HAL
open science

Characterisation of limit measures of higher-dimensional cellular automata

Martin Delacourt, Benjamin Hellouin de Menibus

To cite this version:

Martin Delacourt, Benjamin Hellouin de Menibus. Characterisation of limit measures of higher-dimensional cellular automata. Theory of Computing Systems, 2017, 61 (4), pp.1178-1213. 10.1007/s00224-017-9753-1 . hal-01719833

HAL Id: hal-01719833

https://hal.science/hal-01719833v1

Submitted on 28 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TOCS manuscript No.
(will be inserted by the editor)

Characterisation of limit measures of higher-dimensional cellular automata

Martin Delacourt · Benjamin Hellouin de Menibus

Abstract We consider the typical asymptotic behaviour of cellular automata of higher dimension (≥ 2). That is, we take an initial configuration at random according to a Bernoulli (i.i.d) probability measure, iterate some cellular automaton, and consider the (set of) limit probability measure(s) as $t \to \infty$. In this paper, we prove that limit measures that can be reached by higher-dimensional cellular automata are completely characterised by computability conditions, as in the one-dimensional case. This implies that cellular automata have the same variety and complexity of typical asymptotic behaviours as Turing machines, and that any nontrivial property in this regard is undecidable (Rice-type theorem). These results extend to connected sets of limit measures and Cesàro mean convergence. The main tool is the implementation of arbitrary computation in the time evolution of a cellular automata in such a way that it emerges and self-organises from a random configuration.

Keywords Symbolic dynamics · Cellular automata · Limit measure · Multidimensional · Computability

Cellular automata are discrete dynamical systems defined by a local rule, introduced in the 40s by John von Neumann [14]. They model a large variety of discrete systems and are linked with various areas of mathematics and computer science, in particular computation theory, complex systems, ergodic theory and combinatorics.

One of the main catalysts of the study of cellular automata was their surprisingly complex and organised behaviours, even when iterated on configurations with no particular structure (e.g. chosen at random). To formalise these observations, many authors tried to describe their asymptotic behaviour by considering the limit set, which is the set of configurations that can be reached after arbitrarily many steps. These sets were shown to have potentially high computational complexity [13,1], and any nontrivial property on them is undecidable [10]. These observations built a bridge between the variety of dynamical behaviours and the computational content of the model. Nevertheless, the problem of characterising which sets can be limit sets of CA remains open.

Research partially supported by the FONDECYT Postdoctorado Proyecto 3130496 and by Basal project No. PFB-03 CMM, Universidad de Chile

Martin Delacourt (Corresponding author)
CMM, Universidad de Chile, CNRS, Av. Blanco Encalada 2120, Santiago, Chile

Benjamin Hellouin de Menibus
CMM, Universidad de Chile, CNRS, Av. Blanco Encalada 2120, Santiago, Chile
Departamento de Matemáticas, Universidad Andres Bello, Chile

In 2000, Kůrka and Maass argued that limit sets did not provide a good description of empirical observations and introduced instead a measure-theoretical version of these sets [12]. The idea of μ-limit sets is to choose the initial configuration at random, according to some probability measure μ, to iterate the cellular automaton on this configuration and to consider all patterns whose probability to appear does not tend to 0. In the one-dimensional case, this approach yielded similar results of high complexity and undecidability [4,3,5,2]. Although these two families of results appear similar and both require sophisticated constructions inside cellular automata, they provide insight about different kinds of dynamics (topological vs. measure-theoretical) and computational power (deterministic vs. probabilistic).

In [9], H. and Sablik extended this approach to consider the limit probability measure(s). Still in the one-dimensional case, they provided a computational characterisation of the limit measures reachable by cellular automata, generalising the previous results.

This article is an extended version of [6]. In *op.cit*, we aimed at extending the previous results to the two-dimensional setting. More precisely, we characterised all subshifts that can be μ-limit sets of CA when μ is the uniform Bernoulli measure. The proof works by an explicit construction inspired by the one-dimensional constructions of [2,9], although the higher dimensional setting has many specific challenges. In the present article, this two-dimensional construction is generalised to a d-dimensional space for any $d > 2$; furthermore, through a more careful analysis, we are able to characterise reachable limit measures, which is a more general result. As a corollary, we obtain an undecidability result on properties of limit measures, and cover as well Cesàro mean convergence and the case where the limit measure is not unique.

Section 1 is devoted to general definitions. In Section 2, we introduce more specific computability tools, and in particular computability restrictions on possible limit probability measures. In Section 3, we describe, process by process, the main technical construction that is the core of the proof of our results. Section 4 contains the main results that are the corollaries of this construction, mainly:

Theorem 1 (Main result) *The measures $\nu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ for which there exist:*

- *an alphabet $\mathcal{B} \supset \mathcal{A}$,*
- *a cellular automaton $F : \mathcal{B}^{\mathbb{Z}^d} \to \mathcal{B}^{\mathbb{Z}^d}$, and*
- *a non-degenerate Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{B}^{\mathbb{Z}^d})$*

such that $F^t\mu \xrightarrow[t\to\infty]{} \nu$, are exactly the limit-computable measures.

This result holds even when the initial measure is chosen to be uniform. As a corollary, we show that any nontrivial property on limit measures is undecidable. These result extend to connected sets of limit measures and convergence in Cesàro mean.

1 Definitions

1.1 Symbols, configurations and cellular automata

Let $\mathcal{A}$ be a finite set of symbols called *alphabet*. For $d > 0$, let $\mathcal{A}^{\mathbb{Z}^d}$ be the space of d-dimensional *configurations*.

On $\mathbb{Z}^d$, we define the basis vectors $e_i = (\delta_i(k))_{0<k\leq d}$ (Kronecker deltas), that is, the vector worth 0 on all coordinates except the i-th where it is worth 1. Denote $\mathcal{U}nit(d) = \{\sum_{1\leq j\leq d} \delta_j e_j \neq 0 : \forall j, \delta_j \in \{-1, 0, 1\}\}$ and $\mathcal{H}yp(d)$ the set of hyperplanes that have a normal vector in $\mathcal{U}nit(d)$; these hyperplanes have a basis of $d-1$ vectors in $\mathcal{U}nit(d)$.

We will use the following distances between points of $\mathbb{Z}^d$:

$$\forall x, y \in \mathbb{Z}^d, \qquad d_\infty(x,y) = \max_{1\le i\le d} |x_i - y_i| \quad \text{and} \quad d_1(x,y) = \sum_{1\le i\le d} |x_i - y_i|.$$

An ∞-path is a sequence of points $z_1, \dots, z_k$ such that $d_\infty(z_i, z_{i+1}) = 1$ for any i. An ∞-connected set is a subset of $\mathbb{Z}^d$ such that any pair of points are connected by an ∞-path. Define similarly 1-paths and 1-connected subsets.

If we endow $\mathcal{A}^{\mathbb{Z}^d}$ with the product topology of the discrete topology on $\mathcal{A}$, then $\mathcal{A}^{\mathbb{Z}^d}$ is a Cantor set (compact, perfect and totally disconnected). This topology is also metrisable, for example using the *Cantor metric*:

$$\forall x, y \in \mathcal{A}^{\mathbb{Z}^d},\ d_C(x,y) = 2^{-\delta_{x,y}} \quad \text{where} \quad \delta_{x,y} = \min\{||i||_\infty \ :\ x_i \neq y_i\}.$$

For a subset $U \subset \mathbb{Z}^d$, denote $x_U \in \mathcal{A}^U$ the restriction of x to U. Denote $\mathcal{A}^* = \bigcup_{\substack{U\subset\mathbb{Z}^d \\ finite}} \mathcal{A}^U$ the set of finite *patterns*. For a pattern $w \in \mathcal{A}^U$, denote its *support* $\mathrm{supp}(w) = U$, and its *dimension* is the smallest d such that $\mathrm{supp}(w)$ is isomorphic to a subset of $\mathbb{Z}^d$. We say a pattern is *cubic*, respectively *rectangular*, if its support is a $d-$*cube*, resp. a $d-$*box* (Cartesian product of intervals).

The *cylinder* defined by a pattern $u \in \mathcal{A}^*$ and a position $i \in \mathbb{Z}^d$ is $[u]_i = \{x \in \mathcal{A}^{\mathbb{Z}^d} \ : \ x_{i+\mathrm{supp}(u)} = u\}$. For simplicity we sometimes write $[u]$ for $[u]_{(0,\dots,0)}$.

The *shift map*, or *shift*, is defined as:

$$\forall i \in \mathbb{Z}^d,\ \sigma_i(x) = (x_{i+j})_{j\in\mathbb{Z}^d}.$$

A *subshift* is a closed subset of $\mathcal{A}^{\mathbb{Z}^d}$ invariant under all shifts. Given a cubic pattern $u \in \mathcal{A}^{[0,n-1]^d}$, define the *periodic configuration* ${}^\infty u^\infty$ by ${}^\infty u^\infty_{[0,n-1]^d} = u$ and $\sigma^n_{e_k}({}^\infty u^\infty) = {}^\infty u^\infty$ for every $k \in [1,d]$.

A *cellular automaton* (or CA) is a continuous function $F : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ that commutes with all shifts ($F \circ \sigma_{e_k} = \sigma_{e_k} \circ F$ for every k). By the Curtis-Hedlund-Lyndon theorem [7], it can be defined equivalently as a function $F(x) = (f((x_j)_{j\in(i+\mathcal{N})}))_{i\in\mathbb{Z}^d}$ where $\mathcal{N} \subset \mathbb{Z}^d$ is a finite set called *neighbourhood* and $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$ is called a *local rule*.

1.2 Probability measures on $\mathcal{A}^{\mathbb{Z}^d}$

Let $\mathfrak{B}$ be the Borel σ-algebra of $\mathcal{A}^{\mathbb{Z}^d}$ and $\mathcal{M}(\mathcal{A}^{\mathbb{Z}^d})$ the set of probability measures on $\mathcal{A}^{\mathbb{Z}^d}$ defined on the σ-algebra $\mathfrak{B}$. In this article, we focus on $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ the set of *σ-invariant probability measures* on $\mathcal{A}^{\mathbb{Z}^d}$, that is to say, the measures μ such that $\mu(\sigma_k^{-1}(B)) = \mu(B)$ for all $B \in \mathfrak{B}$ and $k \in \mathbb{Z}^d$. Cylinders corresponding to finite patterns form a base of the topology. Since $\mu([u]_i) = \mu([u])$ for any $i \in \mathbb{Z}^d$ and $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$, μ is entirely characterised by $\{\mu([u]) \ : \ u \in \mathcal{A}^*\}$; actually, considering only cubic patterns is enough.

We endow $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ with the *weak** (or *weak convergence*) topology:

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

In the weak* topology, $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is compact and metrisable. A metric is defined by

$$d_{\mathcal{M}}(\mu,\nu) = \sum_{n\in\mathbb{N}} \frac{1}{2^n} \max_{u\in\mathcal{A}^{[0,n]^d}} |\mu([u]) - \nu([u])|.$$

Define the *ball* centred on $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ of radius $\varepsilon > 0$ as

$$\mathbf{B}(\mu, \varepsilon) = \left\{ \nu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d}) : d_{\mathcal{M}}(\mu, \nu) \leq \varepsilon \right\}.$$

Let us define some examples that we use throughout the article.

The *Bernoulli measure* μ_λ associated to some vector $\lambda = (\lambda_a) \in [0;1]^{\mathcal{A}}$ satisfying $\sum_{a \in \mathcal{A}} \lambda_a = 1$ is defined by

$$\mu_\lambda([u_0 \dots u_n]) = \lambda_{u_0} \cdots \lambda_{u_n} \qquad \text{for all } u_0 \dots u_n \in \mathcal{A}^*.$$

The *Dirac measure* supported by $x \in \mathcal{A}^{\mathbb{Z}^d}$ is defined as $\delta_x(A) = \mathbf{1}_{x \in A}$. Generally δ_x is not σ-invariant. However, for any cubic pattern $w \in \mathcal{A}^{[0,n]^d}$, it is possible to define the *σ-invariant measure supported by* ${}^\infty w^\infty$ by taking the mean of the Dirac measures on the orbit under σ:

$$\widehat{\delta_w} = \frac{1}{(n+1)^d} \sum_{i \in [0,n]^d} \delta_{\sigma_i({}^\infty w^\infty)}.$$

The set of measures $\left\{ \widehat{\delta_w} : w \in \mathcal{A}^* \right\}$ is dense in $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ [15].

1.2.1 Action of a cellular automaton on $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ and limit measures

Let $F : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ be a cellular automaton and $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$. Define the *image measure* $F_*\mu$ by $F_*\mu(A) = \mu(F^{-1}(A))$ for all $A \in \mathfrak{B}$. Since F is σ-invariant, that is to say $F \circ \sigma = \sigma \circ F$, one deduces that $F_*(\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})) \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$. This defines a continuous application $F_* : \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d}) \to \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$.

We consider in particular $F_*^t\mu$ the iterated image of μ by F_*. Since $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is compact in the weak$*$ topology, the sequence $(F_*^t\mu)_{t \in \mathbb{N}}$ admits a set of limit points denoted $\mathcal{V}(F, \mu)$ and called the *μ-limit set of measures of F*. When $\mathcal{V}(F, \mu)$ is a singleton, i.e. when $F_*^t\mu \underset{n \to \infty}{\longrightarrow} \nu$, we say ν is the *limit measure* of F starting on μ.

This name stems from the standard μ-limit set of F defined as $\bigcup_{\nu \in \mathcal{V}(F,\mu)} \operatorname{supp}(\nu)$.

2 Computability

We now introduce the computability notions that are needed to state our main results. This exposition is very similar to the one found in [9], which was later expanded in [8], for the one-dimensional case. Indeed, most of the definitions and proofs only rely on the fact that the space is metric and separable, properties for which the increase in dimension is irrelevant. We omit those proofs that can be obtained by a straightforward substitution ($\mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}^d}$) from the proofs found in these references.

2.1 Turing machines

Turing machines are a standard and robust tool to define the computability of mathematical operations. In the usual model, they have access to a one-dimensional, one- or two-sided infinite memory tape. In order to simplify some constructions, we consider in this article that the tape is d-dimensional and infinite in all directions. This does not affect the computing power of the model.

A *Turing machine* $\mathcal{TM} = (Q, \Gamma, \#, q_0, \delta, Q_F)$ is defined by:

- Γ a finite alphabet, with a blank symbol $\# \notin \Gamma$. Initially, a d-dimensional infinite memory tape is filled with $\#$, except for a finite region (the input), and a computing head is located at coordinate $(0, \dots, 0)$;
- Q a finite set of states, with an initial state $q_0 \in Q$;
- $\delta : (Q \cup \#) \times \Gamma \to (Q \cup \#) \times \Gamma \times \{\pm e_i\}_{1 \leq i \leq d}$ the transition function. Given the current state and the letter it reads on the tape — which depends on its current position — the function returns the new state, the letter to be written on the tape at current position, and the vector by which the head moves.
- $Q_F \subset Q$ the set of final states — when a final state is reached, the computation stops and the output is the contents of the tape.

A function $f : \mathcal{A}^* \to \mathcal{A}^*$ is *computable* if there exists a Turing machine working on an alphabet $\Gamma \supset \mathcal{A}$ that, on any input $w \in \mathcal{A}^*$, eventually stops and outputs $f(w)$.

2.2 Computability of functions mapping countable sets

To generalise this definition to functions mapping arbitrary countable sets $X \to Y$, we need to define an *encoding*, that is, an alphabet $\mathcal{A}_X$ together with a bijection between X and some subset of $\mathcal{A}_X^*$, and similarly for Y. Then the computability of a function $X \to Y$ is defined up to some encoding. However, in practice, reasonable encodings yield the same notion of computability. To simplify notations, we fix some canonical encodings for the rest of the paper :

$\mathbb{Z}$ (or $\mathbb{N}$): Take $\mathcal{A}_{\mathbb{Z}} = \{0, 1\}$ and encode an element $k \in \mathbb{Z}$ as its binary expansion surrounded by blank symbols.

Product $X \times Y$: Take $\mathcal{A}_{X \times Y} = \mathcal{A}_X \times \mathcal{A}_Y$ and encode (x, y) as the product of encodings for x and y.

Using this last case, we define a canonical encoding for $\mathbb{Q}$ as the canonical encoding for $\mathbb{N} \times \mathbb{Z}$, up to the bijection $\frac{p}{q} \mapsto (p, q)$ (with p, q irreducible).

Furthermore, we define the computability of a set $K \subset X$ as the computability of the function $1_K : X \to \mathbb{N}$.

2.3 Computability of probability measures

As we mentioned above, a probability measure $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is entirely described by the value of $\mu([u])$ for all $u \in \mathcal{A}^*$. In other words, an element of $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is described by a function $\mathcal{A}^* \to \mathbb{R}$. Since $\mathbb{R}$ is not countable, the standard ways to define notions of computability is to consider approximations by elements of $\mathbb{Q}$.

A measure $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is *computable* if there exists a computable function $f : \mathcal{A}^* \times \mathbb{N} \to \mathbb{Q}$ such that

$$|\mu([u]) - f(u, n)| < 2^{-n} \qquad \text{for all } u \in \mathcal{A}^* \text{ and } n \in \mathbb{N}.$$

It is *limit-computable* if there exists a computable function $f : \mathcal{A}^* \times \mathbb{N} \to \mathbb{Q}$ such that

$$|\mu([u]) - f(u, n)| \underset{n \to \infty}{\longrightarrow} 0 \qquad \text{for all } u \in \mathcal{A}^*.$$

Additionally we define the notion of a *uniformly computable sequence*. Informally, it means that a sequence of objects can be computed by a single algorithm which, given $n \in \mathbb{N}$ as input, returns a description of the n-th object of the sequence.

Formally, a sequence of measures $(\mu_i)_{i\in\mathbb{N}}$ is *uniformly computable* iff there exists a computable function $f : \mathbb{N} \times \mathcal{A}^* \times \mathbb{N} \to \mathbb{Q}$ such that:

$$|\mu_i([u]) - f(n, u, i)| < 2^{-n} \qquad \text{for all } u \in \mathcal{A}^* \text{ and } n, i \in \mathbb{N}^2.$$

It is easy to see that the limit of a uniformly computable sequence of measures is limit-computable (but not necessarily computable since the rate of convergence of μ_i to μ is not known).

Proposition 1 (Approximation by measures supported by periodic orbits) *These notions can be defined in another equivalent way:*

(i) A measure $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is computable if and only if there exists a computable function $f : \mathbb{N} \to \mathcal{A}^*$ such that $d_{\mathcal{M}}\left(\mu, \widehat{\delta_{f(n)}}\right) \leq 2^{-n}$ for all $n \in \mathbb{N}$.

(ii) A measure $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is limit-computable if and only if there exists a computable function $f : \mathbb{N} \to \mathcal{A}^*$ such that $\lim_{n\to\infty} \widehat{\delta_{f(n)}} = \mu$.

Notice the parallel with the definition of the computability of a real: in both cases, an object is computable if it is approximated by a uniformly computable sequence of elements taken from a dense subset ($\mathbb{Q}$ and the measures supported by periodic orbits, respectively) with a known rate of convergence.

2.4 Action of a cellular automaton on computable measures

Proposition 2 (First computability obstruction) *Let $F : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ be a cellular automaton and $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ be a computable measure. Then $(F_*^t\mu)_{t\in\mathbb{N}}$ is a uniformly computable sequence of measures. In particular, if $F_*^t\mu \underset{t\to\infty}{\longrightarrow} \nu$ then ν is limit-computable.*

In general, $F_*^t\mu$ does not have a single limit point, but a compact set of accumulation points. To obtain a similar obstruction, we extend our computability definitions to those objects.

2.5 Compact sets in computable analysis

Extending naively the definition for countable sets using the characteristic function does not work since the set of inputs would not be countable. Instead, we use a general definition for metric spaces that possess a countable dense subset, $(\widehat{\delta_w})_{w\in\mathcal{A}^*}$ in the case of $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$.

A closed set $\mathcal{K} \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is *computable* if the set $\left\{(w, r) \in \mathcal{A}^* \times \mathbb{Q} \ : \ \overline{\mathbf{B}(\widehat{\delta_w}, r)} \cap \mathcal{K} \neq \emptyset\right\}$ is computable (as a countable set), that is, if its characteristic function is.

However, the set of limit points of the sequence $(F_*^t\mu)_{t\in\mathbb{N}}$, where μ is computable, is not necessarily computable (or even limit-computable). We extend our definitions to the first steps of the so-called arithmetical hierarchy, first on countable spaces, then on closed subsets of $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$.

Let X, Y be two countable sets, with Y being ordered.

A sequence of functions $(f_i : X \to Y)_{i\in\mathbb{N}}$ is *uniformly computable* if $(i, x) \mapsto f_i(x)$ is computable.

A function $f : X \to Y$ is *Π_2-computable* (resp. *Σ_2-computable*) if $f = \inf_{i\in\mathbb{N}} \sup_{j\in\mathbb{N}} f_{i,j}$ (resp. $f = \sup_{i\in\mathbb{N}} \inf_{j\in\mathbb{N}} f_{i,j}$), where $(f_{i,j})_{(i,j)\in\mathbb{N}^2}$ is a uniformly computable sequence of functions.

A closed set $\mathcal{K} \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ is *Π_2-computable* if the set

$$\left\{(w, r) \in \mathcal{A}^* \times \mathbb{Q} \ : \ \overline{\mathbf{B}(\widehat{\delta_w}, r)} \cap \mathcal{K} \neq \emptyset\right\}$$

is Π_2-computable, that is, its characteristic function is.

Remark 1 The symmetric notions of Π_2- and Σ_2-computability come from an analogy with the real arithmetical hierarchy [16,17]. These definitions extend naturally to Π_n- and Σ_n-computability. Other equivalent definitions exists, see for example [9] for Π_2-computability or [8] for a more general result.

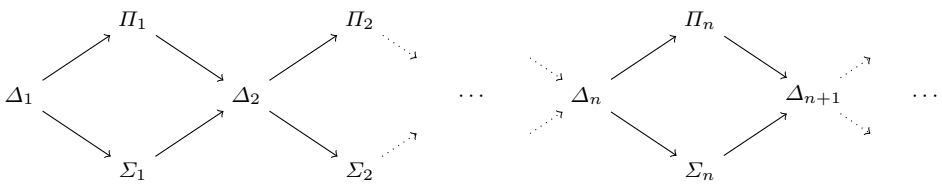

Fig. 1 Representation of the computability hierarchy of closed subsets of $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$. Arrows indicate strict inclusion relations [16].

Proposition 3 (Second computability obstruction)
Let $F : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ be a cellular automaton and μ be a computable measure. Then $\mathcal{V}(F, \mu)$ is a nonempty Π_2-computable compact set.

Aiming at a reciprocal, notice that Π_2-computable compact sets can be all be described as the set of limit points of a sequence of measures $(\widehat{\delta_{w_n}})_{n\in\mathbb{N}}$, where the sequence of patterns $(w_n)_{n\in\mathbb{N}}$ is uniformly computable. However, our construction cannot do better that following such a sequence along a polygonal path, that is, along segments of the form $\left[\widehat{\delta_{w_i}}, \widehat{\delta_{w_{i+1}}}\right] = \left\{t\widehat{\delta_{w_i}} + (1-t)\widehat{\delta_{w_{i+1}}} : t \in [0,1]\right\}$. The following proposition shows that this corresponds to connected limit sets of measures (not necessarily path-connected).

Proposition 4 (Technical characterisation of Π_2-CCC sets)
Let $\mathcal{K} \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ be a non-empty Π_2-computable, compact, connected set (Π_2-CCC for short). Then there exists a uniformly computable sequence of cubic patterns $(w_n)_{n\in\mathbb{N}}$ such that $\mathcal{K}$ is the limit of the polygonal path defined by $(w_n)_{n\in\mathbb{N}}$, that is,

$$\mathcal{K} = \bigcap_{N>0} \overline{\bigcup_{n\geq N} \left[\widehat{\delta_{w_n}}, \widehat{\delta_{w_{n+1}}}\right]}.$$

As we mentioned before, the proof of these statements can be found in [9] or [8] for an extended version.

3 Construction

To obtain the results announced in the introduction, we prove the following result in conjunction to Proposition 4.

Theorem 2 *For any uniformly computable sequence $(w_n)_{n\in\mathbb{N}}$ of cubic patterns of $\mathcal{B}^*$ of dimension at most d, there exists a larger alphabet $\mathcal{A} \supset \mathcal{B}$ and a cellular automaton $F : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ such that for any Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$:*

$$\mathcal{V}(F,\mu) = \bigcap_{N>0} \overline{\bigcup_{n\geq N} \left[\widehat{\delta_{w_n}}, \widehat{\delta_{w_{n+1}}}\right]}.$$

In the rest of the article, assume some fixed alphabet $\mathcal{B}$ and some uniformly computable sequence $(w_n)_{n\in\mathbb{N}}$ of patterns of $\mathcal{B}^*$. We present the construction of the alphabet $\mathcal{A}$ and cellular automaton F.

3.1 Sketch of the construction

We detail the construction of $\mathcal{A}$ and F by describing the tasks to be performed on the initial configuration. Each letter of $\mathcal{A}$ is a product of seven layers separated in three groups, each group representing some information needed to perform a given task. The alphabet of each layer contains a special *blank* symbol # to denote the absence of information.

- The first group is dedicated to the colonising of the configuration. Since we have no control over the contents of the initial configuration, we want to erase (almost) all symbols present initially in favour of various processes that we can control and synchronise. To do this, the *birth layer* contains a *seed* symbol $\boxed{*}$ that can only appear in the initial configuration. Each seed gives birth to a stationary *heart* $\boxed{\heartsuit}$ on the same layer, and to a *membrane* on the *growth layer* which grows in every direction. As it grows, the membrane erases everything in its path, except for other membranes issued from a seed with which it merges.
- The second group is used to divide the colonised space into mostly independent areas called *organisms*, each organism having at its centre a heart issued from a seed. The borders between organisms are redefined regularly by processes on the *organism layer*. Furthermore, organisms need to grow in size regularly, which is achieved by merging organisms whose hearts are close using the *evolution layer*.
- The third group deals with the internal metabolism of the organisms. The goal is first to compute each w_n in succession, which is achieved by simulating a Turing machine in the *computing layer*; then, the main layer of the whole body of the organism is filled with concatenated copies of w_n by using a copying process taking place on the *copying layer*. The above is done synchronously in all organisms, at some time t_n for each w_n.

Copies of w_n are written on the *main layer*, which implies that the corresponding alphabet is $\mathcal{B} \cup \{\#\}$. To sum up, the global alphabet is $\mathcal{A} = \mathcal{A}_{\text{birth}} \times \mathcal{A}_{\text{growth}} \times \mathcal{A}_{\text{orga}} \times \mathcal{A}_{\text{evol}} \times \mathcal{A}_{\text{comp}} \times \mathcal{A}_{\text{copy}} \times (\mathcal{B} \cup \{\#\})$. We check that $\mathcal{B} \subset \mathcal{A}$ up to the bijection $b \mapsto (\#,\#,\#,\#,\#,\#,b)$. Denote $p_{\text{birth}}, p_{\text{growth}}, p_{\text{orga}}, p_{\text{evol}}, p_{\text{comp}}, p_{\text{copy}}, p_{\text{main}}$ the projections on each coordinate.

During the description of F, we will treat each layer successively. The layers were introduced in order of dependency, in the sense that the time evolution of symbols in a given layer only depends on the contents of layers in the same group and the one immediately preceding it. Furthermore, the main layer is only affected by the copying layer.

3.2 Space colonisation: Seeds and membranes

In this section, we describe the cleaning of the configuration through the seeds and the birth, growth and fusion of membranes. We deal only with the birth layer and alphabet $\mathcal{A}_{\text{birth}}$ for

the moment. Section 3.2.1 gives the general ideas of the process, while Section 3.2.2 focuses on technical difficulties of the cellular implementation.

3.2.1 Creation myth: a sketch

Seeds and hearts The alphabet $\mathcal{A}_{\text{birth}}$ contains the seed symbol $\boxed{*}$, which can only appear in the initial configuration and cannot be produced by the local rule of F. At the first step, each seed spawns a number of processes and turns into a heart $\boxed{\heartsuit} \in \mathcal{A}_{\text{birth}}$. This heart and those processes (and those spawned from them) are called *initialised*, which means that their behaviour is well controlled and synchronised (since they are all born at time 1). All other symbols are *uninitialised*.

If two seeds are too close from each other (d_∞ less than 5), the largest (in lexicographic order) is erased to give enough space to the other seed to spawn its processes. A seed that is not destroyed at time 1 in this way is called *viable*. By abuse of notation we write $p_{\text{birth}}(c_x) = \boxed{*}^V$ to mean that the configuration c has a viable seed at coordinate x.

Birth of membranes Each occurrence of $\boxed{*}$ triggers the birth at time 1 of a living *membrane*. The membrane consists in membrane symbols $\square\!\!|$ and $\overline{\square\!\!|}$ (and all their rotations) that form initially the surface of an hypercube of edge length 5 centred on the seed. The membrane is oriented, being able to distinguish inside from outside through orientation vectors.

Definition 1 A membrane m at time t is a maximal 1-connected set of coordinates containing membrane symbols $\square\!\!|$ or $\overline{\square\!\!|}$ at time t with consistent outward orientation; i.e., orientation of neighbouring membrane symbols differ in at most one coordinate, and at most by 1.

When a membrane m forms a closed curve (which is the case for initialised membranes), we denote $\mathbf{Supp}(m)$ its support. In this case, $\mathbf{Supp}(m)$ partitions $\mathbb{Z}^d$ into a finite set $\overline{\mathbf{Int}}(m)$ and an infinite set $\mathbf{Ext}(m)$ which are ∞-connected. We also denote $\mathbf{Int}(m) = \overline{\mathbf{Int}(m)} \setminus \mathbf{Supp}(m)$. By "outward" in the previous definition, we mean that the orientation vectors of m are directed towards $\mathbf{Ext}(m)$.

If a membrane has a malformation that can be detected locally around a symbol (e.g. no neighbours, inconsistent orientations), a death process is spawned. Since uninitialised membranes can be locally well-formed, this is not enough to discriminate them from initialised membranes.

Breathing, growing, getting older Each membrane symbol is associated with an *age counter*, which is a binary counter initialised at 0 and increasing at each step, whose aim is to keep track of the value of t. Notice that in an initialised membrane all age counters are equal. Figure 2 represents some part of a membrane with arrows and counters.

From time 1 onward, the membrane grows slowly towards the outside, erasing the content of other layers as it progresses with the exception of other membranes (see next paragraph). This is governed by the *respiration process*: each time the age stored in its counter is the square of an integer, the membrane grows to the outside, making one step in every direction. Technical details related to the implementation of age counters and the respiration process are the object of the next section.

Fight for survival and death When the growing membranes meet other membrane symbols, they try to determine locally whether they are part of an initialised membrane (in which case the two should merge), or some uninitialised symbols which should be erased. We call *dead* an uninitialised group of membrane symbols that present some locally detectable malformation, such as non-connexity, the absence of or inconsistencies between age counters/respiration processes,

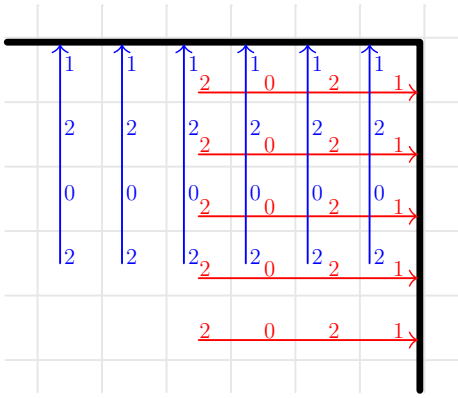

Fig. 2 A corner of a membrane in dimension 2. The arrows give the orientation and the counters store the age of the membrane.

inconsistencies between inner and outer orientation for neighbours, etc. In this case, the malformation generates a death signal ☠ that spreads through the whole membrane erasing it. However, such a group can also form a *zombie membrane*, that is apparently well-formed though uninitialised. Initialised (living) and zombie membranes are distinguished through age counters.

Fact 1 *At time t, all age counters associated with a well-formed membrane have value at least $t-1$, the minimum being reached only for initialised membranes.*

Indeed, age counters of initialised membranes are initialised at 0 at time 1, while age counters of zombie membranes were already present (with a positive value) at time 0, and both are incremented by 1 at each step.

3.2.2 Implementation of age counters

This section is dedicated to the details of the cellular implementation of the age counters in the growth layer. First we show how to implement age counters with binary counters using logarithmic space. $\mathcal{A}_{\text{growth}}$ contains membranes symbols and age counter symbols. Each membrane symbol ◨ or ◩ contains an outward orientation label consisting of a vector of $\mathcal{U}nit(d)$.

Basis and carry propagation We use a redundant binary basis in all counters. Let $c = c_{n-1} \dots c_0 \in \{0, 1, 2, \overline{1}\}^n$ be a counter. The **value** of c is $\sum_{i=0}^{n-1} c_i 2^i$ (reverse order) where $\overline{1}$ has value -1. Since $2 = 10$, 2 can be seen as a 0 with a carry, and $\overline{1}$ as a 0 with a "negative" carry.

At each time step, carries are propagated along the counter, which can be done in a local manner ($02 \to 10, 12 \to 20, \#2 \to 10, 1\overline{1} \to 01, 0\overline{1} \to \overline{1}1$). Additional zeroes at the beginning of the counter are erased ($\#1\overline{1} \to \#1$).

In order to increment or decrement the counter by one, which is the case for age counters, the rule is adapted at the least significant bit of the counter (for incrementing: $0 \to 1, 1 \to 2, 2 \to 1$).

Age counters The age counters are implemented as follows. The least significant bit of each counter is next to its corresponding membrane symbol, and the following bits lie on a line directed towards the inside of the membrane. To each possible direction (corresponding to some $\pm e_j$) corresponds a different sublayer, which allows counters to cross near the corners. Thus the age counters use $2d$ sublayers, each sublayer containing symbols $\{\#, \overline{1}, 0, 1, 2\}$.

Recall that any inconsistency – the absence of an age counter for some membrane symbols, parallel counters containing different symbols, etc. – spawns a death process, which spreads in the whole membrane and erases all layers of these cells.

3.2.3 The respiration process

The goal of the respiration process is to govern a slow growth of the membrane.

Breathing counters Along with the age counter, on two other sublayers of $\mathcal{A}_{\text{growth}}$, two counters A and B are initialised at time 0 with values 1 and 0, respectively. From there on three phases alternate, the current phase being labelled on the membrane symbol:

Phase $+$: A is decremented while B is incremented. When A reaches 0, the phase passes to $-$;
Phase $-$: A is incremented while B is decremented. When B reaches 0, the membrane breathes;
Breath: For one step A is incremented while B is unchanged, then the phase passes to $+$.

The value of $A + B$ is constant during a cycle (1 for the first cycle) except for the last step where it is incremented by one. The cycle takes a total time $2(A + B) + 1$. Therefore a breath occurs at each time t^2 for $t > 1$. In Figure 3 we represent the update operation of all three counters, that is, incrementing the age and updating A and B according to the phase.

t	*phase*	A	B	*age*	$5 + 2\lfloor\sqrt{t}\rfloor$
1	$+$	1	0	0	5
2	$-$	0	1	1	5
3	$-$	1	0	2	5
4	*breath*	2	0	11	7
5	$+$	$1\bar{1}$	1	12	7
6	$-$	0	2	21	7
7	$-$	1	$1\bar{1}$	102	7
8	$-$	2	0	111	7
9	*breath*	11	0	112	9
10	$+$	10	1	121	9

Fig. 3 Values of the three counters for $t \leq 10$.

Lemma 1 *The counter update can be performed locally with radius* 2.

Proof The incrementations and decrementations described above can be achieved with radius 1. The least significant bit can be distinguished by being next to the membrane symbol which contains the information on the current phase and the most significant bit is next to a blank symbol (on its layer).

We show that the fact that a counter is worth 0 is detectable with radius 2 (to see why this is nontrivial, consider the update $\#1\bar{1}\bar{1}\ldots\bar{1} \to \#00\ldots0$). During a decrementation the least significant bit alterns between 0 and $\bar{1}$. Since carries progress at "speed" one, two negative carries can never be next to each other. Therefore the only possible representations of 1 are $\#1$ and $\#1\bar{1}$, and both yield $\#0$ at the next step. Therefore detecting when the counter is worth 0 requires radius two, in order for the membrane symbol to "see" the word $\#0$. □

Breathing process Each breath makes the membrane progress by one cell on every direction, a process we detail below. If such a symbol is not produced synchronously by the whole membrane, then ☠ signals spawn and spread to erase the membrane.

Recall that each membrane symbol at coordinates x is labelled with an outward growth direction, which is a vector $v = \sum \varepsilon_j e_j \in \mathcal{U}nit$. The membrane symbol is the border between $\mathbf{Int}(m)$ and $\mathbf{Ext}(m)$, with the orientation vector indicating which part $\mathbf{Ext}(m)$ is. When a breath symbol appears, the membrane symbol is removed and new symbols are created on all cells of $\mathbf{Ext}(m)$ that were ∞-adjacent to any membrane symbol of m. The new orientation vectors are determined by remaining coherent with the old orientation.

The remaining task is to reproduce the counters for the new symbols.

First consider the case of a face symbol x and orientation e_j. Right after a breath, when a new symbol is created at coordinate $x + e_j$, the symbol at x is replaced by a placeholder symbol s_{lim}. This symbol progressively shifts the counters of x by one cell in direction e_j, marking at each step the limit between the part which is to be shifted and the part already shifted. The counters keep updating by ignoring this symbol, which increases the radius to 3. Figure 4 illustrates the shift.

From this section we deduce the following:

Fact 2 *Each initialised membrane forms an hypercube of edge length* $5 + 2\lfloor\sqrt{t}\rfloor$ *at time* t.

Since the counters of an initialised membrane are initialised to 0 at time 1, the maximum size of the counters is $\lceil \log_2(t-1) \rceil$ at time t and membranes have enough space to contain them. For technical reasons that will become clear later, we need to quantify the maximal number of breaths of any membrane symbol (not necessarily initialised) in a given time:

Lemma 2 *The number of breaths triggered by any membrane symbol between times* t *and* $t + k$ *is at most* $\lfloor\sqrt{t+k}\rfloor - \lfloor\sqrt{t}\rfloor$.

Proof Apart from time 0 (when a breath symbol could be present), a breath is only triggered when the B counter of a membrane symbol without local malformations reaches 0. This symbol must be issued from a seed or from a membrane symbol already present at time 0. In the first case, since a breath is triggered at each step when the time t is the square of an integer (except for 1), the number of breaths before time t is $\lfloor\sqrt{t}\rfloor - 1$. The lemma follows.

In the second case, the membrane symbol had at time 0 counters A and B with some positive values a_0 and b_0 and some phase ε_0, values which correspond to those of an initialised set of counters at some time $t_0 > 0$. From there on the time evolution of the membrane symbol is similar to the evolution of an initialised membrane symbol of age $t_0 + t$, which means that the number of breaths between times t and $t+k$ is $\lfloor\sqrt{t+t_0+k}\rfloor - \lfloor\sqrt{t+t_0}\rfloor \leq \lfloor\sqrt{t+k}\rfloor - \lfloor\sqrt{t}\rfloor$. □

Forming colonies For an initial configuration c, define $M_t(c)$ to be the set of initialised membranes at time t. Then the *colonised space* at time t is:

$$\mathbf{Col}_t(s) = \bigcup_{m \in M_t(c)} \overline{\mathbf{Int}}(m).$$

When a membrane grows, it erases the content of every other layer of the cells it encounters, except when the birth layer contains the outer border of a membrane. In this case, the comparison process starts, which is the topic of the next section.

To sum up, the alphabet $\mathcal{A}_{\text{birth}}$ contains seeds and hearts, and $\mathcal{A}_{\text{growth}}$ contains the states used for membranes (including counter sublayers). As we will see in the next section, 2^d different membranes can share the same cell, so this alphabet will be duplicated this many times.

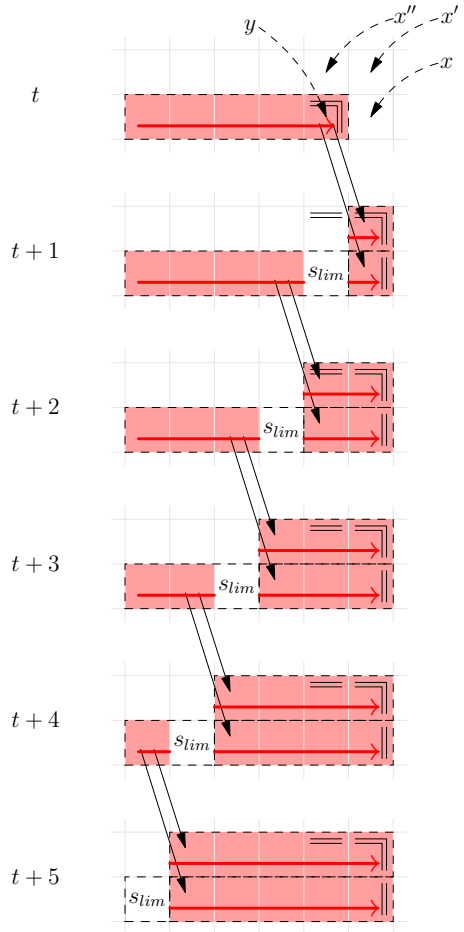

Fig. 4 After a breath, a membrane corner symbol at cell y is erased and new membrane symbols appear at x, x' and x''. For the sake of readability, the vertical counters are not represented, hence we draw only the new horizontal counters in x and x'. At each step between $t+1$ and $t+4$, the red cells represent superposition of the age, A and B counters. They are copied to the new membrane symbols, but the incrementation does not stop.

3.2.4 Survival of the youngest

As membranes grow and tend to cover the whole space, different membranes eventually meet. The result of the encounter should depend on the nature of the membranes: two initialised membranes should merge while an initialised membrane should erase an uninitialised membrane (what happens between uninitialised membrane is irrelevant). In this section, we devise a comparison process to distinguish initialised from uninitialised membranes, using the growth layer and its alphabet $\mathcal{A}_{\text{growth}}$.

From Fact 1 we know that initialised membranes have the youngest age counters, and only tie with other initialised membranes. The value of the age counters are compared to let the younger membrane survive, with merging occurring in case of equality.

Many membranes meeting When we say that two membranes m and m' meet at time t in cells x and x', we mean that there exists $1 \leq j \leq d$ such that:

- either $x \in \mathbf{Supp}(m) \cap \mathbf{Supp}(m')$, $x + e_j \in \mathbf{Ext}(m)$ and $x - e_j \in \mathbf{Ext}(m')$, in which case take $x' = x$;
- or $x \in \mathbf{Supp}(m) \cap \mathbf{Ext}(m')$ and $x + e_j \in \mathbf{Supp}(m') \cap \mathbf{Ext}(m)$, in which case take $x' = x + e_j$.

The two possible situations are illustrated in Figure 5.

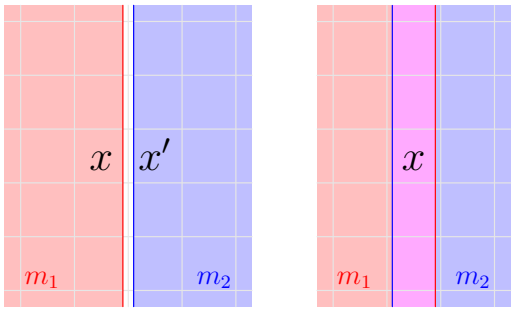

Fig. 5 Depending on the parity of the distance between membranes, they will meet either when they share some border cells ($x' = x$), or when the borders are adjacent($x' = x + e_j$).

In particular, the membranes arriving from opposite directions, they have (at least) an age counter in opposite directions, say e_j and $-e_j$. These age counters are to be copied on a dedicated sublayer of $\mathcal{A}_{\text{growth}}$ and compared.

Copying phase At positions x and x', two symbols $\boxed{C^1_j}$ and $\boxed{C^1_{-j}}$ are written on the growth layer to trigger the process (if $x = x'$, a symbol $\boxed{C^1_{\pm j}}$ represents the superposition of those symbols) and progress at speed one in the corresponding direction, copying at each step one bit from the age counter to a sublayer of $\mathcal{A}_{\text{growth}}$. Carries 2 are copied as 0: indeed, the copy is performed at the same speed as the carry progresses, so it would be copied at each step otherwise (the carry is taken into account once it turns a 0 into a 1). More generally, only carries that appeared before the beginning of the copy can influence the copied counter, which is not incremented. Thus the copied counter have the same value as the age counter at the beginning of the copy. When it reaches the end of its counter, each copy symbol turns into a comparison symbol $\boxed{C^2_j}$ (resp. $\boxed{C^2_{-j}}$), which triggers the comparison phase.

Comparison phase The comparison symbols return towards the meeting point, "pushing" the copied bits in front of them in a caterpillar-like movement, starting from the most significant bit. The returning bits use another sublayer of $\mathcal{A}_{\text{growth}}$. The process is represented in Figure 6.

As the returning bits reach the meeting point, one of the following situations occur:

- the most significant bit from one side arrives earlier than the most significant bit from the other side. In this case the age counter of the corresponding side is shorter, which means that the membrane of this side is younger;
- both most significant bits arrive simultaneously at the meeting points x and x'. Then bits are compared as they arrive. The first bit that is smaller than its counterpart corresponds to the side of the younger membrane;
- in the previous case, if all bits are equal until the end, both membranes have exactly the same age.

Those three possibilities are tested locally at the meeting point and the result is written as a symbol (on its own sublayer) marking the direction of the younger membrane, with $=$ in case of

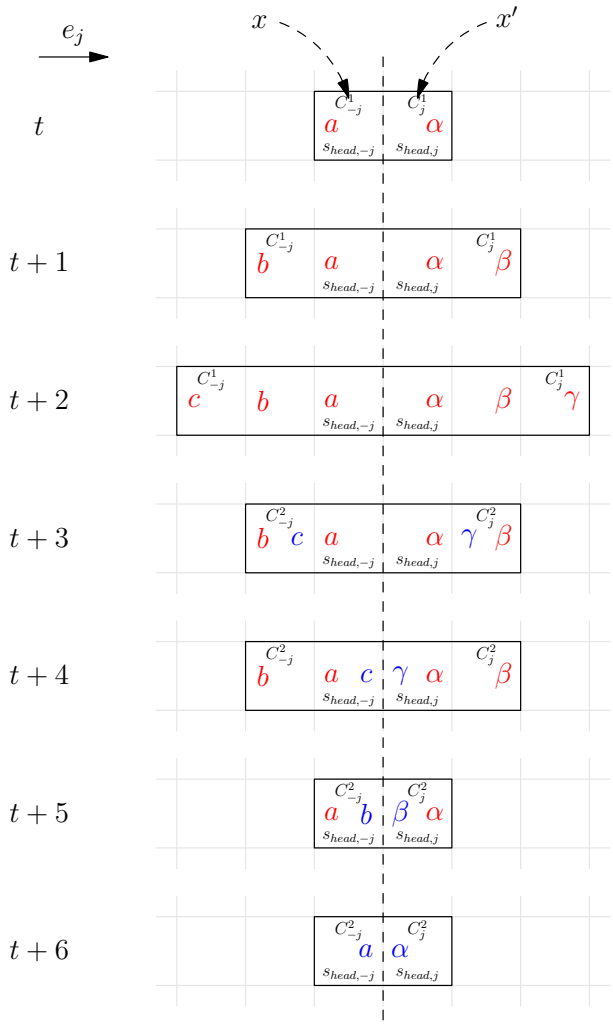

Fig. 6 In this example, two membranes meet in cells x and x' at time t. Their age counters are respectively abc and $\alpha\beta\gamma$ of same length 3. Only the growth layer is represented. At the end $(t+6)$, the decision can be made in both x and x'. In this particular case, the symbols s_{head} have not been moved, which means neither of the membranes did extend during the comparison.

a tie. If for some reason a symbol ☠ (death process) reaches the symbol of one of the sides, the comparison stops and the surviving membrane is marked as younger "by default".

If a membrane is declared younger, all auxiliary symbols used for comparison are erased and a death process triggers in the older membrane. The younger membrane will resume its growth naturally at the next breath. If the result is a tie, both membrane symbols are erased along with all associated auxiliary states: the membranes are merged.

Remark 2

– In general two membranes may have more than one meeting point. In that case, comparisons are performed simultaneously at every point and in every concerned direction. In the case of a tie, all symbols participating in the meeting are erased simultaneously; any local discrepancy results in the spawn of a death process.

- At most 2^d different membranes can meet in the same cell x (corners of hypercubes arriving from all possible directions). To solve this problem, we duplicate each sublayer (in $\mathcal{A}_{\text{birth}}$ and $\mathcal{A}_{\text{growth}}$) used in the comparison into 2^d copies, each copy being able to perform a comparison independently of the others. If the membrane is older than at least another membrane, a death process is spawned; it merges if it is tied for youngest.
- Let ℓ be the length of the shortest age counter. The previous process needs ℓ steps to copy this age counter on the growth layer, and 2ℓ steps to send them one by one to the meeting point. Regardless of the length of the other counter, the comparison reaches a result after the last bit of the shortest counter arrives. Therefore the whole process takes at most 3ℓ steps. Remember that $\ell = \lceil \log t - 1 \rceil$ if one of the membranes is initialised (assuming the comparison process began at time t).

Breathing during comparisons We did not take into account the possibility that one of the membranes breathes during the comparison. For each meeting of a pair of membranes, call *instigating membrane* the one whose breath has triggered the meeting (possibly both if they moved simultaneously; this is the case for initialised membranes).

Lemma 3 *Let m be a living membrane meeting another membrane at time t. During the comparison process, m may breath at most one time if it is not instigating, and cannot breath at all if it is instigating.*

Proof Using the above remark, we know that the comparison process takes at most $k = 3\lceil \log(t-1) \rceil$ steps. Using Lemma 2, the number of breaths of m between times $t-1$ and $t+k$ is at most:

$$\begin{aligned}
\left\lfloor \sqrt{t+k} \right\rfloor - \left\lfloor \sqrt{t-1} \right\rfloor &\leq \left\lceil \sqrt{t+k} - \sqrt{t-1} \right\rceil \\
&\leq \left\lceil \frac{k+1}{2\sqrt{t-1}} \right\rceil \quad \text{(since the derivative of } \sqrt{t-1} \text{ is decreasing)} \\
&\leq \left\lceil \frac{3\log(t-1)+1}{2\sqrt{t-1}} \right\rceil \leq 1.
\end{aligned}$$

If m is instigating, then by definition m breathed at time $t-1$ and cannot breath again before time $t+k$. Otherwise, m breathes at most one time. □

Therefore, if one or both membranes move during the comparison because of the respiration process, it writes a head symbol $s_{head,j+}$ to recall its new position. If a membrane extends more than twice before the end of the comparison, a death process is triggered for this membrane. As the radius of F is more than 2, the moved membrane can still read the result of the comparison.

Lemma 4 *Take any $t > 0$ and initial configuration $c \in \mathcal{A}^{\mathbb{Z}^d}$. Then:*

$$\mathbf{Col}_t(c) = \{x \in \mathbb{Z}^d : \exists y \in \mathbb{Z}^d, d_\infty(x,y) \leq 1 + \sqrt{t}, p_{\text{birth}}(c_y) = \boxed{*}^V\}.$$

In other words, the colonised space at time t is exactly the set of cells that, at time 0, are at distance less than $1+\sqrt{t}$ from a viable seed.

Proof We prove this result by structural induction. If $t = 1$, then the colonised space is the set of all initialised membranes which are hypercubes of edge side 5 around each viable seed, and the result is proved.

Now suppose that the hypothesis holds at time t. Notice than $\mathbf{Col}_t(c) \subset \mathbf{Col}_{t+1}(c)$, and that merging does not add any cell to the colonised space: an initialised membrane cannot be erased,

and the colony obtained after merging two membranes is the union of the colonies defined by the two merged membranes. Only the breathing process may add new cells to the colonised space.

Consequently, the induction step is empty if $(t+1)$ is not a square since $\mathbf{Col}_t(c) = \mathbf{Col}_{t+1}(c)$ and $d_\infty(x,y) \leq 1+\sqrt{t} \Leftrightarrow d_\infty(x,y) \leq 1+\sqrt{t+1}$ (distances are integers). If $(t+1)$ is a square, then all membrane symbols in initialised membranes take a breath and extend by one cell towards the outside. Now, take a cell y at distance $1+\sqrt{t+1}$ from the nearest viable seed. By the induction hypothesis, $y \notin \mathbf{Col}_t(c)$, but y has a neighbour $y+v$ with $v \in \mathcal{U}nit$ at distance $1+\sqrt{t}$ from that seed, so that $y+v \in \mathbf{Col}_t(c)$. Therefore $y+v$ must be a membrane symbol in the support of an initialised membrane that breathes at time $t+1$, and therefore $y \in \mathbf{Col}_{t+1}(c)$. Conversely, if a cell z is at distance greater than $1+\sqrt{t+1}$ from the nearest viable seed, it cannot have a membrane symbol belonging to an initialised membrane as a neighbour, so that $z \notin \mathbf{Col}_{t+1}(c)$. □

3.3 Colonies: evolution of the population

From Lemma 4 we can see that only the contents of the colonised space matter asymptotically. In this section, we describe the interaction of organisms inside colonies. In all the following, we assume we are inside a colony, and the support of the surrounding membrane acts as an impassable wall for any symbol in the second group layers: the organism and evolution layers.

3.3.1 Hearts and organisms

As we saw, seeds $\boxed{*}$ at time 1 spawn a membrane and turn into hearts $\boxed{\heartsuit}$. Each heart will be the centre of an *organism* which is itself a subset of the colony. At first each colony have only one heart, but as initialised membranes merge together, the colonies become multi-hearted, and the colony space is partitioned into organisms (except possibly a negligible part). For various reasons, the size of the organisms should grow in a controlled way, which requires some hearts to be progressively removed.

In the present section, we present the cycle of division of colony space and life of the organisms, and all symbols presented here belongs to $\mathcal{A}_{\text{orga}}$.

The life of an organism consists in a succession of generations. We introduce a sequence of times $(t_n)_{n\geq 1}$ (to be fixed later), marking the limit between the $n-1$-th and n-th generation. Time is tracked by the heart through a binary *time counter* similar to age counters, initialised at 1 at time 1 (along with the heart) and remaining stationary next to the heart. Details on the implementation and the way to determine when $t = t_n$ will be given in Section 3.4.1.

Organisms expanding At each time t_n, *organism-building signals* spread from every heart, progressing as membrane symbols but with speed 1 (although they do not carry any counters). While progressing, they erase the old contents of the second and third group layers except for the main layer. When they meet a membrane or another organism-building signal, they vanish leaving behind a *neutral border symbol* $\boxed{\$}$. For parity reasons, if two signals emitted by hearts in x and x' arrive simultaneously in two neighbour cells y and y', they leave behind two *pseudo-border symbol* $\boxed{\$'}$ with an orientation vector towards the interior of their organism, that is, the direction opposite to the initial organism-building signal. Just as membrane symbols, 3^d-1 different organism-building symbols and pseudo border symbols are required (one for each orientation).

The *territory* of a heart is the maximal set of 1-connected cells containing the heart and no neutral border symbol $\boxed{\$}$ nor pseudo border symbol pointing towards another organism; in other words, the set of cells reached first by organism-building signals emitted by this heart. At time

$t_n + k$ (assuming $t_n + k < t_{n+1}$), the only cells of the colony that are not part of some organism are either at distance more than k from the nearest heart, or were outside the membrane at time t_n (and a breath had included them since).

Fact 3 *Let x be a cell containing a heart at time t with $t_n \leq t \leq t_{n+1}$, and let y be a cell in its territory. Then the organism-building signal emitted by x reached y at time $t_n + d_\infty(x, y)$ and no other organism-building signal reached a neighbour of y before that time.*

The following lemma gives insight about the shape of the territories, namely, that they are a (discrete) star domain, whether the borders are included or not.

Lemma 5 *If a cell y belongs to the territory of a heart in cell x, then each cell y' such that $d_\infty(x, y') + d_\infty(y', y) = d_\infty(x, y)$ is also in this territory. Furthermore, y' can be a border only if y is a border.*

Proof For the first part of the lemma, suppose such a y' is not in the territory of x. We can build an ∞-path between y' and y consisting of cells $(y^{(i)})_{0 \leq i \leq N}$ such that $d_\infty(x, y^{(i)}) + d_\infty(y^{(i)}, y) = d_\infty(x, y)$. Take $y^{(j)}$ the first $y^{(i)}$ that belongs to the territory of x.

Denote $T = t_n + d_\infty(x, y^{(j-1)})$ the time when the organism-building signal emitted by x should have reached $y^{(j-1)}$ in the absence of any other heart. Since $y^{(j-1)}$ is not in the territory of x, there must exist another heart x' that emitted an organism-building signal that arrived in $y^{(j-1)}$ before time T (recall that the pseudo-borders are considered as parts of organisms).

Since $y^{(j)}$ is adjacent to $y^{(j-1)}$, $y^{(j)}$ is reached by some signal before time $T + 1$. But the signal from x cannot reach $y^{(j)}$ before time $t_n + d_\infty(x, y^{(j)}) = T + 1$. Therefore $y^{(j)}$ is not in the territory of x, a contradiction.

For the second case, notice that y' is a border if and only if both signals reached this cell simultaneously. Then the same reasoning along the path $(y^{(i)})_{0 \leq i \leq N}$ shows that these cells cannot be inside the territory of x. □

3.3.2 Natural selection

In this section we consider the evolution layer and the alphabet $\mathcal{A}_{\text{evol}}$.

To have enough computation space and ensure that the auxiliary symbols are in negligible density, the minimal size of the organisms should grow regularly. More precisely, we require that the territory of any organism during the n-th generation contains at least a hypercube of side length $2n + 1$ centred at its heart. If two hearts are at distance less than $2n + 1$, they are said to be in *conflict*. In this section, we devise a selection process to detect this fact and to erase one of them.

The first lemma is related to the quantity and position of hearts conflicting with a given heart.

Definition 2 For each point x and each vector $\mathcal{R} \in \{>, <, =\}^d$ that is not $=^d$ we introduce the corresponding *quadrant* defined as $\{y \in \mathbb{Z}^d \ : \ \forall i, \ y_i \mathcal{R}_i x_i\}$.

For $n \in \mathbb{N}$, each quadrant contains a unique *n-extremal point*, which is a point y such that $y_i - x_i \in \{-n, n, 0\}$ for all i. Denote $Ext_n(y) = \{\varepsilon e_i : \varepsilon \in \{1, -1\}, y_i - x_i = n\varepsilon\}$ the set of its directions of extremality relative to x. Notice that this notion is defined for any vector y such that $d_\infty(x, y) = n$.

Lemma 6 *At any given generation, a heart conflicts with at most one other heart in each quadrant. In particular, it conflicts with at most $3^d - 1$ hearts.*

Proof Let x be the position of the heart and $t_n \leq t < t_{n+1}$ the current time. The heart in x conflicts with another heart at y if and only if $d_\infty(x, y) \in \{2n, 2n+1\}$. If two conflicting hearts were in the same quadrant, then they would be at distance at most $2n-1$, a contradiction. The bound is tight and reached by the configuration where there is a heart per quadrant at each $2n$-extremal point relative to the heart in x. □

Settling conflicts at random Each heart keeps in memory a random *central bit* and the winner of each conflict is decided through the value of these bits: if the bits are equal, destroy the largest heart (in lexicographic order), and the smallest otherwise. As long as the central bits of all hearts are independent of each other, this process will ensure that the probability of each heart to be destroyed in any conflict has a constant lower bound that only depends on the initial Bernoulli.

Remark 3 In dimension 1 [9,2], the heart in the smallest cell was systematically killed. The reason why we adopt a more sophisticated kill choice method, similar to the one used in [6], is that we later need to control the growth rate of the organisms. This growth rate analysis under the simpler method proved to be more difficult than in dimension 1 and could not be carried out.

Independent bit harvesting In addition to its central bit, each heart maintains a collection of $3^d - 1$ *side* bits, one for each quadrant, that must remain independent of each other and of every other central or side bits of other hearts. This is achieved by using the independence of the initial measure. At time 0, each seed looks at its $3^d - 1$ adjacent cells (which are in one-to-one correspondence with its quadrants): if a cell contains the symbol $\boxed{\mathbb{I}}$ (the symbol choice is arbitrary) then the corresponding bit in the newborn heart is put to 1, and 0 otherwise. Remember that if two seeds are too close, one is erased, hence those bits are really independent from one another and from other hearts.

Each time a heart kills another in a conflict, the dying heart transmits one of his side bits to its killer (the bit corresponding to which quadrant the killer belongs to). The new central bit of the killer is the sum of the side bits received from its victims. As we will see, this process lets us maintain probability bounds on the value of central bits while preserving independence.

Body building with bits In order to detect conflicting hearts, each heart in position x builds at time t_n a *body*: a hypercube of side length 5 centred around the heart. Each body symbol in position y carries various pieces of information: the central bit of the heart, which quadrant (relative to the heart) it belongs to, the corresponding side bit, and the set of directions in which it is 2-extremal ($Ext_2(y)$). Next to the heart, a dedicated process writes two copies of n as binary counters from the computation layer (see next section).

The three phases of conflict resolution are:

Body building Each heart sends n (hypercube-shaped) signals that progress at speed 1 in every direction, one every n steps. The count is kept by decrementing the binary counters. As each signal reaches the body, it pushes it outwards by one cell. The set of directions in which it is k-extremal (where k is the number of steps) is kept updated.

Conflicting If two bodies intersect, the corresponding hearts are in conflict. Body symbols at n-extremal points determine locally the relative positions (quadrants) of the two hearts (see next paragraph), and using the values of the central bits, the winner of the conflict. This phase takes only one step.

Body shrinking Each heart keeps sending (hypercube-shaped) signals every n steps, but these signals now pull the body inwards as they reach it. The body is destroyed as it reaches size 5, transmitting its information to the heart. The heart stops sending signals.

If a heart has fought no conflict during a generation, nothing happens. If it loses at least one conflict, it self-destroys. Not being able to send border-building signals, its territory will be occupied by other organisms at the next phase. If it wins all its conflicts, it replaces its central bit by the sum (modulo 2) of all the bits it receives from its victims.

Quadrant determination During the conflicting phase, only the extremal points of its body settle the conflict. Take an extremal point of some body in position y such that there is a body symbol in position y' (not necessarily extremal), belonging to another body. Denote $\mathcal{R}$ and $\mathcal{R}'$ the respective quadrants of y and y'. The extremal point y deals with the conflict if and only if these properties hold:

- (I) $y' = y + \sum_{v \in D} v,\ D \subseteq Ext_n(y)$ (note that D may be empty, that is, $y' = y$) ;
- (II) $\forall \varepsilon e_i \in Ext_n(y) \cap Ext_n(y'),\ y_i \neq y'_i$;
- (III) for each i, if $\mathcal{R}_i$ is $=$ then $\mathcal{R}'_i$ is $=$ as well.

The first condition means that extremal points are only interested in directions (or sets of directions) they are extremal in. Both other conditions are illustrated in Fig 7.

Lemma 7 *A conflict is settled exactly once by each involved heart, at the extremal point corresponding to the relative position (quadrant) of the other involved heart.*

Proof Consider some heart in position x, and a conflicting heart x' located in quadrant $\mathcal{R}$. Since $d(x, x') \in \{2n, 2n+1\}$, we can check that $d(y, x') \leq n+1$ where y is the n-extremal point for x in quadrant $\mathcal{R}$. We check that y satisfies (I) with some neighbour y' and they satisfy both (II) and (III).

Since $d(x, x') = 2n+\alpha$ for some $\alpha \in \{0, 1\}$, there is some $1 \leq i \leq d$ such that $|x_i - x'_i| = 2n+\alpha$. For any such i, fix $y'_i - x'_i = -(y_i - x_i) \in \{-n, n\}$ so that $|x_i - y'_i| = n + \alpha$.

For all others $1 \leq i \leq d$:

- if $\mathcal{R}_i$ is $>$ take $y'_i = x_i + n = x'_i + (n - x'_i + x_i) \in [x'_i - n, x'_i + n]$
- if $\mathcal{R}_i$ is $<$ take $y'_i = x_i - n = x'_i - (n + x'_i - x_i) \in [x'_i - n, x'_i + n]$
- if $\mathcal{R}_i$ is $=$ take $y'_i = x'_i$.

Hence $d(x', y') = n$ and y' belongs to the body of x'.

(I) For every $1 \leq i \leq d$ such that $|x_i - x'_i| = 2n+\alpha$, $|y'_i - y_i| = |(y'_i - x'_i) + (x'_i - x_i) + (x_i - y_i)| = \alpha$. For all others $1 \leq i \leq d$:

- if $\mathcal{R}_i$ is $>$, then $y'_i - y_i = (x_i + n) - (x_i + n) = 0$.
- if $\mathcal{R}_i$ is $<$, then $y'_i - y_i = (x_i - n) - (x_i - n) = 0$.
- if $\mathcal{R}_i$ is $=$, then $y'_i - y_i = (y'_i - x'_i) + (x'_i - x_i) + (x_i - y_i) = 0 + 0 + 0$.

This proves that $d(y, y') \leq 1$ and (I) is verified.

(II) Assume y and y' share an extremality direction e_i and $y_i = y'_i$. That is, $y_i - x_i = n$ and $y'_i - x'_i = n$, so that $x_i = x'_i$. Furthermore, notice that $\mathcal{R}_i$ is $>$ since y belongs to quadrant $\mathcal{R}$. But since $x_i = x'_i$, x' wouldn't be located in quadrant $\mathcal{R}$, a contradiction. The same argument works for an extremality direction $-e_i$.

(III) If $\mathcal{R}_i$ is $=$, since x' is located in quadrant $\mathcal{R}$ relative to x, this means that $x_i = x'_i$. Since y is extremal for x in quadrant $\mathcal{R}$, we have $y_i = x_i$. Therefore y' is located relative to x' in a quadrant $\mathcal{R}'$ such that $y'_i = x'_i$, i.e. $\mathcal{R}'_i$ is $=$.

We show that no other extremal point handles the conflict. Let z be another extremal point in quadrant $\mathcal{R}$" for x, next to a body symbol z' for x'. We distinguish various cases:

- For some i, assume $\mathcal{R}_i$ is $=$ but not $\mathcal{R}"_i$. Then $x_i = x'_i$ but $z_i = x_i \pm n$, i.e. z has an extremality direction along e_i. Similarly z' has an extremality direction along e_i and $z_i = z'_i$, contradicting condition (II).

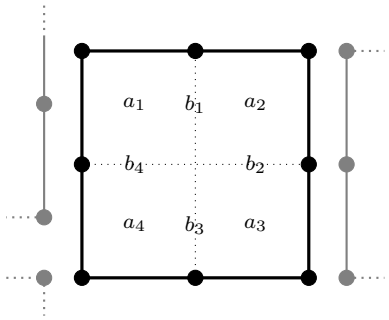

Fig. 7 Body of a heart, with quadrants and extremal points. The heart is involved in conflicts in quadrants a_1, b_2 and a_4. The extremal point for b_4 is uninvolved because of rule (III), and extremal points for a_2 and a_3 are uninvolved because of rule (II). Extremal points for b_2 and a_4 satisfy (II) and (III).

- For some i, assume $\mathcal{R}_i$ is $<$ but $\mathcal{R}"_i$ is $>$. Then $x'_i < x_i$ and $z_i = x_i + n$, so that $z_i > x'_i + n \geq z'_i$ and z does not handle the conflict according to (I).
- For some i, assume $\mathcal{R}_i$ is $<$ but $\mathcal{R}"_i$ is $=$. Then $z_i = x_i > x'_i$ and as $\varepsilon e_i \notin Ext_n(z)$, (I) ensures that $z_i = z'_i$. Hence $z'_i > x'_i$ and $\mathcal{R}"_i$ is not $=$, which contradicts condition (III).

□

Lemma 8 *At any time t, the values of every central and side bit are independent of each other. Furthermore there exists a constant $0 < \alpha \leq 1/2$ such that for any of these bits, the probability that it is worth 1 at time t is bounded between α and $1 - \alpha$.*

Proof At time 0, the central and side bits of each heart are determined by the initial Bernoulli measure, so the lemma is verified and the value of α is determined by the initial measure. A side bit is never changed and only interacts with the outside when its heart is destroyed, so we only prove the result for central bits. Since the only opportunity for central bits to be influenced by or to influence other bits is during conflict resolution, we prove the result by induction on the generation n (with the basic case of the 0-generation already proved).

Assume the result holds for some generation n and consider a heart alive at time t_{n+1}. If the heart is not involved in a conflict at generation $n+1$, then its central bit is unchanged and doesn't influence any outside process, so the property is maintained. If it is involved in k conflicts, then it either dies (in which case there is nothing to prove) or wins them all.

Denote $\beta_1, \ldots, \beta_k$ the side bits it received from his victims. The value of these side bits did not influence the conflict resolution and remains independent of every other bit since they are not sent to any other conflict winner (there is at most one conflict per quadrant). Furthermore:

$$\begin{aligned}
\mu\left(\sum_{i\geq 1}\beta_i = 1 \mod 2\right) &= \mu(\beta_1 = 1)\cdot\mu\left(\sum_{i\geq 2}\beta_i = 0 \mod 2\right) + \mu(\beta_1 = 0)\cdot\mu\left(\sum_{i\geq 2}\beta_i = 1 \mod 2\right) \\
&\geq \alpha\cdot\mu\left(\sum_{i\geq 2}\beta_i = 0 \mod 2\right) + (1-\alpha)\cdot\mu\left(\sum_{i\geq 2}\beta_i = 1 \mod 2\right) \\
&\geq \min(\alpha, 1-\alpha) \geq \alpha
\end{aligned}$$

and symmetrically, using the induction hypothesis of independence between the values of all β_i. □

Thanks to this process we can bound the *radius* of an organism, which is the largest distance from a cell of its territory to its heart.

Definition 3 An organism is *healthy* if its radius is less than $K_n = 2^{n^{d-\frac{1}{2}}}$.

Lemma 9 *Any cell belongs to a healthy organism with asymptotic probability* 1*, i.e.:*

$$\max_{t_n \leq t \leq t_{n+1}} \mu\left(\left\{c \in \mathcal{A}^{\mathbb{Z}^d} : \exists x \in \mathbb{Z}^d, d_\infty(x,0) \leq K_n,\ p_{\text{orga}}(F^t(c)_x) = \boxed{\heartsuit}\right\}\right) \xrightarrow[n\to\infty]{} 1.$$

Proof Consider a heart in $x \in \mathbb{Z}^d$ at time t_n. Its survival until time t_{n+1} only depends on the hearts that are present at distance $2n$ or $2n+1$. Inductively, the survival of an initialised heart until the n-th generation depends only on hearts located at distance at most $\frac{(2n+1)(2n+2)}{2}$: the survival of distant enough hearts is independent.

Denote:

$$\nu_n = \max_{t_n \leq t \leq t_{n+1}} \mu\left(\left\{c \in \mathcal{A}^{\mathbb{Z}^d} : p_{\text{orga}}(F^t(c)_0) = \boxed{\heartsuit}\right\}\right).$$

Take the hypercube of side length $2K_n+1$ centred on 0 and cut it into $\lambda_n = \left(\left\lfloor \frac{2K_n+1}{(2n+1)(2n+2)+1} \right\rfloor\right)^d$ hypercubes of size $(2n+1)(2n+2)+1$. Consider the centres of these hypercubes $x_1, \ldots, x_{\lambda_n}$. Then we have:

$$1 - \max_{t_n \leq t \leq t_{n+1}} \mu\left(\left\{c \in \mathcal{A}^{\mathbb{Z}^d} : \exists x \in \mathbb{Z}^d, d_\infty(x,0) \leq K_n, p_{\text{orga}}(F^t(c)_x) = \boxed{\heartsuit}\right\}\right) \leq (1-\nu_n)^{\lambda_n}$$

where the second step uses the last remark and the shift-invariance of μ.

For any $k \in \mathbb{N}$, a living heart at time t_{k-1} survives until generation k if it wins every conflict it is engaged in. Using Lemma 8, we have $\nu_k \geq \alpha^{3^d-1}\nu_{k-1}$ for some fixed $0 < \alpha \leq 1/2$. Then by recurrence $\nu_n \geq \alpha^{n(3^d-1)}\mu(\{\boxed{*}\})$.

We conclude that $(1-\nu_n)^{\lambda_n} \leq \left(1 - \alpha^{n(3^d-1)}\mu(\{\boxed{*}\})\right)^{\lambda_n} \to 0$ since $\lambda_n \alpha^{n(3^d-1)}\mu(\{\boxed{*}\}) \to \infty$. □

3.4 Individual organisms: internal metabolism

The last group of layers is used to govern the internal metabolism of the organisms. In this section, we consider some organism and describe how it behaves during a generation.

3.4.1 Computing

In this section, we describe the computational layer using the alphabet $\mathcal{A}_{\text{comp}}$. Let $(w_n)_n$ be the uniformly computable sequence of patterns given as an hypothesis of the theorem. Our goal is to delimit a small computation space around the heart where each w_n will be computed in succession.

We use standard techniques to embed the time evolution of any Turing machine $TM = (Q, \Gamma, \#, q_0, \delta, Q_F)$ inside our cellular automaton. The alphabet used for the simulation is $(\Gamma \cup \#) \times (Q \cup \#)$: the left part contains the tape symbol, and the right part contains the current state for the cell where the head is located, and # everywhere else. Then each step of the Turing machine moves the head and modifies the tape around the head according to local information, which can be done through the local rule of a CA.

A brain around each heart The alphabet $\mathcal{A}_{\text{comp}}$ is divided into 4 sublayers containing various computational processes taking place in parallel next to the heart, including Turing machines with a d-dimensional tape. Each layer can read the contents of another sublayer when indicated. Assume that we are at time $t = t_n$.

1. The first layer contains a binary counter keeping track of the current generation;
2. The second layer contains a binary counter keeping track of the current value of t, similarly to binary age counters;
3. The machine on the third layer computes the value of t_{n+1}, then keeps watch on the time counter on the second layer. When it reaches $t = t_{n+1}$, the generation counter is incremented by one, which triggers many other processes.
4. The machine of the fourth layer reads the generation counter and computes the hypercubic patterns w_n along with its side length k. w_n is output on the main layer.

From now on, we fix $t_n = \sum_{k<n} 2^{k^{d-\frac{1}{4}}}$.

Complexity analysis We want to ensure that these computations can be performed between times t_n and t_{n+1} without leaving a hypercube centred on the heart of side length $n^{\frac{d-\frac{1}{4}}{d}}$, that is, that they can be performed in time $t_{n+1} - t_n = 2^{n^{d-\frac{1}{4}}}$ and space $n^{d-\frac{1}{4}}$.

To get rid of the multiplicative constant contained in the O notation, we use the standard techniques of linear speedup and tape compression for Turing machines. For any fixed constant C, by grouping cubes C^d tapes cells together in a single letter and performing C computation steps at once, we can divide required time and space by C. As downside, the tape alphabet of the Turing machines increases exponentially (in C).

First and second layers: This is obvious for the generation counter. The time counter occupies a space $\lceil \log t \rceil \leq \lceil \log t_{n+1} \rceil \sim \log(2^{n^{d-\frac{1}{4}}}) = O(n^{d-\frac{1}{4}})$, and the multiplicative constant is removed by using a base-b counter with b large enough.

Third layer: Computing the value of $t_{n+1} = \sum_{k<n+1} 2^{k^{d-\frac{1}{4}}}$ takes space and time $O(2^{n^{d-\frac{1}{4}}})$.

Fourth layer: Without loss of generality, w_n satisfies the time and space constraints, as the following Lemma shows.

We also assume that $w_n \in \mathcal{A}^{[0,k]^d}$ for some $\frac{1}{2} n^{\frac{d-\frac{1}{2}}{d}} < k \leq n^{\frac{d-\frac{1}{2}}{d}}$, by replacing w_n by concatenating copies of itself if necessary.

Lemma 10 *Given a computable sequence $(w_n)_n$ of hypercubes, there exists another computable sequence $(w'_n = w_{g(n)})_n$ such that:*

- *$g : \mathbb{N} \to \mathbb{N}$ is surjective and non-decreasing;*
- *w'_n is computable in time $O(2^{n^{d-\frac{1}{4}}})$ and space $O(n^{d-\frac{1}{4}})$.*

Proof Consider a TM ϕ_0 that computes the sequence $(w_n)_n$. We describe another machine ϕ on two tapes that computes $(w'_n)_n$.

We define the computation of ϕ on input n inductively:

- compute $\phi(n-1)$ to obtain the value of w'_{n-1} and $g(n-1)$ (they may be respectively empty and 0, for example when $n = 0$);
- draw an hypercube of side $s_n^{lim} = n^{\frac{1}{d}(d-\frac{1}{4})}$;
- on the second tape, compute and store the value $t_n^{lim} = n^{-d} 2^{n^{d-\frac{1}{4}}}$ and initialise a counter with value 0;

- simulate ϕ_0 on input $g(n-1)+1$. Between each simulated step, increment the counter and compare it with t_n^{lim};
- if the counter reaches the value t_n^{lim} before ϕ_0 halts, output $\phi(n-1)$;
- if on the other case ϕ_0 halts first, output w'_n (the output of ϕ_0) and $g(n) = g(n-1)+1$.

Hence g is surjective and non-decreasing, and $(w'_n)_n$ is by construction computable in space $n^{d-\frac{1}{4}}$. Computing the value of s_n^{lim} and t_n^{lim} can be done in $O(2^{n^{d-\frac{1}{4}}})$ operations. Since t_n^{lim} and the counter have length less than n^d, each incrementation and comparison step takes $O(n^d)$ steps at most. Therefore we have $O(n^d t_n^{lim}) = O(2^{n^{d-\frac{1}{4}}})$ operations outside of the recursive call, and the recursive call takes $O(\sum_i^{n-1} 2^{i^{d-\frac{1}{4}}}) = O(2^{n^{d-\frac{1}{4}}})$ operations. □

To conclude, the described computations are doable within these time and space constraints, and w_n is computed before time t_{n+1}. At time t_{n+1} the second machine enters a special set of states that triggers various processes: organism-building signals, body-building, and the object of the next section, a copying process that will write concatenated copies of the pattern w_n all over the main layer of the territory of the organism.

The alphabet $\mathcal{A}_{\text{comp}}$ is thus $\{0,1,2,\#\}\times\{0,1,2,\#\}\times(Q_3\cup\#)\times(\Gamma_3\cup\#)\times(Q_4\cup\#)\times(\Gamma_4\cup\#)$, where Q_i, Γ_i are the state space and the tape alphabet of (the compressed version of) the i-th Turing machine described above.

3.4.2 Copying

The copy layer aims at copying the pattern w_n output by the computational layer on the whole territory of the organism. In this section, auxiliary symbols belong to the copy layer $\mathcal{A}_{\text{copy}}$ but the pattern is written in the main layer with alphabet $\mathcal{B}$.

Writing grid Remember that we assume $w_n \in \mathcal{A}^{[0,k]^d}$ for some $\frac{1}{2}n^{\frac{d-\frac{1}{2}}{d}} < k \leq n^{\frac{d-\frac{1}{2}}{d}}$. Assume the central heart is located at 0 for readability, and that the borders of w_n have been marked with a special symbol $\boxed{G}$ (on the copy layer) by the Turing machine. The copying process relies on an (imaginary) cubic grid of side length k that covers the whole territory of the organism. Starting from the cells centred on the heart, the pattern w_n is copied in each cell of this grid passing from neighbour to neighbour, through a translation of vector ke_j or $-ke_j$ for each $1 \leq j \leq d$.

For some coordinates $i \in \mathbb{Z}^d$, define the corresponding grid element $\Sigma_i = \{\sum_{1\leq j\leq d}\alpha_j e_j : \forall 1 \leq j \leq d, ki_j \leq \alpha_j \leq ki_j + k\}$, and $\overline{\Sigma_i}$ its border (extremal cells). Notice that the computed pattern is supported by $\Sigma_{0,\dots,0}$.

Local copy operation For $u \in \mathcal{U}nit(d)$ and $i \in \mathbb{Z}^d$, we define the *copy operation* $\mathcal{C}_i(u)$ that copies the contents of the main layer from Σ_i to Σ_{i+u}. It consists in simulating a Turing machine (see previous section) that receives as input k the side length of w_n and works as follows:

Reproducing the borders: The first step is to write $\boxed{G}$ in every cell of $\overline{\Sigma_{i+u}}$. It then travels to the coordinate $k(i+u)$ and builds an hypercube of symbols $\boxed{G}$ of side length k (corresponding to $\overline{\Sigma_{i+u}}$). This takes $O(k^d)$ time steps. If the new hypercube is not entirely included in the territory of the organism, the copy process stops.

Reproducing the pattern: The second step is to copy the pattern letter by letter. The machine copies each letter in lexicographic order, marking with a symbol letters already copied. Each letter needs at most $O(k)$ steps to be copied, so the whole process takes $O(k^{d+1})$ steps.

Cleaning the auxiliary states: The third step is to remove all the auxiliary states that remain on the tape in the original grid hypercube Σ_i (including $\boxed{G}$). This is done by going through all k^d cells of Σ_i, taking $O(k^d)$ steps.

Selecting heirs: The process spawns new copy processes from the hypercube Σ_{i+u}, transmitting along the value of k. The direction of the next processes are the following:

$$\mathcal{C}_i(u) \to \begin{cases} \{C_{i+u}(v) : v = \lambda_j e_j + \sum_{k \neq j} \lambda_k e_k, \lambda_k \in \{-1, 0, +1\}\} & \text{if } u = \lambda_j e_j, \lambda_j \in \{-1, +1\} \\ C_{i+u}(u) & \text{otherwise} \end{cases}$$

Those new processes are performed in parallel by duplicating the copy layer 2^d times.

At the initial step, it is enough to trigger a copy process in all directions $u \in \mathcal{U}nit(d)$. The copying operations then progressively fill the whole organism, as can be seen in Figure 8.

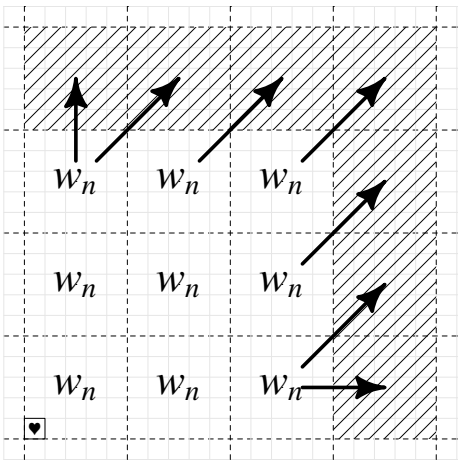

Fig. 8 In this 2-dimensional example, the pattern w_n is copied from the heart of the organism towards its boundaries in successive steps.

Each copying operation takes $O(k^{d+1})$ steps, and the active copying operations expand outward from the heart as a (thick) hypercube. Therefore, if the radius of the organism is r, the total time needed to finish the copying process is $\frac{r}{k} \cdot O(k^{d+1})$. We can take $r \leq 2^{n^{d-\frac{1}{2}}}$ by Lemma 9 and $k \leq n^{\frac{d-\frac{1}{2}}{d}}$, which gives a total time of $O(n^d 2^{n^{d-\frac{1}{2}}}) = O(2^{n^{d-\frac{1}{4}}})$. Lowering if needed the multiplicative constant by the linear speedup theorem, we see that the process ends before time t_{n+1}.

3.5 Proof of the main theorem

We first prove that the density of auxiliary states tend to 0 as time tends to infinity, which ensures they are not charged by any limit measure.

Lemma 11 *Border symbols* $\boxed{\$}$ *have negligible density asymptotically, i.e., for any nondegenerate Bernoulli measure* μ*:*

$$F^t\mu\left(\left\{c \in \mathcal{A}^{\mathbb{Z}^d} : p_{\text{orga}}(c)_0 = \boxed{\$}\right\}\right) \to_t 0.$$

Proof By Lemma 4, we can consider only the cells in the colonised space, i.e. inside a living membrane. Given a configuration c and some time $t_n \leq t < t_{n+1}$ during the n-th generation, denote $S_{\boxed{\$}}(t) = \{x \in \mathbb{Z}^d : p_{\text{orga}}(F^t(c)_x) = \boxed{\$}\} \cap \mathbf{Col}_t(c)$ the set of colonised cells containing a border, and $\overline{S}_{\boxed{\$}}(t)$ the complement of the previous set. We show that there exists a constant λ such that $\frac{|S_{\boxed{\$}}(t)|}{|\overline{S}_{\boxed{\$}}(t)|} \leq \frac{\lambda}{n}$. This property being true for every initial configuration c, the lemma follows using Birkhoff's theorem.

To do this, we show in the next Lemma that borders between organisms can be partitioned into (subsets of) hyperplanes. To each such hyperplane subset bordering two organisms, we associate some volume inside the territory of one of the organisms that is $\frac{n}{\lambda}$ times larger than the subset itself.

Lemma 12 *The common border of two hearts is included in a (finite) union of hyperplanes of $\mathcal{H}yp(d)$. (Recall $\mathcal{H}yp(d)$ is the set of hyperplanes with a normal vector in $\mathcal{U}nit(d)$.)*

Proof Let x and x' be two cells containing a heart each. If no other heart existed in the whole space, any point y in the border between these two hearts would verify $|d_\infty(x,y) - d_\infty(x',y)| \leq 1$. Taking i and j such that $d_\infty(x,y) = |y_i - x_i|$ and $d_\infty(x',y) = |y_j - x'_j|$, this border would be included in $\bigcup_{\substack{1 \leq i,j \leq d \\ \varepsilon = \pm 1}} H^\varepsilon_{i,j}$, where $H^\varepsilon_{i,j}$ are defined as:

$$y \in H^\varepsilon_{i,j} \iff \begin{cases} |2y_j - x_j - x'_j| \leq 1 & \text{if } i = j \text{ and } x_j = x'_j \\ |y_i - x_i - \varepsilon(y_j - x'_j)| \leq 1 & \text{otherwise} \end{cases}$$

All these sets are unions of one or two hyperplanes of $\mathcal{H}yp(d)$ (depending on the parity of $x_i - \varepsilon x'_j$). In the presence of other hearts, the border between x and x' is a subset of this "ideal border", which proves the Lemma. □

Given two hearts at cells x_0 and x_1, denote $B(x_0, x_1)$ the set of cells corresponding to their common border. Partition this set into a finite collection $\{H_1, \ldots, H_k\}$ of disjoint subset of hyperplanes according to the previous lemma. For each such H_i, as $d_\infty(x_0, x_1) \geq 2n$, we have either $d_\infty(x_0, H_i) \geq n$ or $d_\infty(x_1, H_i) \geq n$.

Take any i and any finite subset s of H_i. Denote $A(s)$ the area of s and V_0 and V_1 the volumes of the d-polytopes limited by the surface s and the points x_0 and x_1, respectively. Then $V_\iota = \frac{1}{d} d_\infty(x_\iota, H_i) A(s)$ for each $\iota \in \{0,1\}$. Denote $V(s) = V_0 + V_1$, then $\frac{A(s)}{V(s)} \leq \frac{d}{n}$.

We now do this operation for every organism, that is split $S_{\boxed{\$}}(t)$ into a collection $\mathcal{S}$ of disjoint hyperplanar surfaces that belong to the common border of two organisms. For every two such different surfaces, the corresponding volumes inside organisms are also disjoint, hence

$$\frac{|S_{\boxed{\$}}(t)|}{|\overline{S}_{\boxed{\$}}(t)|} \leq \frac{\sum_{\mathcal{S}} A(s)}{\sum_{\mathcal{S}} V(s)} \tag{1}$$

$$\leq \frac{d}{n} \tag{2}$$

□

Lemma 13 *For any nondegenerate Bernoulli measure μ and $z \in \mathbb{Z}^d$,*

$$\mu\left(F^t(c)_z \in \mathcal{A} \setminus \mathcal{B}\right) \underset{t \to \infty}{\longrightarrow} 0.$$

Proof We handle each layer separately.

Uncolonised space and membranes First, by Lemma 4, we can see that c_0 belongs to the uncolonised space at time t only if the nearest viable seed at time 0 is at distance more than $\sqrt{t}$. For the same reason, c_0 can be part of a living membrane or a related process (age counter, respiration process, comparison process) only if the nearest viable seed is at distance more than $\sqrt{t} - \log t$. Since a viable seed appear with a nonzero probability, the probability of this event tends to 0 as t tends to infinity.

It remains to handle symbols appearing inside the colonised space on the layers dealing with internal affairs of the colonies: organism, evolution, computing, copying and main layers. By Birkhoff's ergodic theorem, it is equivalent to prove that the density of auxiliary states in a configuration tends to 0 almost surely when time tends to infinity.

Hearts, computing symbols In the colonised space hearts $\boxed{\heartsuit}$ must be issued from a seed, and as explained in Section 3.3 they each have at time t_n a body, which are non-overlapping hypercubes of side $2n+1$ centred on the heart (more precisely, they can overlap shortly but are destroyed before the next t_n). Thus the density of hearts $\boxed{\heartsuit}$ in c between times t_n and t_{n+1} is less than $\frac{1}{(2n-1)^d}$. Since the computing process taking place around the heart is contained in a hypercube of side $\sqrt{n}$, the density of cells with nonempty computing layer is almost surely less than $\frac{1}{(2n-1)^{d/2}}$ in this period.

Bodies and bodybuilding signals Body symbols form the surface of an hypercube of side $2n+1$ (when it is fully built) or less (during the construction), and therefore there are less than $2d(2n+1)^{d-1}$ such symbols for each heart. The impulses used to grow the body being sent one at a time, they occupy at most as much space as the body itself at any given time. Therefore all those symbols have density less than $2d(2n+1)^{d-1} \cdot \frac{1}{(2n-1)^d} = O\left(\frac{1}{n}\right)$.

Borders and border-building signals Borders $\boxed{\$}$ were handled in Lemma 11. We use a similar argument to show that the density of symbols in signals used to build borders is asymptotically negligible. The signal is born around the heart and progresses at speed one. Therefore, m steps after its birth, the set of cells in the organism containing the signal is an hypercube of side $2m+1$ centred on the heart (intersected with the inside of the organism). In particular, in an organism of healthy size, the signal sent at time t_n has disappeared before time $t_n + 2^{n^d} \leq t_{n+1}$, so at most one signal appears in a given organism at the same time.

If $m \leq n$, since the organism contains at least n^d cells, signal symbols have density less than $\frac{2d(2m+1)^{d-1}}{n^d} = O(\frac{1}{n})$. If $m > n$, notice that for each cell z of the organism satisfying $d_\infty(\boxed{\heartsuit}, z) = m$, the line between $\boxed{\heartsuit}$ and z is contained in the organism (by Lemma 5) and does not contain other signal symbols (since its distance to the heart is less than m). For any part P of the surface area of the hypercube which is inside the organism, the convex hull of P and $\boxed{\heartsuit}$ is inside the organism as well and its interior does not contain any symbol. The proportion of P to the total surface area is the same as the proportion of its convex hull to the total volume. Therefore the symbol density is at most $\frac{2d(2m+1)^{d+1}}{m^d} = O\left(\frac{1}{n}\right)$ (since $m > n$).

Copying processes The copying grid $\boxed{G}$ is simply a grid of side length $\sqrt{n}$, and therefore the density of symbols $\boxed{G}$ is less than $\frac{2d(2n+1)^{d-1}}{n^d} = O\left(\frac{1}{n}\right)$. Each copying operation contains symbols in at most two squares at any given time: one from which it copies and one to which it copies. Furthermore, because all copying operations take the same amount of time $C(n)$ to copy one square, the whole copying process of an organism in the time interval $[t_n + kC(n), t_n + (k+1)C(n)]$ is contained in the squares located at "distance" k and $k+1$ from the heart, i.e. the cells

whose distance from the heart is between $k\sqrt{n}$ and $(k+2)\sqrt{n}$. Furthermore, by Lemma 9, the probability that an organism contains only one copying process tends to 1. The previous argument (used for the border-building signals) shows that copying symbols have density at most $O\left(\frac{2\sqrt{n}}{n}\right) = O\left(\frac{1}{\sqrt{n}}\right)$. □

From this lemma, we see that no limit measure can assign a non-zero probability to any pattern with a non-empty auxiliary layer.

Lemma 14

$$d_{\mathcal{M}}(F^{t_n}\mu, \widehat{\delta_{w_n}}) \underset{n\to\infty}{\longrightarrow} 0 \qquad \textit{and} \qquad \max_{t_n \le t \le t_{n+1}} d_{\mathcal{M}}\left(F^t\mu, [\widehat{\delta_{w_n}}, \widehat{\delta_{w_{n+1}}}]\right) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof Take any finite square pattern $u \in \mathcal{A}^{[0,\ell]^d}$. From Lemma 13 and by σ-invariance, we can see that if c is drawn according to μ then the probability that $F^t(c)_{[0,\ell]^d}$ has any part outside the colonised space or with a nonempty auxiliary layer is $O\left(\frac{1}{\sqrt{n}}\right)$. Inside any organism at time $t = t_n$, the main layer contains concatenated copies of w_{n-1} in all directions except for those cells at distance more than $\frac{t_n - t_{n-1}}{C(n)}\sqrt{n}$ from the heart (see the last paragraph of the previous proof), which forms an asymptotically negligible set by Lemma 9. By σ-invariance, we obtain that:

$$|F^{t_n}\mu([u]) - \widehat{\delta_{w_n}}([u])| \underset{n\to\infty}{\longrightarrow} 0.$$

Since this is true for any square pattern, we get the first part of the result.

At time t_n, the copying process for w_n is triggered. As explained in the last paragraph of the previous proof, between times $t_n + kC(n)$ and $t_n + (k+1)C(n)$ the copying process is contained in cells at distance $k\sqrt{n}$ to $(k+2)\sqrt{n}$ from the nearest heart. In particular, the main layers of cells at distance less than $k\sqrt{n}$ from the nearest heart contain concatenated copies of w_n while those at distance more than $(k+2)\sqrt{n}$ still contain concatenated copies of w_{n-1}.

Therefore, denoting by $h(c)$ the minimum distance between 0 and an heart in c, we have for any $t_n \le t \le t_{n+1}$:

$$F^t\mu([u]) = \mu\left(h(c) \le \frac{t-t_n}{C(n)}\sqrt{n}\right) \cdot \widehat{\delta_{w_n}}([u]) + \mu\left(h(c) > \frac{t-t_n}{C(n)}\sqrt{n}\right) \cdot \widehat{\delta_{w_{n-1}}}([u]) + \underset{n\to\infty}{o}(1).$$

The second term contains $\mu\left(h(c) > \frac{t-t_n}{C(n)}\sqrt{n}\right)$ instead of the expected $\mu\left(h(c) > \left(\frac{t-t_n}{C(n)} + 2\right)\sqrt{n}\right)$ to get an actual barycentre, the difference between them is asymptotically negligible in n. This equation holding for any square pattern u, we obtain:

$$d_{\mathcal{M}}\left(F^t\mu\ ,\ \mu\left(h(c) \le \frac{t-t_n}{C(n)}\sqrt{n}\right) \cdot \widehat{\delta_{w_n}} + \mu\left(h(c) > \frac{t-t_n}{C(n)}\sqrt{n}\right) \cdot \widehat{\delta_{w_{n-1}}}\right) \underset{n\to\infty}{\longrightarrow} 0. \qquad (3)$$

The right-hand measure belonging to the segment $[\widehat{\delta_{w_{n-1}}}, \widehat{\delta_{w_n}}]$, and this being true for any $t_n \le t \le t_{n+1}$, we obtain the desired result. □

Proof (of Theorem 2) By the right-hand part of Lemma 14, we see that $\mathcal{V}(F,\mu)$ is included in the closure of the polygonal path delineated by the sequence $(\widehat{\delta_{w_n}})_{n\in\mathbb{N}}$. We prove the other inclusion.

Take any $\nu \in [\widehat{\delta_{w_{n-1}}}, \widehat{\delta_{w_n}}]$. For $t_{n-1} \le t \le t_n$, denote μ_t the closest point to $F^t\mu$ in $[\widehat{\delta_{w_{n-1}}}, \widehat{\delta_{w_n}}]$; by Lemma 14, $d_{\mathcal{M}}(F^t\mu, \mu_t) \to 0$. We prove that ν is close to one of the μ_t. By Equation (3),

we have $d_{\mathcal{M}}\left(F^t\mu, F^{t+1}\mu\right) \leq 2\mu\left(\frac{t-t_n}{C(n)} \leq h(c) \leq \frac{t+1-t_n}{C(n)}\right) + \underset{n\to\infty}{o}(1) \to 0$. Since $d_{\mathcal{M}}\left(F^t\mu, \mu_t\right) = \underset{n\to\infty}{o}(1)$, it follows that $d_{\mathcal{M}}(\mu_t, \mu_{t+1}) \underset{n\to\infty}{o}(1)$ as well.

Since $d_{\mathcal{M}}\left(\mu_{t_i}, \widehat{\delta_{w_i}}\right) = \underset{n\to\infty}{o}(1)$ for any i and $d_{\mathcal{M}}(\mu_t, \mu_{t+1}) \to 0$, it follows that

$$\min_{t_n \leq t \leq t_{n+1}} d_{\mathcal{M}}(\nu, \mu_t) = \underset{n\to\infty}{o}(1) \qquad \text{and thus} \qquad \min_{t_n \leq t \leq t_{n+1}} d_{\mathcal{M}}(\nu, F^t\mu) = \underset{n\to\infty}{o}(1).$$

This proves the other inclusion. □

4 Statement of the results

From Theorem 2 we deduce a number of results which are our main contributions.

Corollary 1 *The measures $\nu \in \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ for which there exist:*

- *an alphabet $\mathcal{B} \supset \mathcal{A}$,*
- *a cellular automaton $F : \mathcal{B}^{\mathbb{Z}^d} \to \mathcal{B}^{\mathbb{Z}^d}$, and*
- *a non-degenerate Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{B}^{\mathbb{Z}^d})$*

such that $F^t\mu \xrightarrow[t\to\infty]{} \nu$, are exactly the limit-computable measures.

Corollary 2 *The connected sets of measures $\mathcal{K} \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ for which there exist:*

- *an alphabet $\mathcal{B} \supset \mathcal{A}$,*
- *a cellular automaton $F : \mathcal{B}^{\mathbb{Z}^d} \to \mathcal{B}^{\mathbb{Z}^d}$, and*
- *a non-degenerate Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{B}^{\mathbb{Z}^d})$*

such that $\mathcal{V}(F, \mu) = \mathcal{K}$, are exactly the Π_2-computable, connected, compact sets of measures.

Furthermore, both corollaries hold if one requires the convergence to hold for all nondegenerate Bernoulli measures.

Proof Apply Theorem 2 to Proposition 4. To get Corollary 1, use the fact that ν is a limit-computable measure if and only if the singleton $\{\nu\}$ is a Π_2-computable set of measures (and of course connected). □

Following [9], we obtain a similar characterisation using convergence in Cesàro mean (Corollary 5 in op.cit.) and a Rice-style theorem on μ-limit measures set (Corollary 7 in op.cit.). Since the proofs of op.cit. only involve finding an appropriate uniformly computable sequence (w_n) without modifying the cellular automaton, they can be carried straightforwardly to the d-dimensional case by replacing $\mathcal{A}^{\mathbb{Z}}$ by $\mathcal{A}^{\mathbb{Z}^d}$ and we do not repeat them here.

Corollary 3 *The sets of measures $\mathcal{K}' \subset \mathcal{K} \subset \mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$ for which there exist:*

- *an alphabet $\mathcal{B} \supset \mathcal{A}$,*
- *a cellular automaton $F : \mathcal{B}^{\mathbb{Z}^d} \to \mathcal{B}^{\mathbb{Z}^d}$, and*
- *a nondegenerate Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{B}^{\mathbb{Z}^d})$*

such that $\mathcal{V}(F, \mu) = \mathcal{K}$ and $\mathcal{V}'(F, \mu) = \mathcal{K}'$, are exactly the Π_2-computable, connected, compact sets of measures.

In particular we characterise all sets of measures reachable at the limit in convergence in Cesàro mean from a Bernoulli measure, since those sets are necessarily connected (Section 1.2.3 in op.cit.). Here again, the result holds if one requires the convergence to hold for all nondegenerate Bernoulli measures.

Corollary 4 *Let P be a nontrivial property (i.e. not always or never true) on non-empty Π_2-computable, compact, connected sets of $\mathcal{M}_\sigma(\mathcal{A}^{\mathbb{Z}^d})$. There is no algorithm that can decide, given an alphabet $\mathcal{B}$, a cellular automaton $F : \mathcal{B}^{\mathbb{Z}^d} \to \mathcal{B}^{\mathbb{Z}^d}$ and a Bernoulli measure $\mu \in \mathcal{M}_\sigma(\mathcal{B}^{\mathbb{Z}^d})$, whether $\mathcal{V}(F, \mu)$ satisfies P.*

Here it is assumed that the Bernoulli measure is finitely described by a list of (rational) parameters. A similar statement follows on nontrivial properties of limit-computable measures. This corollary would also hold if the property was required to hold, not only for one, but for some or all nondegenerate Bernoulli measure(s).

5 Open questions

The main questions that remain open concern the characterisation of non-connected sets of limit measures and the extension to more general sets of initial measures. In particular, the result in the 1-dimensional case holds for a large diversity of initial measures (σ-mixing with full support), which we could not obtain for the lack a finer analysis of the disappearance rate of the hearts.

A longer-term research direction concerns surjective cellular automata. The construction developed here is intrinsically non-surjective, and it does not seem that it can be adapted easily (in particular due to the key role of a time 0). Surjective cellular automata are known to be Turing-universal in the classical sense, but surjectivity has a deep impact on the dynamics of the model which is specific to the probabilistic setting [11]. In some sense, the question is whether this dynamical restriction is strong enough to lower the computing power of the model. Even seemingly simple questions, such that the existence of a non fully-supported set of limit measures, remain open.

References

1. Ballier, A., Guillon, P., Kari, J.: Limit sets of stable and unstable cellular automata. Fundamenta Informaticae **110**(1-4), 45–57 (2011)
2. Boyer, L., Delacourt, M., Poupet, V., Sablik, M., Theyssier, G.: μ-limit sets of cellular automata from a computational complexity perspective. Journal of Computer and System Sciences **81**(8), 1623–1647 (2015)
3. Boyer, L., Delacourt, M., Sablik, M.: Construction of μ-limit sets. In: Proc. JAC 2010, pp. 76–87 (2010)
4. Boyer, L., Poupet, V., Theyssier, G.: On the complexity of limit sets of cellular automata associated with probability measures. In: Proc. MFCS 2006, pp. 190–201 (2006)
5. Delacourt, M.: Rice's theorem for μ-limit sets of cellular automata. In: Proc. ICALP 2011 (2), pp. 89–100 (2011)
6. Delacourt, M., Hellouin de Menibus, B.: Construction of μ-limit sets of two-dimensional cellular automata. In: Proc. STACS 2015, pp. 262–274 (2015)
7. Hedlund, G.A.: Endomorphisms and automorphisms of the shift dynamical system. Mathematical Systems Theory **3**(4), 320–375 (1969)
8. Hellouin de Menibus, B.: Asymptotic behaviour of cellular automata: computation and randomness. Ph.D. thesis, Aix-Marseille University (2010)
9. Hellouin de Menibus, B., Sablik, M.: Characterisation of sets of limit measures after iteration of a cellular automaton on an initial measure (2013). Accepted to *Ergodic Theory and Dynamical Systems*
10. Kari, J.: Rice's theorem for the limit sets of cellular automata. Theoretical Computer Science **127**, 229–254 (1994)
11. Kari, J., Taati, S.: Statistical mechanics of surjective cellular automata. Journal of Statistical Physics **160**(5), 1198–1243 (2015). DOI 10.1007/s10955-015-1281-2. URL `http://dx.doi.org/10.1007/s10955-015-1281-2`

12. Kůrka, P., Maass, A.: Limit sets of cellular automata associated to probability measures. Journal of Statistical Physics **100**(5-6), 1031–1047 (2000)
13. Maass, A.: On the sofic limit sets of cellular automata. Ergodic Theory and Dynamical Systems **15**(4), 663–684 (1995)
14. von Neumann, J.: Theory of self-reproducing automata. University of Illinois Press (1966)
15. Petersen, K.: Ergodic theory, *Cambridge Studies in Advanced Mathematics*, vol. 2. Cambridge University Press (1983)
16. Zheng, X., Weihrauch, K.: The arithmetical hierarchy of real numbers. Mathematical Logic Quarterly **47**(1), 51–65 (2001)
17. Ziegler, M.: Computability and continuity on the real arithmetic hierarchy and the power of type-2 nondeterminism. In: Proc. CiE 2005, pp. 562–571 (2005)