fixates).

Since changes of values corresponds to times when a particle  $p_+$  or  $p_-$  crosses the column, this result can be interpreted in terms of limit measures. For  $n \geq 5$ , some particles  $p_-$  (“walls”) survive asymptotically ( $C_n^t \lambda([p_-]) \not\rightarrow 0$ ) and delimit walled areas where the remaining moving particles  $p_-$  or  $p_+$  cannot enter; for  $n \leq 4$ ,  $C_n^t \lambda([p_-]) \rightarrow 0$  and moving particles cross each column infinitely often. This result can be intuited on Figure 2.

Notice that the previous result only applies when the initial measure is uniform. The following result follows from [14], Corollary 1; it is weaker but applies on the much more general setting of  $\sigma$ -ergodic measures:

**Proposition 4.** *Let  $\mu$  be any  $\sigma$ -ergodic measure on  $(\mathbb{Z}/n\mathbb{Z})^\mathbb{Z}$ . Then at least two of the following are true:*

- $C_n^t \mu([p_+]) \rightarrow 0$ ;
- $C_n^t \mu([p_-]) \rightarrow 0$ ;
- $C_n^t \mu([p_=]) \rightarrow 0$ .

For Bernoulli measures, the state of the art is summed up in the following proposition:

**Proposition 5.** *If  $\mu$  is a Bernoulli measure, then  $C_n^t \mu([p_+]) \rightarrow 0$  and  $C_n^t \mu([p_-]) \rightarrow 0$ . In particular, if  $n = 3$ , any limit point of  $(C_n^t \mu)_{t \in \mathbb{N}}$  is a convex combination of the measures  $\delta_{\infty i \infty}$ .*

*If furthermore  $\mu = \lambda$  the Bernoulli uniform measure, the unique limit point of  $(C_n^t \mu)_{t \in \mathbb{N}}$  is  $\frac{1}{n} \sum_i \delta_{\infty i \infty}$  for both cases  $n \in \{3, 4\}$ .*

*Proof.* In the case where  $\mu$  is a Bernoulli measure, or more generally a measure invariant by the mirror involution  $\gamma : (x_i)_{i \in \mathbb{Z}} \mapsto (x_{-i})_{i \in \mathbb{Z}}$ , the only possible non-zero case is  $C_n^t \mu([p_-]) \not\rightarrow 0$ . Indeed, since  $C_n \circ \gamma = \gamma \circ C_n$  and the mirror operation sends  $p_+$  to  $p_-$  and conversely, we have  $C_n^t \mu([p_+]) = C_n^t \mu([p_-])$ .

For  $n = 3$ , since  $p_= = \emptyset$ , there is asymptotically no particle at all, so all limit points must be some convex combination of the measures  $\delta_{\infty i \infty}$ .

If furthermore  $\mu = \lambda$  the Bernoulli uniform measure, Theorem 3 gives us  $C_n^t \mu([p_-]) \rightarrow 0$  in the case  $n = 4$  as well. Since this measure is invariant by the state-transposing operation  $\kappa : (x_i)_{i \in \mathbb{Z}} \mapsto (x_i + 1)_{i \in \mathbb{Z}}$  and  $C_n \circ \kappa = \kappa \circ C_n$ , the unique limit point is  $\frac{1}{n} \sum_i \delta_{\infty i \infty}$  for both cases  $n \in \{3, 4\}$ .  $\square$

The previous results, fluctuation in particular, can be interpreted in terms of heteroclinic cycles. For  $\lambda$ -almost every configuration  $x$ , no state ever dominates the whole space in the sense that (by Corollary 1)  $\text{freq}(i, C_3^t(x)) = C_3^t \lambda[i] \rightarrow \frac{1}{3}$  for every state  $i \in \mathbb{Z}/3\mathbb{Z}$  (we use the fact that the image under  $C_3$  of a  $\sigma$ -ergodic measure is  $\sigma$ -ergodic).

However, Proposition 5 implies that, for any fixed window  $[-N, N]$  and  $\lambda$ -almost every  $x$ ,  $C_3^t(x)_{[-N, N]}$  is monochromatic (in topological terms, it is close to one of the  ${}^\infty i^\infty$ ,  $i \in \mathbb{Z}/3\mathbb{Z}$ ) except for some sequence of times of zero density. Theorem 3 further shows that  $C_3^t(x)$  does not converge to one of the  ${}^\infty i^\infty$  as  $t \rightarrow \infty$ , but that the window keeps changing state (as a particle crosses the central column), less and less often, letting each state dominate the central window in turn. In this sense, the 3-state cyclic cellular automaton exhibits heteroclinic cycles in local regions.

Our main new result determines the unique limit point for non-uniform Bernoulli measures:

**Theorem 6** (Main result).

Let  $\mu$  be a Bernoulli measure on  $(\mathbb{Z}/3\mathbb{Z})^{\mathbb{Z}}$  with nonzero parameters  $(p_0, p_1, p_2)$ . Then:

$$C_3^t \mu \xrightarrow[t \rightarrow \infty]{} p_2 \delta_{\infty 0^\infty} + p_0 \delta_{\infty 1^\infty} + p_1 \delta_{\infty 2^\infty}.$$

Theorem 6 can be interpreted as follows. Draw an initial configuration according to a Bernoulli measure with nonzero parameters  $(p_0, p_1, p_2)$ , and consider a fixed arbitrary window  $[-N, N]$ . By Proposition 5, the probability that  $C_3^t(x)_{[-N, N]}$  contains at least two different states (i.e. a particle) tends to 0. Theorem 6 further shows that the probability that  $C_3^t(x)_{[-N, N]} = i^{2N+1}$  for  $i \in \mathbb{Z}/3\mathbb{Z}$  tends to  $p_{i-1}$  as  $t$  tends to infinity.

Remarkably, the parameters of the limit measure are a simple cyclic permutation of the parameters of the initial Bernoulli measure: each state  $i$  reaches asymptotically the initial frequency of its “prey”  $p_{i-1}$ . This is illustrated on Figure 3.



Figure 3: The 3-state cyclic cellular automaton iterated on an initial configuration drawn according to the Bernoulli measure of parameters  $(\frac{1}{10}, \frac{3}{10}, \frac{6}{10})$ . State 0 =  $\square$ , initially present with probability  $\frac{1}{10}$  at each site, is present in the topmost configuration with a frequency approximately  $\frac{6}{10}$ .

### 3 Proof of the main result

This section is dedicated to the proof of Theorem 6. Since we already know by Proposition 5 that any limit point of  $(C_3^t \mu)_{t \in \mathbb{N}}$  is a convex combination of  $\delta_{\infty 0^\infty}$ ,  $\delta_{\infty 1^\infty}$  and  $\delta_{\infty 2^\infty}$ , it remains to show that for each  $i$ ,  $\mu(C_3^t(x)_0 = i) \rightarrow p_{i-1}$ .

In this section, we use the one-sided version of  $C_3$  to simplify proofs:

**Definition 7** (One-sided cyclic CA).  $(\mathbb{Z}/3\mathbb{Z}, C_{3+})$  is the one-sided 3-state cyclic cellular automaton defined on the neighbourhood  $N = \{0, 1\}$  by the local rule:

$$c_{3+}(u_0, u_1) = \begin{cases} u_0 + 1 & \text{if } u_1 = u_0 + 1 \pmod{3}, \\ u_0 & \text{otherwise.} \end{cases}$$

A (computer assisted) proof by enumeration of all  $3^3 = 27$  factors of length 3, shows that

$$C_3 = C_{3+}^2 \circ \sigma.$$

Hence proving Theorem 6 on  $C_{3+}$  implies a similar result on  $C_3$ .

The proof proceeds in 4 steps:

**Section 3.1** where we associate a random walk to each configuration and relate the properties of this random walk to the orbit of the configuration under  $C_{3+}$ ;

**Section 3.2** where we translate Theorem 6 on the random walk and establish the objects that will be relevant to the proof.

**Section 3.3** where we introduce a second random walk “embedded” in the previous one, which is symmetric (hence easier to analyse) and captures its large-scale behaviour.

**Section 3.4** where we bring back the results from the embedded walk to the initial walk and bring all tools together to conclude the proof.

### 3.1 Random walk associated with a configuration

In this section, we introduce tools to turn the study of the dynamics of the 3-state cyclic automaton, in particular of  $C_{3+}^t(x)_0$  (defined above), into the study of some random walk built from the initial configuration  $x$ .

**Definition 8.** To a configuration  $x \in \{0, 1, 2\}^{\mathbb{Z}}$  we associate a random walk  $W[x] := (w_i)_{i \in \mathbb{Z}}$  on  $\mathbb{Z}$  such that  $w_0 \in \{0, 1, 2\}$  and made up of steps in  $\{-1, 0, 1\}$  as follows:

- $w_0 = x_0$ ,
- for all  $i \geq 0$ ,  $w_{i+1}$  is the value in  $\{w_i - 1, w_i, w_i + 1\}$  such that  $w_{i+1} \equiv x_{i+1} \pmod{3}$ ,
- and for  $i \leq 0$ ,  $w_{i-1}$  is the value in  $\{w_i - 1, w_i, w_i + 1\}$  such that  $w_{i-1} \equiv x_{i-1} \pmod{3}$ .

and this encoding is an injection.

Figure 4 provides an example of this encoding (black configuration to black walk).

We denote by  $W_{[a,b]}[x] := (w_a, w_{a+1}, \dots, w_b)$  the positions of the walk on  $\mathbb{Z}$  from time  $a$  to time  $b$ . Notice that we call time in the context of the random walk what corresponds to space in the configuration  $x$ , which is different from the time corresponding to the iteration of cellular automaton. Context should make clear which notion of time we refer to.

The main interest of this correspondence is to deduce the state of a cell after  $n$  iterations from the maximal height in the first  $n$  steps of the walk associated to the initial configuration  $W[x]$ :

**Proposition 9.** For  $n \geq 0$ , we have

$$C_{3+}^n(x)_0 = (\max W_{[0,n]}[x]) \pmod{3}.$$

*Proof.* We will prove that the iterations of  $C_{3+}$  keep for  $t = 0, \dots, n - 1$  the following invariant:

$$(\max W_{[0, n-t]}[C_{3+}^t(x)]) \bmod 3 = (\max W_{[0, n-(t+1)]}[C_{3+}^{t+1}(x)]) \bmod 3.$$

When this invariant is expressed for  $t = 0$  and  $t = n$ , we deduce the expected identity:

$$(\max W_{[0, n]}[x]) \bmod 3 = (\max W_{[0, 0]}[C_{3+}^n(x)]) \bmod 3 = \max\{C_{3+}^n(x)_0\} \bmod 3 = C_{3+}^n(x)_0.$$

We prove this invariant in the case  $t = 0$  and any  $n \geq 1$ . The cases  $t > 0$  follow by replacing  $x := C_{3+}^t(x)$  and  $n := n - t$ .

We describe how to obtain  $W_{[0, n-1]}[C_{3+}(x)] = (w'_i)_{i=0, \dots, n} =: w'$  from  $W_{[0, n]}[x] = (w_i)_{i=0, \dots, n+1} =: w$  by a 3-step transformation:  $w \mapsto w^1 := (w_i^1)_{i=0, \dots, n+1} \mapsto w^2 := (w_i^2)_{i=0, \dots, n+1} \mapsto w'$ . Each of these steps, illustrated in Figure 4, preserves the invariant.

The preserved invariant is the height (modulo 3) of the first maximum of paths represented by  $\circ, \circ, \circ$ , respectively.

The initial configuration given by  $(x_i)_{i=0 \dots 12}$  is also represented by the black path  $(w_i)_{i=0 \dots 12}$ .

Step 1: Its image  $(y_i)_{i=0 \dots 12}$  by  $c_{3+}$  is represented by the red path  $(w_i^1)_{i=0 \dots 12}$ , obtained from the black path by rising in parallel the lower vertex of each rise.

Step 2: Since this red path starts at  $3 \notin \{0, 1, 2\}$ , it is shifted downwards by 3, leading to the blue path  $(w_i^2)_{i=0 \dots 12}$  starting at  $0 \in \{0, 1, 2\}$  that is associated with  $(y_i)_{i=0 \dots 12}$ .

Step 3: The green area indicates that the last vertex of the blue path is deleted, yielding  $(w'_i)_{i=0 \dots 11}$ .

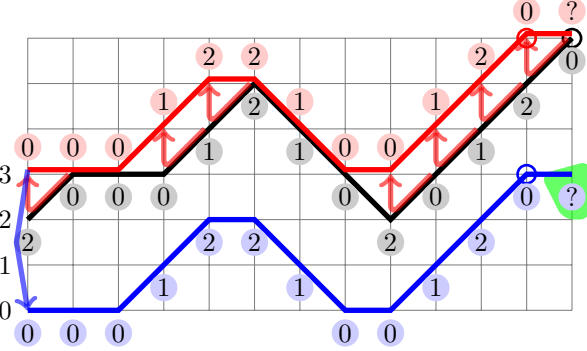


Figure 4: The 3-step transformation preserving the invariant.

By definition, for  $0 \leq i \leq n - 1$ ,  $C_{3+}(x)_i = c_{3+}(x_i, x_{i+1})$ . We notice that cases where  $x_i$  becomes  $x_i + 1 \bmod 3$  are exactly the steps  $+1$  in the walk  $w$  (factors 01, 12 or 20 in  $x$ ).

- Step 1 For  $0 \leq i < n$ , define  $w_i^1 := w_i + 1$  if  $w_i < w_{i+1}$  and  $w_i^1 := w_i$  otherwise. In addition  $w_n^1 := w_n$ . Notice that  $w^1$  is also a walk on  $\mathbb{Z}$  made up of steps  $\{-1, 0, +1\}$ . The maximal height is preserved since any visit at maximal height in  $w$  may not be followed by a  $+1$  step.
- Step 2 The only case where  $w_0^1 = 3 \notin \{0, 1, 2\}$  is when  $w_0 = 2 \bmod 3$  and  $w_1 = 0 \bmod 3$ . In this case, for  $i = 0, \dots, n$ , define  $w_i^2 = w_i^1 - 3$  and  $w^2 = w^1$  otherwise. The maximal height may be decreased by 3, but it is preserved mod 3.
- Step 3 We remove the last position  $w_n^2$  in the walk  $w^2$  to obtain  $W_{[0, n-1]}[C_{3+}(x)]$ . This preserves the maximal height: if  $w_n$  was the first visit to the maximal height, the first step ensures that  $w_{n-1}^1 = w_{n-1} + 1 = w_n = w_n^1$ . Therefore  $w_{n-1}^2 = w_n^2$ , so  $w_n^2$  can not be the first occurrence of the maximal height and can be safely removed.

□



### 3.2 Analysing the random walk

Recall that the measure on the initial configuration is the Bernoulli measure  $\mu$  of parameters  $(p_0, p_1, p_2)$ . From this and the bijection with walks on  $\mathbb{Z}$  we forget its relationship with  $x$  to study it for itself as a random variable, directly sampling  $W[x]$  as follows (each choice being independent):

- $w_0 = j$  with probability  $p_j$  for  $j \in \{0, 1, 2\}$ ,
- then for all  $i \geq 0$ , with probability  $p_j$ ,  $w_{i+1}$  is the value in  $\{w_i - 1, w_i, w_i + 1\}$  such that  $w_{i+1} \equiv j \pmod{3}$  for  $j \in \{0, 1, 2\}$ ,
- and for  $i \leq 0$ , with probability  $p_j$ ,  $w_{i-1}$  is the value in  $\{w_i - 1, w_i, w_i + 1\}$  such that  $w_{i-1} \equiv j \pmod{3}$  for  $j \in \{0, 1, 2\}$ .

Similarly, we can sample the factor  $W_{[0,n]}[x] = (w_i)_{i=0,\dots,n}$  by assuming by convention that  $w_{-1} = 1$  to ensure that  $w_0 \in \{0, 1, 2\}$ . Then the only rule is  $w_{i+1} \in \{w_i - 1, w_i, w_i + 1\}$  with probability  $p_{w_{i+1} \pmod{3}}$ , independently from other choices.

In the proofs, we will need such walks starting from an arbitrary  $k \in \mathbb{Z}$ . Formally, define  $\mathcal{W}_{k,n}$  a random walk on  $\mathbb{Z}$  of length  $n \in \mathbb{N}$  and starting from  $k \in \mathbb{Z}$  as:

$$\mathcal{W}_{k,n} := (W_t)_{t=0\dots n} \quad \text{where} \quad \begin{array}{l} W_0 = k \\ W_t = W_{t-1} - 1 + ((Z_t - W_{t-1} + 1) \pmod{3}) \quad \text{for } t = 1 \dots n, \end{array}$$

where  $(Z_t)_{t=1\dots n}$  are i.i.d. random variables in  $\mathbb{Z}/3\mathbb{Z} := \{0, 1, 2\}$  for all  $t$ , and  $\mathbb{P}(Z_t = j) = p_j$  for all  $j \in \mathbb{Z}/3\mathbb{Z}$ .

**Theorem 10** (Main result of this section). *For any  $i \in \mathbb{Z}/3\mathbb{Z}$  and any  $k \in \mathbb{Z}$ ,*

$$\lim_{n \rightarrow +\infty} \mathbb{P}(\max(\mathcal{W}_{k,n}) \pmod{3} = i) = p_{(i-1) \pmod{3}}$$

where  $\max(\mathcal{W}_{k,n}) := \max_{t=0\dots n} W_t$ .

We first consider the case  $i = 0$  (and  $k = 0$ ), *i.e.*  $\lim_{n \rightarrow +\infty} \mathbb{P}(\max(\mathcal{W}_{0,n}) \pmod{3} = 0) = p_2$ ; the other cases will follow.

Our proof proceeds by conditioning this event to the length  $m$  of the 3-tail (defined below), and describing the probability in terms of the value of other probabilities  $(P_{k,m}^{<H})_{k,m,H}$  (also defined below).

**Definition 11** (Record, tail). *A record occurs at time  $t'$  in the random walk  $\mathcal{W}_{k,n}$  if  $W_{t'} = \max_{t=0\dots t'} W_t$ ; notice a walk can have multiple records  $t_i$  sharing the same value  $W_{t_i}$ .*

*The  $h$ -tail of  $\mathcal{W}_{k,n}$  is the suffix  $W_{[t' \dots n]}$ , where  $t'$  is the last occurrence of a record  $W_{t'}$  divisible by  $h$ ; the  $h$ -tail for  $h > 1$  may not exist.*

We make use of the 3-tail and the 1-tail in the proof. The length of the 3-tail  $\text{tail}_3(\mathcal{W}_{k,n}) := n - t'$  is usually denoted by  $m$ .

**Notations:**

- $\mathbb{W}_{k,n}^{<H}$  is the set of walks on  $n \in \mathbb{N}$  steps which start from  $k \in \mathbb{Z}$  and remain on values strictly lower than  $H \in \mathbb{Z}$ .
- $P_{k,n}^{<H}$  is the probability that a random walk  $\mathcal{W}_{k,n}$  belongs to  $\mathbb{W}_{k,n}^{<H}$ .

**Proposition 12** (Description conditioned by 3-tail). *For any  $n \in \mathbb{N}$  and any possible 3-tail length  $m \geq 1$ , we have:*

$$\mathbb{P}(\max(\mathcal{W}_{k,n}) \pmod{3} = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) = m) = p_2 K_m$$

where  $K_m := \frac{P_{-1,m-1}^{<0}}{P_{-1,m}^{<0}}$ .

*Proof.* By the definition of conditional probability:

$$\mathbb{P}(\max(\mathcal{W}_{k,n}) \bmod 3 = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) = m) = \frac{\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m \text{ and } \max(\mathcal{W}_{k,n}) \bmod 3 = 0)}{\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m)}.$$

We now evaluate the denominator and then the numerator of the right-hand side.

Evaluation of  $\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m)$ :

By definition  $\text{tail}_3(\mathcal{W}_{k,n}) = m > 1$  implies that  $W_{n-m} = 3H$  is the last record divisible by 3 in  $\mathcal{W}_{k,n}$ , and that  $W_{n-m+1}$  exists and is in the 3-tail. Assume therefore that  $W_{n-m} = 3H$  for some  $H$ . Since  $m \geq 1$ , we may discuss the possible values of  $W_{n-m+1} \in \{3H-1, 3H, 3H+1\}$  for any walk of  $\mathcal{W}_{k,n}$ . We identify below which of these values are allowed.

- If  $W_{n-m+1} = 3H$  (with probability  $p_0$ ), then  $W_{n-m}$  is not the last record divisible by 3, a contradiction.
- If  $W_{n-m+1} = 3H+1$  (with probability  $p_1$ ), future visits of height  $3H$  will not be a new record, so  $W_{n-m}$  is the last record divisible by 3 if and only if the walk never reaches  $3H+3$  (the next record divisible by 3). This corresponds to  $W_{[n-m+1, \dots, n]} \in \mathbb{W}_{3H+1, m-1}^{<3H+3}$  happening with probability  $P_{3H+1, m-1}^{<3H+3}$ .
- If  $W_{n-m+1} = 3H-1$  (with probability  $p_2$ ), the next visit of height  $3H$  would be a new occurrence of a record divisible by 3, so  $W_{n-m}$  is the last record divisible by 3 if and only if the walk never reaches again  $3H$ . This corresponds to  $W_{[n-m+1, \dots, n]} \in \mathbb{W}_{3H-1, m-1}^{<3H}$  happening with probability  $P_{3H-1, m-1}^{<3H}$ .

In the definition of  $\mathcal{W}_{k,n}$  it appears that any realisation  $(W_t)_t \in \mathcal{W}_{k,n}$  can be translated into  $(W_t + 3T)_t$  for any  $T \in \mathbb{Z}$  without changing the probability of steps. This implies that for any  $3T \in 3\mathbb{Z}$  we have

$$\forall H, k \in \mathbb{Z}^2, \forall n \in \mathbb{N}, P_{k,n}^{<H} = P_{k+3T,n}^{<H+3T}.$$

Therefore the probabilities in the previous discussion does not depend on  $3H$ . By the law of total probability:

$$\begin{aligned} \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m) &= \sum_{H \in \mathbb{Z}} \mathbb{P}(W_{n-m} = 3H) \cdot \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m \mid W_{n-m} = 3H) \\ &= \sum_{H \in \mathbb{Z}} \mathbb{P}(W_{n-m} = 3H) \cdot (p_1 P_{3H+1, m-1}^{<3H+3} + p_2 P_{3H-1, m-1}^{<3H}) \\ &= \mathbb{P}(W_{n-m} = 0 \bmod 3) \cdot (p_1 P_{-2, m-1}^{<0} + p_2 P_{-1, m-1}^{<0}) \\ &= p_0 \cdot (p_1 P_{-2, m-1}^{<0} + p_2 P_{-1, m-1}^{<0}) \end{aligned}$$

where we use the fact that any step in the walk leads to a height divisible by 3 with probability  $p_0$ , regardless of the previous position.

The first step of a walk in  $\mathbb{W}_{-1, m}^{<0}$  leads from  $-1$  to either  $-1$  or  $-2$ , so we get the following partition:

$$\mathbb{W}_{-1, m}^{<0} = \{-1\} \times \mathbb{W}_{-1, m-1}^{<0} \cup \{-1\} \times \mathbb{W}_{-2, m-1}^{<0}$$

In terms of probabilities this identity turns into:

$$P_{-1, m}^{<0} = p_2 P_{-1, m-1}^{<0} + p_1 P_{-2, m-1}^{<0}.$$

Hence  $\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m) = p_0 P_{-1, m}^{<0}$ .

Evaluation of  $\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m \text{ and } \max(\mathcal{W}_{k,n}) \bmod 3 = 0)$ : We reconsider the previous discussion on  $W_{n-m+1}$  under the additional condition  $\max(\mathcal{W}_{k,n}) \bmod 3 = 0$ . Again assume that  $W_{n-m} = 3H$

- $W_{n-m+1} = 3H$  is still impossible by definition of the 3-tail.

- $W_{m-n+1} = 3H+1$  is impossible since it would imply the maximum is strictly greater than  $3H$ . However, no new record divisible by three is now reachable by definition of 3-tail.

Hence the only possible choice for  $W_{n-m+1}$  is  $3H - 1$ , and the walk must avoid  $3H$  from time  $n - m + 1$  onwards: this happens with probability  $p_2 P_{3H-1, m-1}^{<3H} = p_2 P_{-1, m-1}^{<0}$ . Therefore the numerator of  $K_m$  is

$$\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) = m \text{ and } \max(\mathcal{W}_{k,n}) \bmod 3 = 0) = p_0 p_2 P_{-1, m-1}^{<0}$$

Since by assumption  $p_0 \neq 0$ , they cancel out in the expression for  $K_m$  to give the desired result.  $\square$

### 3.3 The embedded walk

The remainder of this section proves that when  $n$  tends to infinity, the length  $m$  of the 3-tail also tends to infinity with high probability (Lemma 17) and  $K_m$  tends to 1 (Lemma 18). Proposition 12 then leads to Theorem 10 for  $i = 0$  asymptotically.

We use a factorisation of the walk into factors forming a symmetric  $\{+3, -3\}$  random walk on  $3\mathbb{Z}$ , that can be scaled to be the usual symmetric  $\{+1, -1\}$  random walk on  $\mathbb{Z}$ .

**Definition 13** (Embedded walk). *Define greedily and recursively a sequence of times  $(t_j)_{j=0, \dots, J-1}$  as follows:*

- $W_{t_0}$  is the first occurrence of a value divisible by 3 in  $\mathcal{W}_{k,n}$ , if any.
- given  $(t_j)_{j=0, \dots, i-1}$ ,  $t_i$  is the next time where  $W_{t_i}$  is divisible by 3 and distinct from  $W_{t_{i-1}}$ , if any such  $t_i \leq n$  exists.

From this sequence we define the embedded random walk  $\mathcal{W}_{k,n}^s = (W_{t_j})_{j=0, \dots, J-1}$  whose length  $J$ , also denoted  $|\mathcal{W}_{k,n}^s|$ , is random.

An example of embedded walk is given in Figure 5.

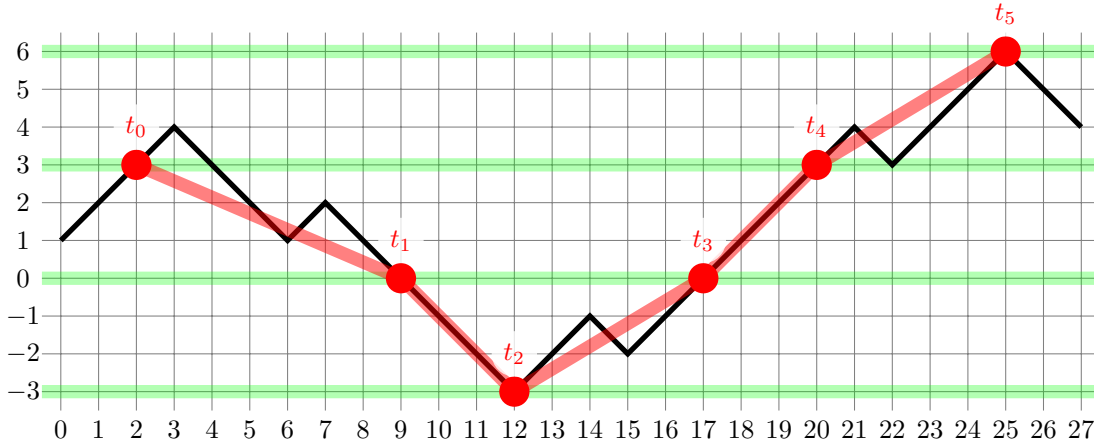


Figure 5: A (black) walk  $\mathcal{W}_{1,27}$  of length 27 and its embedded (red) walk  $\mathcal{W}_{1,27}^s$  of length  $J = 6$ .

First we show that the embedded walk on  $3\mathbb{Z}$  is symmetric (Lemma 14), then that its length is at least linear in the size of the original walk with high probability (Lemma 15) and finally provide an upper bound on the probability of a small 1-tail in this symmetric  $\{+1, -1\}$  walk on  $\mathbb{Z}$  (Lemma 16).

**Lemma 14** (The embedded walk on  $3\mathbb{Z}$  is symmetric). *For any value  $J \geq 0$  and any  $n \geq 0$ , the walk  $\mathcal{W}_{k,n}^s$  conditioned by  $|\mathcal{W}_{k,n}^s| = J$  is a symmetric random walk of  $J - 1$  steps<sup>1</sup>  $\{-3, +3\}$  on  $3\mathbb{Z}$  with independent steps.*

<sup>1</sup>By convention, 0 steps means one vertex,  $-1$  steps means no vertex at all.

*Proof.* The vertex  $W_{t_0}$  exists since the length of the embedded walk is conditioned to be  $J \geq 0$ . We prove that  $\mathcal{W}_{k,n}^s$  is symmetric, that is, for any  $i$  such that  $1 \leq i \leq J - 1$ :

$$\mathbb{P}(W_{t_i} = W_{t_{i-1}} + 3) = \mathbb{P}(W_{t_i} = W_{t_{i-1}} - 3)$$

The proof of this equality relies on an involution on the atomic events  $(Z_k)_{k=t_{i-1}+1, \dots, t_i} =: Z_{]t_{i-1}, t_i]}$  describing the two probabilities. We illustrate this involution on Figure 6.

First denote, for any  $a < b$ ,  $\bar{Z}_{]a, b]}$   $= (Z_{a+b+1-k})_{k=a+1, \dots, b}$  the mirror image of the sequence  $Z_{]a, b]}$ . Second, for any sequence  $Z_{]t_{i-1}, t_i]}$  defining an event where  $W_{t_i} = W_{t_{i-1}} + 3$ , define  $s_i$  as the index of the last occurrence of the event  $Z_t = 0$  (that is,  $W_t = 0 \pmod 3$ ) before  $t_i$ ; notice  $s_i \geq t_{i-1}$  since  $Z_{t_{i-1}} = 0$ . It appears that the map

$$\tau : Z_{]t_{i-1}, t_i]} = Z_{]t_{i-1}, s_i]} Z_{]s_i, t_i-1]} Z_{t_i} \longrightarrow Z_{]t_{i-1}, s_i]} \bar{Z}_{]s_i, t_i-1]} Z_{t_i}$$

is an involution from the events where  $W_{t_i} = W_{t_{i-1}} + 3$  to the events where  $W_{t_i} = W_{t_{i-1}} - 3$  (see more details in Figure 6). Furthermore, since  $Z_{]a, b]}$  and  $\bar{Z}_{]a, b]}$  have the same probability,  $\tau$  preserves probabilities.  $\square$

We consider an initial walk (black then red) from  $t_{i-1}$  to  $t_i$  defined by the  $(Z_t)_{t=t_{i-1}+1, \dots, t_i} = (\mathbb{x})_t \cup (\mathbb{x}')_t$ . It defines a step  $+3$  in the embedded walk. The magenta interval starts at time  $s_i$ , which is the last occurrence of 0, and ends at time  $t_i$ . In the magenta interval,  $\tau$  transforms the (red)  $(\mathbb{x}')_t$  into the (blue)  $(\mathbb{x})_t$ .

Observe that the blue part is up to translation the symmetric of the red path via reflection with respect to a vertical axis and hence now defines a step from  $t_{i-1}$  to  $t_i$  in the embedded walk.

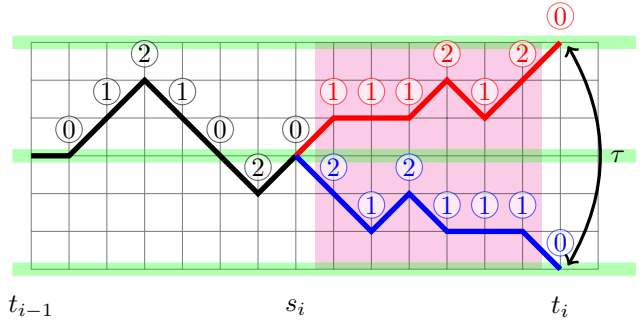


Figure 6: Involution used in the proof of symmetry of the embedded walk

The following lemma ensures that the length of the embedded walk grows almost surely at least linearly in  $n$ . This helps us later to convert bounds on 1-tail length in  $\frac{1}{3}\mathcal{W}_{k,n}^s$  into bounds on 3-tail length in  $\mathcal{W}_{k,n}$ .

**Lemma 15** (The embedded walk's length is almost always linear). *There exists  $E > 0$  and  $N \geq 0$ , such that for any  $\beta \in ]0, 1/(16E)[$  and  $n \geq N$  we have*

$$\mathbb{P}(|\mathcal{W}_{k,n}^s| < \beta n) < C(\beta) \exp(-D(\beta)n)$$

where  $C(\beta)$  and  $D(\beta) > 0$  are non-negative functions of  $\beta$  made explicit in the proof.

*Proof.* Split the variables  $(Z_i)_{i=1, \dots, n}$  used in the definition of the walk into independent factors of 4 variables  $((Z_{]4a, 4(a+1)]})_{a=0, \dots, \lfloor n/4 \rfloor}$ . Now define variables  $(X_a)_{a=0, \dots, \lfloor n/4 \rfloor}$  as follows:

$$X_a = \begin{cases} 1 & \text{if } Z_{]4a, 4(a+1)]} \in \{(0, 1, 2, 0), (0, 2, 1, 0)\} \\ 0 & \text{otherwise} \end{cases}$$

The motivation for this definition is that, as illustrated in Figure 7, each occurrence of a factor  $Z_{]a, a+4]} = (0, 1, 2, 0)$  or  $Z_{]a, a+4]} = (0, 2, 1, 0)$  in the atomic events  $(Z_i)_{i=1, \dots, m}$  implies the existence of at least one embedded vertex at  $t_i = a + 4$ . This is obviously a rough bound and many embedded vertices are missed.

It follows that  $|\mathcal{W}_{k,n}^s| \geq \sum_{a=0}^{\lfloor n/4 \rfloor} X_a$ . Notice that  $(X_a)_{a=0, \dots, \lfloor n/4 \rfloor}$  is a family of i.i.d Bernoulli variables of parameter  $q := 2p_0^2 p_1 p_2 < 1$ .

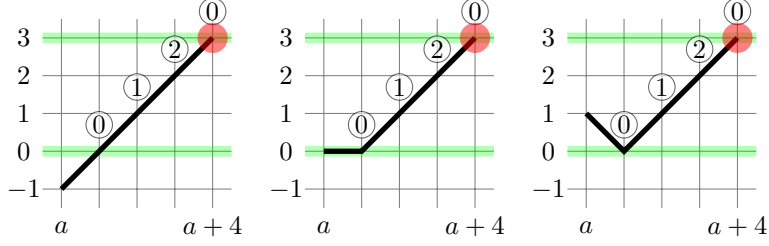


Figure 7: If  $Z_{[a, a+4]} = (0, 1, 2, 0)$  then an embedded step occurs at  $t_i = a + 4$  (and possibly another at  $t_{i-1} = a + 1$ )

Now define a time sequence  $(t'_i)_{i=0, \dots, k}$  recursively as:

$$t'_0 = \min\{a > 0 \mid X_a = 1\} \text{ and } t'_i + 1 = \min\{a > t'_i \mid X_a = 1\},$$

and  $t'_i$  is left undefined if the minimum is taken on an empty set. Furthermore put  $Y_i = t'_{i+1} - t'_i$ , if it is defined. Notice that if  $\sum_{i=0}^{\beta n - 1} Y_i \leq \lfloor n/4 \rfloor$ , then  $t'_{\beta n}$  is defined, which implies that  $\sum_{a=0}^{\lfloor n/4 \rfloor} X_a \geq \beta n$  and therefore  $|\mathcal{W}_{k,n}^s| \geq \beta n$ . In other words,

$$\mathbb{P}(|\mathcal{W}_{k,n}^s| < \beta n) \leq \mathbb{P}\left(\sum_{i=0}^{\beta n} Y_i \geq n/4\right)$$

Since  $Y_i$  is a family of i.i.d. geometric variables of parameter  $q = 2p_0^2 p_1 p_2 < 1$ , we use a bound on sums of geometric variables found in [16], Theorem 2.3. Denote  $E$  the (finite) expectation of each  $Y_i$ . Using the notation from this paper, we are in the particular case of i.i.d geometric variables of parameter  $p_i = 2p_0^2 p_1 p_2$ , so  $p_* := \min_i p_i = 2p_0^2 p_1 p_2$ . With our notations,  $\mu = 4\beta n E$  and  $\lambda \mu = n/4$ , so  $\lambda = \frac{1}{16\beta E}$ . Fixing any  $\beta \leq \frac{1}{16E}$ , we ensure that  $\lambda \geq 1$  for  $n$  large enough, so by *op. cit.* we have

$$\mathbb{P}\left(\sum_{i=0}^{\beta n - 1} Y_i \geq n/4\right) \leq \lambda^{-1} (1 - p_*)^{(\lambda - 1 - \ln \lambda)\mu} = 16\beta E (1 - p_*)^{(1 - 16\beta E + \ln 16\beta E)n} = C(\beta) \exp(-D(\beta)n)$$

where  $C(\beta) := 16\beta E$  and  $D(\beta) := \ln(1 - p_*)(16\beta E - 1 + \ln(16\beta E))$ . Since  $0 < 16\beta E < 1$ , the sign of  $(16\beta E - 1 + \ln(16\beta E)) = (16\beta - 1) + (\ln(16\beta E))$  is negative like the sign of  $\ln(1 - p_*)$  so  $D(\beta)$  is positive.  $\square$

To show that the 3-tail of  $\mathcal{W}_{k,n}$  is almost always at least of order  $n^{1/4}$ , we consider the 1-tail of the symmetric embedded walk  $\mathcal{W}_{k,n}^s$  of length  $|\mathcal{W}_{k,n}^s|$ . Denote  $\mathcal{S}^n = (S_t^n)_{t=0, \dots, n}$  the usual symmetric random walk on  $\mathbb{Z}$  made up of  $n$  steps  $\{-1, 1\}$  (where  $S_0^n = 0$ ). This object is well studied for convergence in law but we need precise bounds that we did not find in the literature, which is the object of the following lemma.

**Lemma 16** (Majoration on probability of a small 1-tail). *There exists  $K$  and  $N$  such that for any  $n \geq N$ , we have*

$$\mathbb{P}(\text{tail}_1(\mathcal{S}^n) \leq n^{1/4}) \leq K n^{-1/4}.$$

The proof of this lemma relies on classical combinatorics for walks on  $\mathbb{Z}$  with steps in  $\{+1, -1\}$ . To follow conventional notation, we represent them in this section as words on the alphabet  $\mathcal{S} := \{a, b\}$  under the convention  $a \mapsto +1, b \mapsto -1$ .

For a word  $w \in \mathcal{S}^*$  and any letter  $x \in \mathcal{S}$  we denote by  $|w|_x$  the number of occurrences of the letter  $x$  in the word  $w$ . A word  $w \in \mathcal{S}^*$  is a:

**bridge** if  $|w|_a = |w|_b$ . Bridges of length  $2n$  are counted by central binomial coefficient  $\binom{2n}{n} = \Theta(n^{-1/2}2^{2n})$  (and there are no bridges of odd length).

**Dyck word** if it is a bridge such that for any prefix  $p$ ,  $|p|_a \geq |p|_b$ . Dyck words of  $2n$  letters are counted by Catalan numbers  $\frac{1}{2n+1}\binom{2n+1}{n} = \Theta(n^{-3/2}2^{2n})$ .

**meander** if it is the prefix of a Dyck word. There is a classical folklore bijection<sup>2</sup> between meanders of length  $2n$  and bridges of length  $2n$  so meanders of even length  $2n$  are also counted by  $\binom{2n}{n} = \Theta(n^{-1/2}2^{2n})$ .

We can deduce the number of meanders of odd length, say  $2n + 1$ , as follows. Decompose this meander into two parts: the prefix of length  $2n$ , which is a meander; and the last letter. If the prefix is a Dyck word, then the last letter must be  $a$ ; otherwise, it may be  $a$  or  $b$ . Hence the number of meanders of length  $2n + 1$  is:

$$2\binom{2n}{n} - \frac{1}{2n+1}\binom{2n+1}{n} = \Theta(n^{-1/2}2^{2n}).$$

In a symmetric random walk, each letter  $a$  or  $b$  has the same independent probability  $1/2$ . The probability that any factor  $(S_t^n)_{t \in [\ell, \ell+k]}$  of  $\mathcal{S}^n$  is a meander is in  $\Theta(k^{-1/2})$ , while the probability that it is a Dyck word (assuming  $k$  is even) is in  $\Theta(k^{-3/2})$ .

*Proof.* We discuss the value of  $\text{tail}_1(\mathcal{S}^n)$ , with a generic example being illustrated in Figure 8.

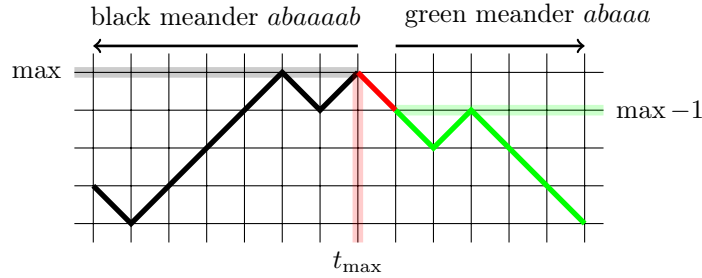


Figure 8: Generic decomposition of  $\mathcal{S}^n$  at its last maximum

Because  $n - \text{tail}_1(\mathcal{S}^n)$  is a maximum, the walk from time 0 to  $n - \text{tail}_1(\mathcal{S}^n)$  is a (black) meander when read in the reverse order.

- $\text{tail}_1(\mathcal{S}^n) = 0$  if and only if  $\mathcal{S}^n$  is itself a reverse meander, which happens with probability  $\Theta(n^{-1/2})$ .
- Otherwise, there is a (red)  $-1$  step just after time  $n - \text{tail}_1(\mathcal{S}^n)$ . The walk from time  $n - \text{tail}_1(\mathcal{S}^n) + 1$  to  $n$  is a (green) meander. For any fixed  $T < n$ , we have a bijection between walks where  $\text{tail}_1(\mathcal{S}^n) = T$  and pairs of two independents (black and green) meanders of length  $n - T - 1$  and  $T$ , respectively. Denoting  $M_i$  the probability that a word of length  $i$  is a meander, we have that the probability that  $n - \text{tail}_1(\mathcal{S}^n) = T$  is  $M_T \cdot M_{n-T-1}$ . Remind that  $M_i = \Theta(i^{-1/2})$ .

<sup>2</sup>A meander of even length ends at an even height  $2h$ : consider the  $h$  steps  $a$  starting at last passage at heights  $k = 0, \dots, h-1$ , turn each of them into a step  $b$ , this leads to a bridge. To recover the meander from the bridge, do the opposite on the steps that leads to the first visits at heights  $-1, \dots, -h$ , where  $h$  is the minimal height reached by the bridge.

Taking the two cases in consideration, it follows:

$$\begin{aligned}
\mathbb{P}(\text{tail}_1(\mathcal{S}^n) \leq n^{1/4}) &\leq \mathbb{P}(\text{tail}_1(\mathcal{S}^n) = 0) + \sum_{T=1}^{n^{1/4}} \mathbb{P}(\text{tail}_1(\mathcal{S}^n) = T) \\
&\leq M_n + \sum_{T=1}^{n^{1/4}} M_T M_{n-T-1} \\
&\leq \Theta(n^{-1/2}) + n^{1/4} \cdot \Theta(n^{-1/2}) \\
&\leq \Theta(n^{-1/4}),
\end{aligned}$$

where on the third line we used the fact that  $M_{n-T-1} = \Theta(n^{-1/2})$  and  $M_T \leq 1$ . Hence we have proved the lemma.  $\square$

### 3.4 Back to the main walk, and end of the proof

We now transfer the bound on the probability of a small 1-tail for the symmetric embedded walk, obtained in Lemma 16, to a similar upper bound for the 3-tail on the initial walk, using the probabilistic lower bound on the length of the embedded walk obtained in Lemma 15.

**Lemma 17** (Upper bound for the probability of a small 3-tail).  $\exists \gamma, \delta, N$  such that  $\forall n \geq N$ ,

$$\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4}) \leq \delta n^{-1/4}.$$

*Proof.* Let  $\beta$  a constant such that  $0 < \beta < 1/E$  where  $E \geq 1$  is the expectation of each variable  $(Y_i)_{i>0}$ , defined in the proof of Lemma 15. We distinguish the cases where the embedded symmetric walk  $\mathcal{W}_{k,n}^s$  has more or less than  $\beta n$  steps:

$$\begin{aligned}
\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4}) &= \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4} \mid |\mathcal{W}_{k,n}^s| < \beta n) \cdot \mathbb{P}(|\mathcal{W}_{k,n}^s| < \beta n) \\
&\quad + \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4} \mid |\mathcal{W}_{k,n}^s| \geq \beta n) \cdot \mathbb{P}(|\mathcal{W}_{k,n}^s| \geq \beta n)
\end{aligned}$$

For the first term, Lemma 15 ensures that the embedded walk has at least linear length with high probability:

$$\mathbb{P}(|\mathcal{W}_{k,n}^s| < \beta n) \leq C(\beta) \exp(-D(\beta)n).$$

For the second term, we use the (very rough) inclusion of events

$$\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4} \implies \text{tail}_1(\mathcal{W}_{k,n}^s) \leq \gamma n^{1/4}.$$

By Lemma 16, and fixing  $\gamma = \beta^{1/4}$ , the 1-tail is small with low probability:

$$\mathbb{P}(\text{tail}_1(\mathcal{W}_{k,n}^s) \leq \gamma n^{1/4} \mid |\mathcal{W}_{k,n}^s| \geq \beta n) \leq K(\beta n)^{-1/4}$$

Summing both contributions, and bounding the other expressions by 1,

$$\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \leq \gamma n^{1/4}) \leq C(\beta) \exp(-D(\beta)n) + K(\beta n)^{-1/4}.$$

For  $n$  larger than an appropriate  $N$ , we have the constant  $\delta$  allowing to bound this probability by  $\delta n^{-1/4}$  as expected.  $\square$

Remember that  $K_m := \frac{P_{-1,m-1}^{<0}}{P_{-1,m}^{<0}}$ , where  $P_{k,n}^{<H}$  is the probability that a random walk starting from  $k$  remains strictly below  $H$  during  $n$  steps.

**Lemma 18** (Bounds for  $K_m$ ).  $\exists \alpha, M$  such that  $\forall m \geq M, 1 \leq K_m \leq 1 + \frac{\alpha}{m}$ .

*Proof.* Reformulation of  $K_m$  and lower bound 1. Let  $E_m$  be the event:

$$\mathcal{W}_{-1,m} \in E_m \Leftrightarrow W_t < 0 \text{ for } 0 \leq t < m, \text{ and } W_m = 0.$$

Among the walks that remains strictly below 0 for  $m - 1$  steps, the event  $E_m$  exactly excludes those that remain strictly below 0 at the  $m$ -th step, so

$$\mathbb{P}(E_m) = P_{-1,m-1}^{<0} - P_{-1,m}^{<0}.$$

Dividing the previous equation by  $P_{-1,m}^{<0}$  we obtain that

$$K_m = 1 + \frac{\mathbb{P}(E_m)}{P_{-1,m-1}^{<0}}.$$

The lower bound  $1 \leq K_m$  follows.

Upper bound on  $K_m$  We now find an upper bound on the numerator  $\mathbb{P}(E_m)$ . First condition the event  $E_m$  by the possible length  $\ell$  of the embedded walk:

$$\mathbb{P}(E_m) = \sum_{\ell=0}^{+\infty} \mathbb{P}(E_m \text{ and } |\mathcal{W}_{-1,m}^s| = \ell).$$

Fix  $\beta$  such that  $0 < \beta < 1/E$ , and consider the terms such that  $\ell < \beta m$ . By Lemma 15, we have:

$$\sum_{\ell=0}^{\beta m-1} \mathbb{P}(E_m \text{ and } |\mathcal{W}_{-1,m}^s| = \ell) \leq \mathbb{P}(|\mathcal{W}_{-1,m}^s| < \beta m) \leq C(\beta) \exp(-D(\beta)m)$$

We now consider the remaining terms ( $\ell \geq \beta m$ ). First translate the event  $E_m$  on the embedded symmetric walk  $\mathcal{W}_{-1,m}^s$ . Define the flip as the morphism  $\{a, b\}^* \rightarrow \{a, b\}^*$  exchanging letters  $a$  and  $b$ . On the one hand, we have:

$$\mathcal{W}_{-1,m} \in E_m \text{ and } |\mathcal{W}_{-1,m}^s| = \ell \iff \mathcal{W}_{-1,m}^s \text{ is a flipped Dyck word of } \ell - 2 \text{ steps followed by a } +1 \text{ step}$$

since the first value divisible by 3 is  $-3$  and the first non-negative value divisible by 3 is  $W_m = 0$  that corresponds to the end of  $\mathcal{W}_{-1,m}^s$ .

On the other hand,

$$\mathcal{W}_{-1,m-1} \in \mathbb{W}_{-1,m-1}^{<0} \text{ and } |\mathcal{W}_{-1,m-1}^s| = \ell \iff \mathcal{W}_{-1,m}^s \text{ is a flipped meander of } \ell - 1 \text{ steps}$$

since avoiding 0 in  $\mathcal{W}_{-1,m-1}$  is by definition equivalent to the fact that  $\mathcal{W}_{-1,m-1}^s$  (starting in  $-3$ ) never visit 0.

Since the walk  $\mathcal{W}_{-1,m}^s$  is symmetric, we can use classical combinatorics giving  $1/2$  probability to each step to estimate the conditional probabilities:

$$\mathbb{P}(E_m \mid |\mathcal{W}_{-1,m}^s| = \ell) = 2^{-(\ell-1)} |\text{number of Dyck words of } \ell - 2 \text{ steps}|$$

since the last step must be a  $+1$  and

$$\mathbb{P}(\mathbb{W}_{-1,m-1}^{<0} \mid |\mathcal{W}_{-1,m-1}^s| = \ell) = 2^{-(\ell-1)} |\text{number of meanders of } \ell - 1 \text{ steps}|.$$

From the counting formula of Dyck words and meanders we get a classical result in combinatorics:

$$\frac{|\text{number of Dyck words of } \ell - 2 \text{ steps}|}{|\text{number of meanders of } \ell - 1 \text{ steps}|} = \begin{cases} \frac{1}{2^{\ell-3}} & \text{if } \ell = 0 \pmod{2} \\ 0 & \text{otherwise} \end{cases}$$



We turn this equality into a uniform upper bound using the assumption  $\ell \geq \beta m$ .

$$\frac{2^{-(\ell-1)} |\text{number of Dyck words of } \ell - 2 \text{ steps}|}{2^{-(\ell-1)} |\text{number of meanders of } \ell - 1 \text{ steps}|} \leq \frac{1}{2\ell - 3} \leq \frac{1}{2\beta m - 3}.$$

Summing up, we have

$$\begin{aligned} \sum_{\ell \geq \beta m} \mathbb{P}(E_m \mid |\mathcal{W}_{-1,m}^s| = \ell) \cdot \mathbb{P}(|\mathcal{W}_{-1,m}^s| = \ell) &\leq \frac{1}{2\beta m - 3} \sum_{\ell \geq \beta m} \mathbb{P}(\mathbb{W}_{-1,m-1}^{<0} \mid |\mathcal{W}_{-1,m-1}^s| = \ell) \cdot \mathbb{P}(|\mathcal{W}_{-1,m-1}^s| = \ell) \\ &\leq \frac{1}{2\beta m - 3} P_{-1,m-1}^{<0}, \end{aligned}$$

and therefore, adding the missing terms for  $\ell < \beta m$ :

$$\mathbb{P}(E_m) \leq C(\beta) \exp(-D(\beta)m) + \frac{1}{2\beta m - 3} P_{-1,m-1}^{<0}.$$

Now fix any  $\alpha < \frac{1}{2\beta}$ , and we get for  $m$  large enough the expected bound:

$$\mathbb{P}(E_m) \leq \frac{\alpha}{m} P_{-1,m-1}^{<0}.$$

From the decomposition of  $K_m$  we deduce that for  $m \geq M$ ,

$$K_m \leq 1 + \frac{\alpha}{m}.$$

□

*Proof. (Main theorem for  $i = 0$  and  $k = 0$ ).* Let  $m_{cut} := \gamma n^{1/4}$  (where  $\gamma$  is defined in Lemma 17). We discuss if  $\text{tail}_3(\mathcal{W}_{k,n})$  is at least  $m_{cut}$  or strictly lower to use our previous bounds:

$$\begin{aligned} \mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0) &= \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) \geq m_{cut}) \cdot \mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) \geq m_{cut}) \\ &\quad + \mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) < m_{cut}) \cdot \mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) < m_{cut}). \end{aligned}$$

We have from Lemma 17

$$\mathbb{P}(\text{tail}_3(\mathcal{W}_{k,n}) < m_{cut}) \leq \delta n^{-1/4}.$$

We know from Proposition 12 that

$$\mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) \geq m_{cut}) = p_2 K_{m_{cut}},$$

and Lemma 18 gives

$$1 \leq K_{m_{cut}} \leq 1 + \frac{\alpha}{\gamma n^{1/4}}$$

Together with the trivial bound  $0 \leq \mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0 \mid \text{tail}_3(\mathcal{W}_{k,n}) < m_{cut}) \leq 1$ , those bounds lead to:

$$(1 - \delta n^{-1/4}) p_2 \leq \mathbb{P}(\max \mathcal{W}_{k,n} \bmod 3 = 0) \leq \delta n^{-1/4} + p_2 \left( 1 + \frac{\alpha}{\gamma n^{1/4}} \right).$$

When  $n$  tends to infinity we obtain the expected limit  $p_2$ .

□

*Proof. (Main theorem for  $i \neq 0$ ).* The proof is similar as the case  $i = 0$ , the single difference being that the embedded symmetric walk corresponds to visits of  $3\mathbb{Z} + i$  instead of  $3\mathbb{Z} = 3\mathbb{Z} + 0$ . The change in the analysis corresponds to the path from 0 to the first occurrence of visit at an height equal to  $i$  modulo 3. Then we apply the map  $(p_0, p_1, p_2) \longrightarrow (p_{(i+0) \bmod 3}, p_{(i+1) \bmod 3}, p_{(i+2) \bmod 3})$  and obtain that the asymptotic probability of the state  $i$  is  $p_{(i-1) \bmod 3}$  as expected.

□

## 4 Conclusion

**Other models** It is natural to ask whether a similar phenomenon occurs in more complex models of cyclic dominance.

The easiest extension is to consider that each predator has a fixed probability  $p < 1$  to replace its prey (probabilistic version). Although the global behaviour seems similar (see Figure 9), the probability that a small region surrounded by a predator and a prey disappears may impact the early dynamics; noise has been shown to create the possibility of this kind of “unlucky extinctions” in non-spatial models [23]. Experimental evidence seems to indicate a much slower convergence and no obvious numerical relationship between the initial parameters and asymptotic probability.

In higher dimension, we are so far from a complete understanding of the dynamics and limit measure [7] that no conjecture seems possible.

We believe that the random walk approach (Proposition 9) can be adapted to more general prey/predator relationship. Consider the prey/predator graph, where the oriented edge  $i \rightarrow j$  means that  $i$  is a predator for  $j$ . Beyond the simple 3-state cyclic dominance, more complex predator/prey graphs have been observed in nature [30, 28, 22]. As a first example, we believe Proposition 9 holds on alphabets of size  $2k + 1$  where each state  $n$  has  $k$  predators  $n + 1, n + 2, \dots, n + k$  and  $k$  preys  $n - 1, \dots, n - k$  (modulo  $2k + 1$ ). The random walk has steps in  $\{-k, +k\}$ , with the same condition that  $W[x]_i$  must be equal to  $x_i$  modulo  $2k + 1$ .

The clearest limit to our approach is the presence of neutral particles, which can interact with other particles in ways that we cannot seem to describe in terms of a simple height function (as an example, in the 4-state cyclic cellular automaton, a neutral particle can turn a positive particle into a negative particle, or the opposite). The absence of neutral particle means that the prey/predator relationship corresponds to an orientation of the complete graph (a *tournament*).

**Max path preservation** We believe that a necessary and sufficient condition for Proposition 9 to hold is the following. Assume the alphabet is  $\mathbb{Z}/n\mathbb{Z}$  and denote  $a < b < c$  if, by incrementing  $a$  by one repeatedly, one reaches  $b$  before  $c$ . A complete graph orientation is *max path preserving* if, for any triplet of distinct vertices/species  $a < b < c$ , if  $a$  and  $b$  are predators for  $c$ , ( $a \rightarrow c$  and  $b \rightarrow c$ ) then  $b$  is a predator for  $a$  ( $b \rightarrow a$ ).

The definition of the walk  $W[x]$  becomes the following:

- $w_0 = x_0$
- if  $x_i = x_{i+1}$  then  $w_{i+1} = w_i$
- if  $x_i$  is a prey for  $x_{i+1}$  then  $w_{i+1}$  is the value equal to  $x_{i+1}$  modulo  $n$  in  $\{x_i + 1, \dots, x_i + n - 1\}$
- if  $x_i$  is a predator for  $x_{i+1}$  then  $w_{i+1}$  is the value equal to  $x_{i+1}$  modulo  $n$  in  $\{x_i - 1, \dots, x_i - (n - 1)\}$ .

To understand the max path preserving assumption, consider the following situation: a factor  $x_i x_{i+1} x_{i+2} = acb$  such that  $a < b < c$  and  $a$  and  $b$  are predators for  $c$ . We have  $W[x]_i > W[x]_{i+1} < W[x]_{i+2}$  and, since  $a < b < c$ ,  $W[x]_i < W[x]_{i+2}$ . Since this factor becomes  $ab$  at the next time step and the walk steps up, we want  $b$  to predate  $a$ .

Brute force enumeration up to  $n = 6$  suggests three families of prey-predator graphs for  $n$  species with the max-path preserving property:

- the  $n$  total orders compatible with the cyclic increments:  $k, k - 1, \dots, 0, n - 1, n - 2, \dots, k + 1$  for  $k$  from  $0$  to  $n - 1$  (the corresponding cellular automata is uninteresting as  $k - 1$  dominates every other state).
- some strongly connected prey/predator graphs where  $0, n - 1, n - 2, \dots, 1, 0$  forms an Hamiltonian cycle.
- some strongly connected prey/predator graphs where  $0, n - 1, n - 2, \dots, 1, 0$  is not an Hamiltonian cycle (but there is at least one Hamiltonian cycle, like for any strongly connected tournament).

The smallest example of the third family corresponds to the 3-state cyclic automaton where prey/predator relations between distinct species are reversed. Notice that the resulting walk on  $\mathbb{Z}$  consists of steps  $\pm 2$  instead of steps  $\pm 1$ . Of course we could consider the minimum on the walk on steps  $\pm 1$  instead, but we do not know if relabelling the prey/predator graphs of the third family according to one of the possibly many Hamiltonian cycles leads to a structure similar to the second family in general.

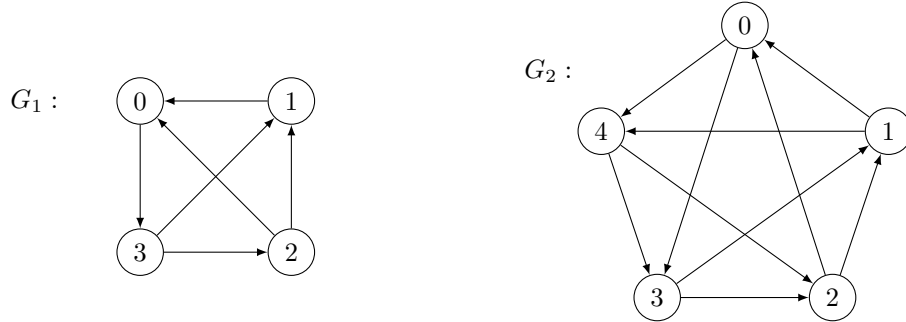
Up to  $n = 6$ , the last two families are counted by the Eulerian numbers (**A000295** in OEIS):

$$e_n := \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}.$$

We conjecture the following characterisation: any orientation in the second family is defined by selecting an odd number of vertices  $0 \leq c_1 < c_2 < \dots < c_{2k+1} \leq n-1$ , where  $3 \leq 2k+1 \leq n$  and the order corresponds to the numerotation in the hamiltonian cycle  $0, n-1, n-2, \dots, 1, 0$ . The convention is that for any state  $j$  in the cyclic interval  $]c_i, c_{i+1}[$ ,  $j$  is a predator for any state in  $[c_{i-k \bmod 2k+1}, j[$ . Since the number of possible cycles of  $2k+1$  vertices is counted by the binomial coefficient  $\binom{n}{2k+1}$ , this characterisation would imply the counting above. Empirically, all orientations defined by this conjectural characterisation satisfy the max path preservation up to  $n = 11$  species.

If Proposition 9 does generalise to these three graph families, it remains to check that the rest of the proof does as well. We expect a more systematic analysis using functional equations.

**Numerical results** In order to conjecture a relationship between the asymptotic probability of each state and the parameters of the initial Bernoulli measure, we performed numerical simulations of various cellular automata: the 3-state cyclic; the 3-state cyclic with probability (invasion rate)  $1/2$ ; the 4-state cyclic; and the two cellular automata corresponding to the predator/prey graphs represented below.



These results did not suggest any clear conjectural relationship, but we include them for possible future work.

For each cellular automaton  $F$ , we fixed values for the parameters of the initial Bernoulli measure that are distinct enough to be clearly distinguished. We computed the values of  $(F^t(x)_0)_{0 \leq t \leq 150}$  for 100000 random configurations  $x$ . The error margin (95% confidence) is  $\pm 0.002$  for all the values.

		3-state cyclic CA			3-state w/ proba 1/2			4-state cyclic CA			
states		0	1	2	0	1	2	0	1	2	3
parameters		0.1	0.3	0.6	0.1	0.3	0.6	0.05	0.15	0.3	0.5
time	0	0.100	0.301	0.599	0.100	0.300	0.600	0.050	0.149	0.300	0.500
	50	0.550	0.087	0.362	0.760	0.087	0.153	0.065	0.578	0.018	0.339
	100	0.565	0.090	0.345	0.780	0.147	0.074	0.122	0.615	0.019	0.354
	150	0.572	0.091	0.336	0.742	0.198	0.060	0.005	0.615	0.123	0.368

		4-state (graph $G_1$ )				5-state (graph $G_2$ )				
states		0	1	2	3	0	1	2	3	4
parameters		0.05	0.15	0.3	0.5	0.025	0.075	0.15	0.3	0.45
time	0	0.051	0.149	0.302	0.498	0.025	0.076	0.152	0.298	0.450
	50	0.369	0.009	0.040	0.580	0.169	0.354	0.026	0.061	0.390
	100	0.407	0.009	0.040	0.544	0.165	0.376	0.028	0.061	0.370
	150	0.424	0.010	0.041	0.525	0.165	0.385	0.027	0.061	0.362

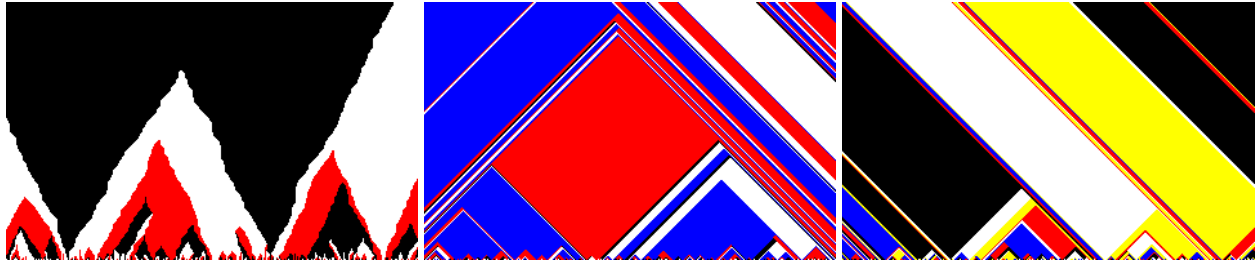


Figure 9: Left to right, the 3-state cyclic cellular automaton with invasion rate  $1/2$ , and the cellular automata corresponding to graphs  $G_1$  and  $G_2$ .

## Acknowledgements

This work benefits of discussions between Jean-François Marckert and the two authors. The first author is grateful towards Antony Quas for his help and remarks on an earlier version of the paper.

## References

- [1] Murray E. Alexander and Seyed M. Moghadas. Bifurcation analysis of an SIRS epidemic model with generalized incidence. *SIAM Journal on Applied Mathematics*, 65(5):1794–1816, 2005.
- [2] Per Bak, Kan Chen, and Chao Tang. A forest-fire model and some thoughts on turbulence. *Physics Letters A*, 147(5-6):297–300, 1990.
- [3] Vladimir Belitsky and Pablo A. Ferrari. Ballistic annihilation and deterministic surface growth. *Journal of Statistical Physics*, 80:517–543, 1995.
- [4] Maury Bramson and David Griffeath. Flux and fixation in cyclic particle systems. *The Annals of Probability*, 17(1):26–45, 1989.
- [5] Duncan D. Cameron, Andy White, and Janis Antonovics. Parasite–grass–forb interactions and rock–paper–scissor dynamics: predicting the effects of the parasitic plant *Rhinanthus minor* on host plant communities. *Journal of Ecology*, 97(6):1311–1319, 2009.
- [6] Rick Durrett and Simon Levin. Spatial aspects of interspecific competition. *Theoretical Population Biology*, 53(1):30–43, 1998.
- [7] Robert Fisch. Cyclic cellular automata and related processes. *Physica D: Nonlinear Phenomena*, 45(1):19–25, 1990.
- [8] Robert Fisch. The one-dimensional cyclic cellular automaton: a system with deterministic dynamics that emulates an interacting particle system with stochastic dynamics. *Journal of Theoretical Probability*, 3(2):311–338, 1990.

- [9] Robert Fisch. Clustering in the one-dimensional three-color cyclic cellular automaton. *The Annals of Probability*, 20:1528–1548, 1992.
- [10] Eric Foxall and Hanbaek Lyu. Clustering in the three and four color cyclic particle systems in one dimension. *Journal of Statistical Physics*, 171(3):1–14, 2018.
- [11] Marcus Frean and Edward R. Abraham. Rock–scissors–paper and the survival of the weakest. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1474):1323–1327, 2001.
- [12] Christian Guill, Barbara Drossel, Wolfram Just, and Eddy Carmack. A three-species model explaining cyclic dominance of Pacific salmon. *Journal of Theoretical Biology*, 276(1):16–21, 2011.
- [13] Christoph Hauert, Silvia De Monte, Josef Hofbauer, and Karl Sigmund. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science*, 296(5570):1129–1132, 2002.
- [14] Benjamin Hellouin de Menibus and Mathieu Sablik. Self-organisation in cellular automata with coalescent particles: qualitative and quantitative approaches. *Journal of Statistical Physics*, 167(5):1180–1220, 2017.
- [15] Lorens A. Imhof, Drew Fudenberg, and Martin A. Nowak. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences*, 102(31):10797–10800, 2005.
- [16] Svante Janson. Tail bounds for sums of geometric and exponential variables. *Statistics & Probability Letters*, 135:1–6, 2018.
- [17] Benjamin Kerr, Margaret A. Riley, Marcus W. Feldman, and Brendan J. M. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature*, 418(6894):171, 2002.
- [18] Benjamin C. Kirkup and Margaret A. Riley. Antibiotic-mediated antagonism leads to a bacterial game of rock–paper–scissors in vivo. *Nature*, 428(6981):412, 2004.
- [19] Benoit Laslier, Jean-François Laslier, et al. Reinforcement learning from comparisons: Three alternatives are enough, two are not. *The Annals of Applied Probability*, 27(5):2907–2925, 2017.
- [20] Hanbaek Lyu and David Sivakoff. Persistence of sums of correlated increments and clustering in cellular automata. *Stochastic Processes and their Applications*, 2018.
- [21] Robert M. May and Warren J. Leonard. Nonlinear aspects of competition between three species. *SIAM Journal on Applied Mathematics*, 29(2):243–253, 1975.
- [22] Daniel S. Maynard, Mark A. Bradford, Daniel L. Lindner, Linda T. A. van Diepen, Serita D. Frey, Jessie A. Glaeser, and Thomas W. Crowther. Diversity begets diversity in competition for space. *Nature ecology & evolution*, 1(6):0156, 2017.
- [23] Tobias Reichenbach, Mauro Mobilia, and Erwin Frey. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Physical Review E*, 74(5):051907, 2006.
- [24] Dirk Semmann, Hans-Jürgen Krambeck, and Manfred Milinski. Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature*, 425(6956):390, 2003.
- [25] Cosma Rohilla Shalizi and Kristina Lisa Shalizi. Quantifying self-organization in cyclic cellular automata. In *Noise in Complex Systems and Stochastic Dynamics*, volume 5114, pages 108–118. International Society for Optics and Photonics, 2003.
- [26] Karl Sigmund. Oscillations in the evolution of reciprocity. *Journal of Theoretical Biology*, 137:21–26, 1989.

- [27] Barry Sinervo and Curt M. Lively. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature*, 380(6571):240, 1996.
- [28] György Szabó and Tamás Czárán. Phase transition in a spatial Lotka-Volterra model. *Physical Review E*, 63(6):061904, 2001.
- [29] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics Reports*, 446(4-6):97–216, 2007.
- [30] Attila Szolnoki, Mauro Mobilia, Luo-Luo Jiang, Bartosz Szczesny, Alastair M. Rucklidge, and Matjaž Perc. Cyclic dominance in evolutionary games: a review. *Journal of the Royal Society Interface*, 11(100):20140735, 2014.
- [31] Kei-ichi Tainaka. Lattice model for the Lotka-Volterra system. *Journal of the Physical Society of Japan*, 57(8):2588–2590, 1988.
- [32] Kei-ichi Tainaka. Paradoxical effect in a three-candidate voter model. *Physics Letters A*, 176(5):303–306, 1993.