

Distortion in one-head machines and cellular automata

Pierre Guillon¹ and Ville Salo²

¹ Université d'Aix-Marseille
CNRS, Centrale Marseille, I2M, UMR 7373
13453 Marseille, France

`pierre.guillon@math.cnrs.fr`

² University of Turku
`vosalo@utu.fi`

Abstract. We give two families of examples of automorphisms of subshifts that are *range-distorted*, that is, the radius of their iterations grows sublinearly. One of these families comes from one-head machines, and allows us to build such automorphisms for the full shift, and to obtain undecidability results. We also give some conditions on the functions that can occur as such growths.

1 Introduction

In this article, ‘distortion’ means that something that typically grows or moves linearly or not at all instead does so at an intermediate rate. In one-head machines, we consider sublinear head movement (the head visits $o(t)$ cells in t steps), and in cellular automata sublinear radius growth (the radius of the iterates grows in $o(t)$), which corresponds to range distortion in the terminology of [1]. In both cases, we show ‘trichotomy’ results: there are logarithmic gaps between periodic and distorted cases, and between distorted and positive-speed machines.

We show that every aperiodic one-head machine is distorted. The existence of aperiodic one-head machines is well-established, in particular [2] shows that they not only exist but form a computationally hard (undecidable) set. The single most beautiful example of an aperiodic machine is probably the SMART machine [3], whose moving tape dynamics is even minimal.

We discuss two ways of achieving distortion in automorphism groups of subshifts. To every one-head machine, we can associate a cellular automaton (on a full shift) whose radius grows at roughly the same speed as the head of the one-head machine moves. Given that there exist distorted one-head machines, there also exist distorted cellular automata. The examples given are reversible, and thus we obtain distorted automorphisms on a transitive subshift, answering an implicit question of [1]. By known embedding theorems, we obtain such examples on all uncountable sofic shifts.

We also construct an example of an automorphism on a general subshift with ‘highly unbalanced distortion’, in the sense that for an infinite set of times $t \in \mathbb{N}$,

f^t has a ‘right-leaning’ neighborhood (one of the form $\llbracket a, \infty \rrbracket$) that contains only slightly more than logarithmically many cells to the left of the origin, and ‘left-leaning’ neighborhood with the symmetric property, yet all its two-sided neighborhoods grow at an almost linear rate. In particular, the intersection of all neighborhoods is far from being a neighborhood, answering Question 3.26 of [1].

2 Definitions

2.1 Subshifts and cellular automata

Let Σ be a finite set called the *alphabet*. Then $\Sigma^{\mathbb{Z}}$ with the product topology is called the *full shift*, and it is a \mathbb{Z} -dynamical system under the shift map $\sigma : \Sigma^{\mathbb{Z}} \rightarrow \Sigma^{\mathbb{Z}}$ defined by $\sigma(x)_i = x_{i+1}$. Closed shift-invariant subsets of it are called *subshifts*.

If X and Y are subshifts, a function $f : X \rightarrow Y$ is called a *morphism* if it is continuous and $\sigma \circ f = f \circ \sigma$. It is an *endomorphism* if $Y = X$ and an *automorphism* if, besides, it is bijective (in which case it automatically has a left and right inverse endomorphism). A *cellular automaton* is another name for an endomorphism, though often this term is reserved for the case $X = Y = \Sigma^{\mathbb{Z}}$. Automorphisms are also called *reversible cellular automata*.

An endomorphism f is *preperiodic* if $f^{p+q} = f^q$ for some preperiod $q \in \mathbb{N}$ and some period $p \in \mathbb{N} \setminus \{0\}$. If an automorphism is preperiodic with $q = 0$, it is periodic.

The trace map $T_f : X \rightarrow \Sigma^{\mathbb{N}}$ is the map defined by $T_f(x)_t = f^t(x)_0$ for all $x \in X$ and $t \in \mathbb{N}$. It is clear that $\tau_f = T_f(X)$ is a one-sided subshift (closed and shift-invariant), which is finite if and only if f is preperiodic.

For X a subshift and $n \in \mathbb{N}$ we define the *complexity* function \mathcal{K}_X by $\mathcal{K}_X(n) = |\{x_{\llbracket 0, n \rrbracket} \mid x \in X\}|$, the number of distinct patterns occurring in configurations of X . It is easy to see that if X is infinite, then \mathcal{K}_X is increasing.

2.2 Neighborhoods and radii

It is quite well-known [4] that if $f : X \rightarrow Y$ is a morphism, then it admits a neighborhood, that is a finite interval $I \subset \mathbb{Z}$ such that $\forall x, y \in X, x_I = y_I \implies f(x)_0 = f(y)_0$.

Let X and Y be \mathbb{Z} -subshifts and $f : X \rightarrow Y$ a morphism. We define the set of *neighborhoods* as

$$N(f) = \{I = \llbracket a, b \rrbracket \subset \mathbb{Z} \mid \forall x, y \in X, x_I = y_I \implies f(x)_0 = f(y)_0\}.$$

The *diameter* $D(f)$ of a morphism f is then the least possible diameter $2r+1$ of a central neighborhood $\llbracket -r, r \rrbracket \in N(f)$.

Remark 1. It is easy to see that $N(f)$ is an upset: $I \in N(f), J \supset I \implies J \in N(f)$.

The case when $N(f)$ is a principal filter, that is when $N(f) = \{J \mid J \supset I\}$ for some finite interval I (it is well-known that this happens for the full shift), is especially desirable. In that case, I must be the intersection of all elements in $N(f)$, and thus we define $I(f) = \bigcap_{J \in N(f)} J$ (it corresponds to $\mathcal{I}(-1, f)$ in the notation of [1]). Let us also define $d(f)$ as the diameter of $\{0\} \cup I(f)$, which is at most $D(f)$ (and is equal if X is the full shift for example). Theorem 5 will give an example of endomorphism where $N(f)$ is far from being a filter.

2.3 Distortion

Let f be an endomorphism of a subshift X . For $t \in \mathbb{N}$, let us define $D_t(f)$ as $\max_{k \leq t} D(f^k)$. It is clear that $D_t(f) \leq t(D(f) - 1) + 1 = O(t)$, and that $D_t(f) \leq D_{p+q}(f)$ is bounded if $f^{p+q} = f^q$. f will be called *range-distorted* (or simply, in this article, *distorted*) if $D_t(f) = o(t)$ but f is not preperiodic.

This definition is equivalent to the one from [1, 5], and comparable to the notion of distortion from group theory: if f^t can be expressed as a product of $o(t)$ generators of some finitely generated endomorphism submonoid, then f is range-distorted.

The distortion function $t \mapsto D_t(f)$ cannot be arbitrarily low. In fact, naively counting the possible local rules gives a $\log \log$ lower bound, but the Morse-Hedlund theorem allows to ‘remove’ one log in the following proposition, which is a direct adaptation from the main argument in [5, Theorem 3.8].

Proposition 1. *If X is a subshift and $f : X \rightarrow X$ an endomorphism. Then exactly one of the following holds:*

- (D_t) is bounded (f is preperiodic);
- $\forall t \in \mathbb{N}, D_t(f) \geq \mathcal{K}_X^{-1}(\mathcal{K}_{\tau_f}(t)) > \mathcal{K}_X^{-1}(t) = \Omega(\log t)$ and $D_t(f) = o(t)$ (f is distorted);
- $D_t(f) = \Theta(t)$, and $d(f^t) = \Theta(t)$ (f has non-0 Lyapunov exponents).

Note that if X has linear complexity or if f has positive entropy, then the central class is empty. Moreover, if $\mathcal{K}_X(n) = O(n^d)$ (resp. $O(2^{n^\epsilon})$), then endomorphisms of this class must even have $D_t(f) = \Omega(t^{1/d})$ (resp. $\Omega((\log t)^{-\epsilon})$).

Proof. Let $t \in \mathbb{N}$. If f is not preperiodic, then neither is its trace τ_f . By the Morse-Hedlund theorem, we must have $\mathcal{K}_{\tau_f}(t) > t$. By definition, D_t can be written as $2r + 1$ such that for all $x, y \in X$ such that $x_{[-r, r]} = y_{[-r, r]}$, we have $T_f(x)_{[0, t]} = T_f(y)_{[0, t]}$, so that $\mathcal{K}_{\tau_f}(t) \leq \mathcal{K}_X(D_t)$. We obtain $D_t \geq \mathcal{K}_X^{-1}(\mathcal{K}_{\tau_f}(t)) > \mathcal{K}_X^{-1}(t)$ because \mathcal{K}_X is increasing.

Now, suppose that D_t is not $o(t)$, then subadditivity and the Fekete lemma imply that $D_t = \Theta(t)$ (and the same for $d(f^t)$). This argument is formalized for example in [6] or [1], and the limit of $D_t/2t$ corresponds to the maximal so-called Lyapunov exponent, in absolute value. \square

It is not known if there is the same lower gap for $d(f^t)$, or in general which kinds of growths are possible.

A natural object is the two-dimensional subshift of (two-sided) space-time diagrams of f : $\mathcal{X}_f = \{x \mid \forall t \in \mathbb{Z}, (x_{n,t})_{n \in \mathbb{Z}} = f((x_{n,t-1})_{n \in \mathbb{Z}}) \in X\}$. Following [7], we say that a vector line (or *direction*) $\ell \subset \mathbb{R}^2$ is expansive in \mathcal{X}_f if there exists a width $r \in \mathbb{N}$ such that:

$$\forall x, y \in \mathcal{X}_f : (\forall v \in \mathbb{Z}^2 : d(v, \ell) < r \implies x_v = y_v) \implies x = y.$$

The following proposition is not difficult. It is for example a particular case of [1, Proposition 4.5].

Proposition 2. *A non-periodic automorphism f and its inverse are distorted if and only if \mathcal{X}_f has the vertical direction as unique direction of nonexpansiveness.*

Actually, any 2D subshift is expansive in every nonvertical direction if and only if it is conjugate to \mathcal{X}_f for some automorphism f such that both f and f^{-1} are periodic or distorted (a particular case of [1, Proposition 5.6]).

Note that Proposition 2 could motivate a notion of directional distortion, corresponding to endomorphisms whose space-times have a unique direction of nonexpansiveness, and whose composition with the corresponding shift is not preperiodic (in particular, if the unique direction of nonexpansiveness is irrational).

Several examples of such *extremely expansive* two-dimensional subshifts are known. A general self-simulating construction is given in [8], and effectivized (so that f is obtained as a partial local rule from the full shift) in [9, 10].

We give in Section 4 a construction which is very similar to a second construction in [8], though independent. But first, in Section 3, we prove a link with one-head machines, which allows us to get distorted automorphisms of the full shift.

3 Distorted one-head machines

3.1 One-head machines

Let $\Delta = \{-1, +1\}$ be the set of directions. A *one-head machine* (or Turing machine) \mathcal{M} is a triple (Q, Σ, δ) where Q is a finite set of states, Σ is a finite set of symbols, and $\delta \subset (Q \times \Delta \times Q) \sqcup (Q \times \Sigma \times Q \times \Sigma)$ is the transition function. This model (for example introduced in [11]) is equivalent to the one in [12], but handles reversibility better.

Noting $\tilde{\Sigma} = \Sigma \sqcup (Q \times \Sigma)$, where elements of $Q \times \Sigma$ are called *heads*, and $X_{\mathcal{M}} = \left\{ x \in \tilde{\Sigma}^{\mathbb{Z}} \mid \forall i, j \in \mathbb{Z}, x_i \in \Sigma \text{ or } x_j \in \Sigma \right\}$ (set of tapes with at most one head somewhere), we can associate to it the so-called *moving-head model* as the closed shift-invariant relation $M_{\mathcal{M}}$ of $X_{\mathcal{M}}$ defined by: $(x, x') \in M_{\mathcal{M}}$ if one of the following occurs:

$$\begin{aligned} & x = x' \in \Sigma^{\mathbb{Z}}; \\ & \exists i \in \mathbb{Z}, (x_i, x'_i) \in \delta \text{ and } \forall j \neq i, x_j = x'_j \in \Sigma; \\ & \exists i, i' \in \mathbb{Z}, (q, i' - i, q') \in \delta, x_i = (q, x'_i) ; x'_{i'} = (q', x_i) ; \forall j \notin \{i, i'\}, x_j = x'_j \in \Sigma. \end{aligned}$$

We actually focus on *total deterministic* machines, that is, machines where every configuration has exactly one successor, which makes $M_{\mathcal{M}}$ induce an endomorphism of $X_{\mathcal{M}}$, also noted $M_{\mathcal{M}}$. A *reversible one-head machine (RTM)* is a deterministic one-head machine for which $M_{\mathcal{M}}$ is actually an automorphism.

A (total deterministic) one-head machine \mathcal{M} is *periodic* or *preperiodic* if the corresponding endomorphism is.

A configuration $x \in X_{\mathcal{M}}$ is *weakly periodic* if $M_{\mathcal{M}}^p(x) = \sigma^j(x)$ for some $p \geq 1$ and $j \in \mathbb{Z}$. We will say that it is *aperiodic* if it has no weakly periodic configuration containing a head³ (that is, no configuration x with a nontrivial $p \geq 1$ and $q \in \mathbb{Z}$ such that $M_{\mathcal{M}}^p(x) = \sigma^q(x)$).

Fix a one-head machine $\mathcal{M} = (Q, \Sigma, \delta)$. If $x \in \tilde{\Sigma}^{\mathbb{Z}}$ contains a head, we write $s_t(x)$ for the number of distinct cells that the head of \mathcal{M} visits in the first t steps starting from configuration x (taking the number of distinct cells rather than the position makes it nondecreasing, which simplifies some arguments). The function $m : \mathbb{N} \rightarrow \mathbb{N}$ defined by $m(t) = \max_x s_t(x)$ is called the *movement bound* of \mathcal{M} . The *speed* of \mathcal{M} is defined in [14] as the limit of $m(t)/t$, which exists by subadditivity. A one-head machine will be called *distorted* if it is not periodic but m is sublinear in t .

Remark 2. It is easy to see that $D_t(M_{\mathcal{M}}) = m(t)$. In particular, \mathcal{M} is distorted if and only if $M_{\mathcal{M}}$ is a distorted endomorphism.

We will prove that aperiodic one-head machines are examples of distorted machines.

3.2 Speed trichotomy

In this section, we give some information on the possible speeds of one-head machines, namely that there are two gaps of impossible movements.

Theorem 1. *Let $\mathcal{M} = (Q, \Sigma, \delta)$ be a one-head machine with movement bound m . Then exactly one of the following holds:*

- m is bounded (\mathcal{M} is preperiodic);
- $m(t) = \Omega(\log t)$ and $m(t) = O(t/\log t)$ (\mathcal{M} is distorted);
- $m(t) = \Theta(t)$ (\mathcal{M} has positive speed).

The preperiodic and positive speed cases are quite well understood. It can even be shown that some periodic configuration achieves the maximal speed [14]. We do not know what kinds of intermediate growth functions can be realized with distorted one-head machines.

Here is a simple counting lemma. If $H : \mathbb{N} \rightarrow \mathbb{N}$ is nondecreasing, we write $\lfloor H^{-1}(n) \rfloor$ for the largest ℓ such that $H(\ell) \leq n$.

Lemma 1. *Let a_0, \dots, a_n be in \mathbb{N} and suppose that $|\{i \in \llbracket 0, n \rrbracket \mid a_i = \ell\}| \leq h(\ell)$ for all ℓ . We have the following.*

³ This corresponds to classical aperiodicity in the so-called *moving-tape model* or *trace subshift* from [12, 13].

1. If $H(\ell) = \sum_{i=0}^{\ell} h(i)$, then $\sum_{i=0}^n a_i \geq \sum_{i=0}^{\lfloor H^{-1}(n) \rfloor} ih(i)$.
2. Moreover, if $h(\ell) = \alpha^\ell$, then $\sum_{i=0}^n a_i = \Omega(n \log n)$.

Proof.

1. Define $b_i = k \iff i \in \llbracket H(k-1), H(k) \rrbracket$ for $k \geq 0$. Then we have

$$\sum_{i=0}^n a_i \geq \sum_{i=0}^n b_i \geq \sum_{i=0}^{\lfloor H^{-1}(n) \rfloor} ih(i),$$

where the first inequality follows by sorting the a_i in increasing order and observing that then necessarily $a_i \geq b_i$ for all i , and the second follows by a direct counting argument.

2. If $\ell \leq \log_\alpha n - 1$, then $H(\ell) = \sum_{i=0}^{\ell} h(i) \leq \alpha^{\ell+1} \leq n$, so $\lfloor H^{-1}(n) \rfloor \geq \log_\alpha n - 2$. Thus

$$\begin{aligned} \sum_{i=0}^n a_i &\geq \sum_{i=0}^{\lfloor H^{-1}(n) \rfloor} ih(i) \\ &\geq \sum_{i=0}^{\log_\alpha n - 2} ih(i) \\ &\geq (\log_\alpha n - 2)h(\log_\alpha n - 2) \\ &= (\log_\alpha n - 2)\alpha^{\log_\alpha n - 2} \\ &= (\log_\alpha n - 2)n\alpha^{-2} \\ &= \Omega(n \log n). \end{aligned} \quad \square$$

The upper bound is achieved by a counting argument, and the right object to count are the crossing sequences, which we now define.

To any machine \mathcal{M} , configuration $x \in X_{\mathcal{M}}$ and position $J \subset \mathbb{Z}$, we can associate the *crossing times* $\theta_J(x)$ as the ordered set of times $k \in \mathbb{N}$ such that $\exists i \in J, M_{\mathcal{M}}^k(x)_i \in Q \times \Sigma$; it is formally a tuple, but sometimes we use set notation, like its cardinality $|\theta_J(x)|$ or diameter $\max \theta_J(x) - \min \theta_J(x)$. Moreover for all steps $t \in \mathbb{N}$, we can associate the (partial) *crossing sequence* $u_{J,t}(x) = (M_{\mathcal{M}}^k(x)_J)_{k \in \theta_{J,t}(x)} \in (\tilde{\Sigma}^J)^*$, where $\theta_{J,t}(x) = \theta_J(x) \cap \llbracket 0, t \rrbracket$. This definition is close to a finitary version of the notion in [15, 14], except we take the sequence at a given cell rather than between two neighboring cells, which makes no difference except for writing. We use notations $\theta_i, \theta_{i,t}, u_{i,t}$ if $J = \{i\}$.

We are now ready to prove the main equivalence of this section.

Proposition 3. *Let \mathcal{M} be a one-head machine. The following are equivalent.*

1. $m(n)$ is not $O(n/\log n)$.

2. There exist a configuration $x \in X_{\mathcal{M}}$, two distinct positions $i, j \in \mathbb{Z}$, and a step $t \in \mathbb{N}$ such that $u_{i,t}(x) = u_{j,t}(x)$ are nonempty.
3. There exists a configuration $x \in X_{\mathcal{M}}$ such that the cardinality of $\theta_i(x)$ is uniformly bounded for $i \in \mathbb{Z}$.
4. There exists a configuration $x \in X_{\mathcal{M}}$ such that the diameter of $\theta_i(x)$ is uniformly bounded for $i \in \mathbb{Z}$.
5. There exists a weakly periodic configuration which is not periodic.
6. \mathcal{M} has positive speed: $m(n) = \Omega(n)$.

Point 2 actually remains equivalent if the first visited crossing sequence admits the other one as a prefix. The implications $2 \Rightarrow 3$ (resp. $3 \Rightarrow 4$) could also have been derived from looking at the countable-state Markov shift built in [14] (resp. from a general result over path spaces [16]), but we give specific proofs for completeness.

Proof.

1 \Rightarrow 2 Let x be a configuration and t a step, and $J = \{i \in \mathbb{Z} \mid \theta_{i,t}(x) \neq \emptyset\}$ the set of visited cells. Since $\llbracket 0, t \rrbracket = \bigsqcup_{i \in J} \theta_{i,t}(x)$, we get that $t = \sum_{i \in J} |\theta_{i,t}|$, which is the sum of lengths of the crossing sequences. Suppose that every crossing sequence $u_{i,t}(x)$ is distinct, for $i \in J$. There are at most $(|\Sigma| |Q|)^\ell$ distinct crossing sequences of length ℓ , so it follows from Lemma 1 that $t = \Omega(|J| \log |J|)$. We get:

$$m(t) = \sup_x s_t(x) = \sup_x s_{\gamma(t) \log \gamma(t)}(x) \leq B\gamma(t) \leq 2B \frac{t}{\log t},$$

for some constant B and all large enough t , and where γ is the inverse of the function $t \mapsto t \log t$, which satisfies $\gamma(t) \leq \frac{t}{\log t - \log \log t} \leq 2 \frac{t}{\log t}$ for large t .

2 \Rightarrow 3 By symmetry, we can assume that $j > i$ and $\min \theta_{j,t}(x) > \min \theta_{i,t}(x)$. By shifting and applying $M_{\mathcal{M}}$, we can assume that $i = 0$ and $\min \theta_{i,t}(x) = 0$. We can also assume that t is minimal for the property that $u_{0,t}(x) = u_{j,t}(x)$ is not empty. Equivalently, t is the first step n for which $k_n = |\theta_{0,n}(x)| - |\theta_{j,n}(x)| = 0$. Since $k_0 = 1$ and for $n \in \mathbb{N}$, $k_{n+1} \in k_n + \{-1, 0, 1\}$, we get that $k_n > 0$ if $n < t$, which means that $t' = \max \theta_{0,t}(x) < t = \max \theta_{j,t}(x)$. Note that $t+1 \in \theta_{j+1}(x)$, because $M_{\mathcal{M}}^t(x)_j = M_{\mathcal{M}}^{t'}(x)_0$ gives a right movement by δ (if the machine head had been going to the left on 0 at $t' = \max \theta_{0,t}(x)$, then it could not have reached position j before time t).

Let $J_{-1} = \llbracket -\infty, 0 \rrbracket$ and, for $n \in \mathbb{N}$, $J_n = jn + \llbracket 0, j \rrbracket$. Let $y \in X_{\mathcal{M}}$ have a tape that is periodic on the right in the following way: $y_{J_{-1}} = x_{J_{-1}}$ and $y_{jn + \llbracket 0, j \rrbracket} = x_{\llbracket 0, j \rrbracket}$.

Let us build inductively a nondecreasing map $\phi : \mathbb{N} \times (\mathbb{N} \sqcup \{-1\}) \rightarrow \llbracket 0, t+1 \rrbracket$ such that for all steps $k \in \mathbb{N}$ and $n \in \mathbb{N} \sqcup \{-1\}$, if $\theta_{J_n, k}(y)$ is nonempty then $M_{\mathcal{M}}^k(y)_{J_n} = M_{\mathcal{M}}^{\phi(k, n)}(x)_{J_{\min(n, 0)}}$; besides, $\phi(\llbracket 0, k \rrbracket \times \{n\}) = \{0\} \cup \theta_{J_{\min(n, 0), \phi(k, n)}} \cup (\theta_{J_{\min(n, 0), \phi(k, n)} + 1}$ (in particular, this gives $u_{J_n, k}(y) = u_{J_{\min(n, 0), \phi(k, n)}}(x)$); moreover, the restriction of ϕ to $\theta_{J_n, k}(y) \times \{n\}$ is an injection onto $\theta_{J_{\min(n, 0), \phi(k, n)}}(x)$.

$\phi(0, n) = 0$ clearly satisfies this. Now, suppose that ϕ has been built up to step $k \in \mathbb{N}$ for all $n \in \mathbb{N} \sqcup \{-1\}$. Let $n \in \mathbb{N} \sqcup \{-1\}$.

- If $k, k+1 \notin \theta_{J_n}(y)$, then $\phi(k+1, n) = \phi(k, n)$ is satisfying because J_n is unchanged at this moment.
- If $k \in \theta_{J_n}(y)$, then it is clear that $\phi(k+1, n) = \phi(k, n) + 1$ satisfies the two properties. By hypothesis, $M_{\mathcal{M}}^k(y)_{J_n} = M_{\mathcal{M}}^{\phi(k, n)}(x)_{J_{\min(n, 0)}}$, so that $\phi(k, n) \in \theta_{J_{\min(n, 0)}}(x)$. By the first remark of the proof, we see that $\phi(k, n) \neq t+1$.
- Now if $k+1 \in \theta_{J_n}$ but $k \notin \theta_{J_n}$, then one can note that $k+1 \in \theta_{j_n} \cup \theta_{j(n+1)}$ (the head should be in one boundary), say $k+1 \in \theta_{j_n}$ (the other case can be dealt with by symmetry), and in that case $k \in \theta_{j_{n-1}}$, so that we have already defined $\phi(k+1, n-1) = \phi(k, n-1) + 1$, which is at most t (by hypothesis and because $\phi(k+1, n-1) \in \theta_{J_{n-1}}(y)$). We know that $M_{\mathcal{M}}^{k+1}(y)_{J_{n-1}} = M_{\mathcal{M}}^{\phi(k+1, n-1)}(x)_{J_{\min(n, 0)}}$. In particular, $M_{\mathcal{M}}^{k+1}(y)_{j_n} = M_{\mathcal{M}}^{\phi(k+1, n-1)}(x)_0$ contains the head, and the main hypothesis gives a corresponding time $\phi(k+1, n) \leq t$ such that $u_{j, \phi(k+1, n-1)}(x) = u_{0, \phi(k+1, n)}(x)$. From the construction of the previous steps, $\phi(k, n)$ is either 0 if $\theta_{J_n, k}$ is empty, or $\phi(\max \theta_{J_n, k}, n) + 1 = \max \theta_{J_n, \phi(k, n)}(x) + 1 < \phi(k+1, n)$ otherwise. In both cases, we get the wanted properties.

The last property of ϕ gives that for all $n \in \mathbb{N}$, $|\theta_{J_n, k}| = \left| \theta_{J_{\min(n, 0)}, \phi(k, n)} \right|$ and the fact that the map is bounded gives that this is at most $|\theta_{J_{\min(n, 0)}, t}| \leq t$. We obtain that the number of visited cells in the first k steps on y is $\Omega(jk/t)$.

3 \Rightarrow 4 Let $\ell \geq 1$ be minimal such that there is a configuration $x \in X_{\mathcal{M}} \setminus \Sigma^{\mathbb{Z}}$ such that $\forall i \in \mathbb{Z}, |\theta_i(x)| \leq \ell$. Assume that for all $n \in \mathbb{N}$, there exists $i_n \in \mathbb{Z}$ for which the diameter of $\theta_{i_n}(x)$ is at least n . Let us consider a limit point y of $(M_{\mathcal{M}}^{\min \theta_{i_n}(x)} \sigma^{i_n}(x))_{n \in \mathbb{N}}$. By minimality of ℓ , we know that $|\theta_i(y)| \geq \ell$ for some $i \in \mathbb{Z}$. Let $t = \max \theta_i(y)$, and n be such that for all $m \geq n, \forall k \leq t, M_{\mathcal{M}}^k \sigma^{i_m}(x)_i = M_{\mathcal{M}}^k(y)_i$. In particular, $u_{i, t}(\sigma^{i_m}(x)) = u_{i, t}(y)$ has length at least ℓ . By assumption, it actually has length ℓ and $\theta_{i+i_m}(x) = \theta_i(\sigma^{i_m}(x)) \subset \llbracket 0, t \rrbracket$. For every $m \geq \max(n, t)$, we have $\max \theta_{i_m}(x) \geq m \geq t$, so that, after time t , the head is in the connected component of $\mathbb{Z} \setminus \{i+i_m\}$ that contains i_m . Let us assume that $i > 0$ (the argument is symmetric), so that this connected component is $\llbracket -\infty, i+i_m \rrbracket$ for every m , and let j be one position taken by the head after time t . Then for all $m \geq \max(n, t), j < i+i_m$, which means $i_m > j-i$. On the other hand, if i_m is itself a position that the head takes after time t , so it must be in $\llbracket -\infty, i+i_{\max(n, t)} \rrbracket$. It results that $\{i_m \mid m \geq \max(n, t)\}$ is included in the finite set $\llbracket j-i, i+i_{\max(n, t)} \rrbracket$, which contradicts its infinity.

4 \Rightarrow 5 Let $n \geq 1$ and $x \in X_{\mathcal{M}} \setminus \Sigma^{\mathbb{Z}}$ such that for all $i \in \mathbb{Z}$, the diameter of $\theta_i(x)$ is at most n . By the pigeonhole principle, there are two distinct positions $i, j \in \mathbb{Z}$ such that $\theta_i(x) - \min \theta_i(x) = \theta_j(x) - \min \theta_j(x)$ and $u_i(x) = u_j(x)$. Assume that $i = 0 < j$ and $\min \theta_i(x) = 0 < p = \theta_j(x)$ (by symmetry). Then our assumption says that $(M_{\mathcal{M}}^t(x)_i)_{t \in \mathbb{N}} = (M_{\mathcal{M}}^t(x)_j)_{t \in \llbracket p, \infty \rrbracket}$. Let us define $y \in X_{\mathcal{M}}$ by $y_{j_m+n} = M_{\mathcal{M}}^{\max(0, -pm)}(x)_n$ if $m \in \mathbb{Z}$ and $n \in \llbracket 1, j \rrbracket$. A standard cellular automata argument (an a little drawing) can convince that the pieces of space-time diagrams fit, so that, by induction on $t \in \llbracket 0, p \rrbracket$: if

$n \in \llbracket 1, j \rrbracket$ and $m > 0$, then $M_{\mathcal{M}}^t(y)_{jm+n} = x_n = y_{j(m-1)+n}$, and if $m \leq 0$, $M_{\mathcal{M}}^t(y)_{jm+n} = M_{\mathcal{M}}^{t-pm}(x)_n$. In particular, we get that $M_{\mathcal{M}}^p(y)_{jm+n} = y_{j(m-1)+n} = \sigma^j(y)_{jm+n}$.

5 \Rightarrow 6 It is clear that any configuration x such that $M_{\mathcal{M}}^p(x) = \sigma^j(x)$ has a speed $s_t(x) \sim jt/p$.

6 \Rightarrow 1 This is obvious. □

Proof (of Theorem 1). From Proposition 1 together with Remark 2, we know that if \mathcal{M} is not preperiodic, then $m(t) = \Omega(\log t)$. The other gap corresponds to the implication 1 \Rightarrow 6 in Proposition 3. □

3.3 Aperiodic machines

Theorem 2. *Every aperiodic one-head machine is distorted.*

In particular, there exist distorted one-head machines: see for example [17, 3] for constructions of aperiodic machines. The latter is even minimal in the moving tape model (which directly implies aperiodicity, except over the trivial alphabet).

Proof. Consider the three cases of Theorem 1. If \mathcal{M} is aperiodic, it naturally cannot be preperiodic. If \mathcal{M} were in the last case of the theorem, its trace subshift would contain a periodic point $y \in (Q \times \Sigma)^{\mathbb{N}}$ with positive speed. On a configuration where this movement is realized, every cell is visited a bounded number of times, during a time interval of bounded length. Thus \mathcal{M} is essentially performing a finite transduction, and it is easy to extract, by the pigeonhole principle, a configuration in $Q \times \Sigma^{\mathbb{Z}}$ where \mathcal{M} acts periodically. □

The machine constructed in [3] also has the property that the trace subshift of the one-head machine (the subshift encoding possible sequences of states that the head can enter when acting on a configuration) has a substitutive structure, and an explicit substitution is given. As the head movement only depends on the trace, it should be possible to compute the movement bound explicitly using spectral properties of the matrix associated to the substitution (see [18]), but this requires a bit of work since the substitution given in [3] is not primitive.

3.4 Distortion on sofic shifts

The question of distortion is most interesting on simple subshifts, as then distortion comes from the automorphism itself and not the structure of the subshift. In [1], it is stated in particular that it is not known whether range-distortion can be achieved on transitive subshifts. In this section, we show that the existence of distorted one-head machines directly implies the existence of distorted automorphisms on all uncountable sofic shifts.

The following lemma is a direct corollary of the construction in [19, Lemma 7] (the result is proved for mixing SFTs in [20], with essentially the same construction).

Lemma 2. *Let X be a full shift and Y an uncountable sofic shift. Then there exist $C, B \in \mathbb{N}$ and an embedding ϕ from the endomorphism monoid of X to that of Y such that $|D(\phi(f)) - BD(f)| \leq C$ for all endomorphisms f of X .*

The number B comes from the fact that individual (pairs of) letters are written as words of length B occurring in Y , and C comes from the fact that the rule is only applied in “safe contexts”.

Lemma 3. *Let $\mathcal{M} = (Q, \Sigma, \delta)$ be a deterministic total one-head machine. Then, letting*

$$\Gamma = ((\Sigma^2 \times \Delta) \cup (Q \times \Sigma) \cup (\Sigma \times Q)),$$

there is a cellular automaton $f : \Gamma^{\mathbb{Z}} \rightarrow \Gamma^{\mathbb{Z}}$ such that if $m : \mathbb{N} \rightarrow \mathbb{N}$ is the movement bound for \mathcal{M} , then $D(f^t) \leq m(t)$ for all $t \in \mathbb{N}$. Moreover, f is reversible (resp. preperiodic) if and only if \mathcal{M} is.

The proof uses so-called ‘conveyor belts’ to deal with configurations with several heads. One could also use the construction of [2] to embed the one-head machine to a cellular automaton, and obtain the same result.

Proof (of Lemma 3). The proof is similar to that of [19, Lemma 7]. For a residual set of points $x \in \Gamma^{\mathbb{Z}}$, we can split x into a product $x = \cdots w_{-2}w_{-1}w_0w_1w_2 \cdots$ such that for each $i \in \mathbb{Z}$, we have

$$w_i \in (\Sigma^2 \times \{+1\})^* ((Q \times \Sigma) \cup (\Sigma \times Q)) (\Sigma^2 \times \{-1\})^* \cup (\Sigma^2 \times \{+1\})^* (\Sigma^2 \times \{-1\})^*,$$

and this factorization is clearly unique: every point in $\Gamma^{\mathbb{Z}}$ can be seen as a point of this form, but the leftmost and/or rightmost words can be degenerate, and have an infinite number of ± 1 . It is enough to define how f transforms these words, and if the resulting map is uniformly continuous (which will be evident from the construction), then f uniquely extends to a continuous function on the full shift. Shift-commutation follows automatically because the decomposition of x is unique and the decomposition process is shift-invariant, and thus we obtain a cellular automaton.

On words in $(\Sigma^2 \times \{+1\})^* (\Sigma^2 \times \{-1\})^*$, we do nothing. If $w \in (\Sigma^2 \times \{+1\})^* ((Q \times \Sigma) \cup (\Sigma \times Q)) (\Sigma^2 \times \{-1\})^*$, let $w' \in (\Sigma^2)^* ((Q \times \Sigma) \cup (\Sigma \times Q)) (\Sigma^2)^*$ be the word obtained from w by erasing the arrows. We see w' as a ‘conveyor belt’, wrapped around which is a word of length $2|w'|$. More precisely, let $u = \pi_1(w')$ and v equals the reversal $\overline{\pi_2(w')}$, of $\pi_2(w')$, and observe that one of these words is in Σ^+ and the other one is in $\Sigma^*Q\Sigma^*$.

Apply the transition function of the one-head machine to the configuration $(uv)^{\mathbb{Z}}$. Note that this configuration contains infinitely many heads, but as they move with the same rule, the movement is still well-defined. Note also that if the machine \mathcal{M} is reversible, then this application is reversible as well, in the sense that the inverse of \mathcal{M}^{-1} applied at every head undoes the transition step of \mathcal{M} even on this periodic configuration. (This justifies the last sentence in the statement of the lemma.)

Now, the resulting configuration $(u'v')^{\mathbb{Z}}$ still contains exactly one head in every pattern $u'v'$. This configuration was obtained by a bijection that unwrapped

a word $w' \in (\Sigma^2)^*((Q \times \Sigma) \cup (\Sigma \times Q))(\Sigma^2)^*$ to a pair of words. Perform the inverse of this bijection, rewrapping u' and v' to a word in $(\Sigma^2)^*((Q \times \Sigma) \cup (\Sigma \times Q))(\Sigma^2)^*$, and add a $+1$ and -1 component pointing towards the machine head to each cell containing a symbol in Σ^2 . This defines f . Note that x and $f(x)$ always have the same decomposition, and if the one-head machine is reversible, its reverse one-head machine defines f^{-1} , so f is reversible if \mathcal{M} is.

To see that $D(f^t) \leq m(t)$, consider any configuration $x \in I^{\mathbb{Z}}$. If there is no machine head in $x_{[-m(t), m(t)]}$, then $f^t(x)_0 = x_0$, since no machine head can travel by more than $m(t)$ cells in t steps. If there is a machine head in this interval in some coordinate $j \in \mathbb{Z}$, we start simulating its movement (also modifying the tape according to its movement). Note that the one-head machines stay neatly in their separate conveyor belts, so no machines crash into each other during this simulation. If a head steps out of the interval $[-m(t), m(t)]$ during the simulation, we can stop simulating it, as it will not reach the origin. After simulating heads for t steps, we know the value of $f^t(x)_0$. (Of course, we really only have to simulate a head if its conveyor belt contains the origin, and there is a unique such a head, but it does not hurt to simulate all of them.) \square

Theorem 3. *Let X be an uncountable sofic shift. Then there exists a distorted automorphism on X .*

Proof. Let \mathcal{M} be a distorted one-head machine. Then the cellular automaton f constructed in the previous lemma is distorted. By Lemma 2, we obtain the same cellular automaton on any uncountable sofic shift. \square

3.5 Undecidability of distortion

In this section, we show that distortion is undecidable.

Theorem 4. *It is undecidable, given a reversible one-head machine \mathcal{M} , whether \mathcal{M} is distorted.*

We can actually see from the proof (and from the reduction in [2]) that it is Π_1^0 -complete.

Proof. Every one-head machine lies in exactly one of the three cases of Theorem 1. We have a semialgorithm for the periodic case (by simply computing powers of \mathcal{M} and checking whether they are the identity map), and we have a semialgorithm for the case when \mathcal{M} has positive speed by the computability of speed, presented in [14, Theorem 2.7]. If we had a semialgorithm for detecting distortion, we would then be able to decide all three classes, contradicting the undecidability of periodicity, established in [2, Theorem 8]. \square

Corollary 1. *For every uncountable sofic subshift X , it is undecidable, given an automorphism Φ of X , whether Φ is distorted.*

Proof. Lemmas 2 and 3 are effective: if some reversible one-head machine is given, an automorphism of X can be built such that the radius growth differs at most by a multiplicative constant. If we could decide whether the corresponding automorphism is distorted, we could then decide whether the original machine is, which contradicts Theorem 4. \square

4 Unbalanced distortion in general subshifts

We give a general construction of a distorted automorphism. The distortion we aim for is ‘highly unbalanced’.

As mentioned above, it is desirable that $N(f)$ is generated by a single interval, that is, $I(f) \in N(f)$. Thus the size of the difference between $I(f)$ and the minimal intervals in $N(f)$ somehow measures the ‘badness’ of $N(f)$. We give an automorphism where this difference grows fast along iterations of f : The following theorem shows that we can have $d(f^t)$ be close to logarithmic, while $D(f^t)$ is close to linear.

We note that our construction is very similar to a construction in [8], though ours is not (at least consciously) based on it. Our proof is based on self-similar mud machinery that allows the construction of tracks that take a long time to walk over, but return to their original state once passed. To organize the behavior required in the theorem is then not difficult, though getting the numbers right requires some care because ϕ and ψ are arbitrary.

Theorem 5. *Let $\phi : \mathbb{N} \rightarrow \mathbb{N}$ be any sublinear function and let $\psi : \mathbb{N} \rightarrow \mathbb{N}$ be any nondecreasing superlogarithmic function. Then there exist a subshift X and an automorphism $f : X \rightarrow X$ such that there exist arbitrarily large t_i such that $d(f^{t_i}) \leq \psi(t_i)$ but $D(f^{t_i}) \geq \phi(t_i)$.*

By Proposition 1, the function ϕ cannot be made linear and ψ cannot be made logarithmic from a subshift X with subexponential complexity.

The following is shown in [1, Theorem 3.24]. Let us say that f is weakly periodic if there exist $p \geq 1$ and $j \in \mathbb{Z}$ such that $f^p = \sigma^j$.

Theorem 6. *Let X be an SFT and $f : X \rightarrow X$ an automorphism which is not weakly periodic. Then $I(f^t) \in N(f^t)$ for all large enough t .*

They ask [1, Question 3.26] whether the assumption that X is an SFT is needed. Theorem 5 answers by showing a general subshift and an automorphism of it for which at infinitely many t the interval $I(f^t)$ is arbitrarily close to logarithmic in size, but all contiguous neighborhoods are arbitrarily close to linear in size.

The proof of Theorem 5 first needs a technical but simple lemma. A function $\phi : \mathbb{N} \rightarrow \mathbb{R}$ is sublinear if $\phi(t) = o(t)$. Functions with this property can have weird local behavior, which complicates the argument. We show that all sublinear functions are majored by sublinear functions with some additional nice properties.

A nondecreasing function $\psi : \mathbb{N} \rightarrow \mathbb{R}$ has asymptotic slope zero if $|\psi(t+1) - \psi(t)|$ tends to zero as $t \rightarrow \infty$. Note that if for a function ϕ , we write $\partial\phi : \mathbb{N} \rightarrow \mathbb{R}$ for its discrete derivative $\partial\phi(t) = \phi(t+1) - \phi(t)$, then asymptotic slope zero, for an increasing function means just that $\partial\phi(t)$ tends to 0 as $t \rightarrow \infty$. If $\psi : \mathbb{R} \rightarrow \mathbb{R}$ is piecewise linear and it is linear on every interval $\llbracket i, i+1 \rrbracket$ where $i \in \mathbb{N}$, then the restriction $\psi : \mathbb{N} \rightarrow \mathbb{R}$ has asymptotic slope zero if and only if the slopes of the linear pieces of ψ tend to zero.

Lemma 4. *If $\phi : \mathbb{N} \rightarrow \mathbb{R}$ is sublinear, then there is a sublinear increasing piecewise linear function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ with asymptotic slope zero such that $\psi(t) \geq \phi(t)$ for all $t \in \mathbb{N}$.*

Proof. First, we may assume that ϕ is nondecreasing, by replacing $\phi(t)$ with $\max_{i=0}^t \phi(i)$, as the resulting function stays sublinear.

By sublinearity, for all $k \geq 1$ there exists t_k such that $t \geq t_k \Rightarrow \phi(t) \leq t/k$. Pick such $t_k \in \mathbb{N}$ for all $k \geq 1$, and observe that we can increase any of the t_k without changing their relevant property. Thus, we can assume the following further properties:

- t_k is increasing in k ,
- $t_{k+1}/k > t_k/(k-1)$ for $k \geq 2$.

Consider the sequence of points $(t_{k+1}, t_{k+1}/k) \in \mathbb{N} \times \mathbb{R}$. The second item makes sure that this sequence of points increases on the first axis, and the second makes sure the sequence also increases on the second axis. On $\llbracket t_2, \infty \llbracket$, define ψ as the piecewise linear function obtained by linearly interpolating values in the interval $\llbracket t_k, t_{k+1} \llbracket$ between $\llbracket t_k/(k-1), t_{k+1}/k \llbracket$.

The function ψ is now increasing $\llbracket t_2, \infty \llbracket$, because the point $(t_{k+1}, t_{k+1}/k)$ is strictly below point $(t_{k+2}, t_{k+2}/(k+1))$ for all $k \geq 1$.

Note that in the interval $\llbracket t_k, \infty \llbracket$ where $k \geq 2$, we have $\psi(t) \leq t/(k-1)$: Each of the points $(t_\ell, t_\ell/(\ell-1))$ for $\ell > k$ are strictly below the line $L_k = \{(x, x/(k-1)) \mid x \in \mathbb{N}\}$ because $x/(k-1) > x/(\ell-1)$. Thus, interpolating linearly between these points, we obtain a path that stays under L_k , and values of $\psi(t)$ are by definition on this path. It follows from this that ψ is sublinear.

Next, observe that $\psi(t+1) \leq \psi(t) + \frac{1}{k}$ whenever $t \geq t_{k+1}$. This is because the slope of the line between $(t_{\ell+1}, t_{\ell+1}/\ell)$ and $(t_{\ell+2}, t_{\ell+2}/(\ell+1))$ is

$$\frac{t_{\ell+2}/(\ell+1) - t_{\ell+1}/\ell}{t_{\ell+2} - t_{\ell+1}} \leq \frac{t_{\ell+2}/\ell - t_{\ell+1}/\ell}{t_{\ell+2} - t_{\ell+1}} = 1/\ell,$$

for all $\ell \geq 1$. Thus, increasing t by one can increase the value of $\psi(t)$ by at most $1/k$ whenever $t \geq t_{k+1}$, since any such t fits in one of the intervals $\llbracket t_\ell + 1, t_{\ell+1} \llbracket$ where $\ell \geq k$.

Finally, we show that $\psi(t) \geq \phi(t)$ for all but finitely many t , from which the claim follows by choosing the first few values of ψ suitably, and then increasing other values by a constant. Suppose then that $k \geq 2$. On the interval $t \in \llbracket t_k, t_{k+1} \llbracket$, $\psi(t)$ is linearly interpolated between $(t_k, t_k/(k-1))$ and $(t_{k+1}, t_{k+1}/k)$. In particular the line between $(t_k, t_k/k)$ and $(t_{k+1}, t_{k+1}/k)$ is strictly below the graph of ψ . But $t \in \llbracket t_k, t_{k+1} \llbracket \Rightarrow \phi(t) \leq t/k$, implying that $(t, \psi(t))$ is below the the point $(t, t/k)$, thus below the graph of ψ . \square

With a slightly more careful construction, ψ could be made to have strictly nonincreasing first difference function (that is, such that the slopes of the linear pieces decrease from piece to piece), though we do not need this.

A function $\psi : \mathbb{N} \rightarrow \mathbb{R}$ is 2-nice if for all $C \in \mathbb{N}$, $\psi(t+C) \leq 2\psi(t)$ for all but finitely many t .

Corollary 2. *If $\phi : \mathbb{N} \rightarrow \mathbb{N}$ is sublinear, then there is a sublinear nondecreasing 2-nice function $\psi : \mathbb{N} \rightarrow \mathbb{N}$ such that $\psi(t) \geq \phi(t)$ for all t .*

Proof. Seeing ϕ as a function $\phi : \mathbb{N} \rightarrow \mathbb{R}$, the previous lemma gives us a sublinear increasing $\psi : \mathbb{N} \rightarrow \mathbb{R}$ with asymptotic slope zero. It is easy to see (by separate easy proofs in the bounded and in the unbounded case) that any nondecreasing function $\psi : \mathbb{N} \rightarrow \mathbb{R}$ with asymptotic slope zero is 2-nice. If $\psi : \mathbb{N} \rightarrow \mathbb{R}$ is increasing and 2-nice, then $t \mapsto \lceil \psi(t) \rceil : \mathbb{N} \rightarrow \mathbb{N}$ is nondecreasing and 2-nice as well, and it clearly majors ψ , thus ϕ . \square

Proof (of Theorem 5). By the previous corollary, we may assume that ϕ is sublinear and 2-nice.

Take the alphabet $\{0, 1, 2, 3, 4, 5\} \times \{0, >, <\}$. The number $\{0, 1, 2, 3, 4, 5\}$ is the *mud state* and $<$ and $>$ are called *runners*. We construct a cellular automaton f that preserves the number of runners in every configuration with the property that on every f -invariant subshift, the map h flipping left and right runners is a time-symmetry for f , assuming that two runners never meet (that is, no word in $\{\ll, \langle \rangle, \rangle \langle, \gg\}$ appears in the configuration).

We only describe how the CA behaves when runners do not meet, as we will construct our subshift so that this does not happen. Our CA is composed of two CA, $f = g_2 \circ g_1$. The CA g_1 moves every occurrence of $<$ to the left and $>$ to the right. The CA g_2 maps

$$\begin{array}{ll} (0, >) \mapsto (2, >) & (1, <) \mapsto (5, <) \\ (1, >) \mapsto (3, <) & (0, <) \mapsto (4, >) \\ (2, <) \mapsto (0, >) & (5, >) \mapsto (1, <) \\ (3, >) \mapsto (1, >) & (4, <) \mapsto (0, <) \end{array}$$

and the local rule is filled arbitrarily so that this is a symbol permutation. It is a good idea to think of a left-going runner $<$ as already being on the left of the symbol, and $>$ as being on the right. We use a shorthand notation reflecting this⁴, and write $(a, 0)$ as simply a , $(a, >)$ as $a>$ and $(a, <)$ as $<a$. We also write $>w$ and $w<$ when the mud state of the cell the runner is on is not important.

Now, the idea is the following: We call a word $w \in \{0, 1\}^*$ a *track* if, when a runner $>$ enters it from one side, it eventually goes out from the other side, leaving w in whatever state it was originally in, with the additional property that the number of times the mud state of a cell in w turns from zero or 1 to a symbol in $\{2, 3, 4, 5\}$ is odd. The first property means that there exists t such that for all $a \in \{0, 1, 2, 3, 4, 5\}$ we have

$$f^t(>0^{|w|}, aw) = (0^{|w|}>, aw)$$

and

$$f^t(0^{|w|}<, wa) = (<0^{|w|}, wa).$$

⁴ We could just as well actually define our cellular automaton this way, but then tracks (defined below) move on the tape when a runner passes over them, which muddies up the global picture.

The importance of the second requirement will be clarified later. Note that on the first step, the runner is not yet on the support of the word, and on the last step, it is on the last symbol of the word w . To a track w , we associate its duration $t(w) \in \mathbb{N}$, which is the least t with the property above.

We show an example of a track. The word 01 is a track, and $t(01) = 4$, since (showing configurations from left to right) we have the evolutions

$$\begin{aligned} &>01; 2>1; 2<3; 0>3; 01> \\ &01<; 0<5; 4>5; 4<1; <01 \end{aligned}$$

where we see that the word returned to its original state, and both 0 and 1 changed their state an odd number of times. Clearly the composition of two tracks is a track, and it is easy to show by induction that if w is a track, then $0w1$ is a track as well, so we have a full Dyck language of tracks. One can check the formulas $t(uv) = t(u) + t(v)$ and $t(0w1) = 4 + 3t(w)$. For example

$$t(001011) = 4 + 3t(0101) = 4 + 3(t(01) + t(01)) = 28.$$

Just for fun, let us show how the head moves through 001011 , representing configurations top-down, then left to right. We have

$$\begin{array}{ccccc} >001011 & 2012>11 & 2014>53 & 2<01013 & 0012>13 \\ 2>01011 & 2012<31 & 2014<13 & 0>01013 & 0012<33 \\ 22>1011 & 2010>31 & 201<013 & 02>1013 & 0010>33 \\ 22<3011 & 20101>1 & 20<5013 & 02<3013 & 00101>3 \\ 20>3011 & 20101<3 & 24>5013 & 00>3013 & 001011> \\ 201>011 & 2010<53 & 24<1013 & 001>013 & \end{array}$$

and one can check that the vector recording the number of times each symbol 0 or 1 was changed is the all-odd vector $(1, 3, 3, 3, 3, 1)$. The right-to-left case is symmetric.

Now fix $w_0 = 01$. Suppose w_i has been defined and define $w_{i+1} \in \{0, 1\}^*$ as

$$w_{i+1} = (w_i 0)^{k_{i+1}} w_i (1 w_i)^{k_{i+1}}$$

where $k_{i+1} \in \mathbb{N}$. Then, writing $\ell_j = |w_j|$ for all $j \in \mathbb{N}$, we have

$$\ell_{i+1} = (2k_{i+1} + 1)\ell_i + 2k_{i+1} \leq 4k_{i+1}\ell_i$$

if k_{i+1} is large enough as a function of ℓ_i (note that $\ell_i \geq 2$). Then w_{i+1} is a track, and writing $t_j = t(w_j)$ for all $j \in \mathbb{N}$, we have

$$t_{i+1} = 3^{k_{i+1}}(2t_i + 2) - t_i - 2 \geq 3^{k_{i+1}}t_i$$

if k_{i+1} is large enough. (The exact formula is provided for completeness, but we only need $t_{i+1} \geq 3^{k_{i+1}}t_i$ which obviously follows by induction from $t(0w1) = 4 + 3t(w)$.)

We pick k_{i+1} so that $\psi(\frac{1}{2}t_{i+1}) > 2\ell_{i+1}$. This is possible because when k_{i+1} grows, ℓ_{i+1} grows at a linear rate, and t_{i+1} grows exponentially, while ψ is superlogarithmic. More precisely, since $t_i > 2$ we have $\log_3 \frac{1}{2}t_{i+1} \geq k_{i+1}$ and $\ell_{i+1} \leq 4k_{i+1}\ell_i$. By the assumption on ψ , if n is large enough, we have $\psi(n) > 8\ell_i \log_3 n$, so in particular if k_{i+1} is large enough, we have

$$\psi(\frac{1}{2}t_{i+1}) > 8\ell_i \log_3 \frac{1}{2}t_{i+1} > 8\ell_i k_{i+1} \geq 2\ell_{i+1}.$$

We have obtained that t_i grows very fast as a function of ℓ , as it must make the function ψ – which can be arbitrarily close to logarithmic – overtake ℓ . Reversing our point of view, we have achieved that ℓ_i grows ‘arbitrarily close to logarithmically’ in t_i .

It is easy to prove by induction that for every $j < i$, we have a decomposition

$$w_i = w_j b_1 w_j b_2 \cdots b_m w_j$$

for some m , where the b_k are individual bits $b_k \in \{0, 1\}$.

Now, we construct our subshift X , which we call the *mud run subshift*. For each $i \in \mathbb{N}$, pick $q_i \in \mathbb{N}$ and $k_i \in \mathbb{N}$ so that $\phi(q_i t_i) < q_i \ell_i$ (which is true for any large enough q_i since ϕ is sublinear), and additionally so that $\frac{1}{2}t_{k_i} < \phi(q_i t_i) \leq t_{k_i}$, using the fact that $\phi(n + t_i) \leq 2\phi(n)$ for all large enough n . (Note that, if we pick $k_i \geq i + 1$, the value of t_{k_i} is not determined by the values k_1, \dots, k_i , but rather values up to k_{k_i} , so it is easy to make sure that t_{k_i} is much larger than $\phi(q_i t_i)$, and we can then increase q_i to get $\phi(q_i t_i)$ in the desired interval.)

For each i , take the periodic points $x_i = (>w_i^{q_i})^{\mathbb{Z}}$. Then x_i is a temporally periodic point for f , and the length of its f -orbit is $q_i t_i$, while the length of its σ -orbit is $q_i \ell_i$. Let $\chi : \{0, 1, 2, 3, 4, 5\} \times \{0, <, >\} \rightarrow \{0, 1, 2, 3, 4, 5\} \times \{0\}$ be the map that removes runners, and define the subshift Y as the closure of

$$\{\sigma^a(f^b(x_i)), \sigma^a(f^b(\chi(x_i))) \mid a \in \mathbb{Z}, b \in \mathbb{Z}, i \in \mathbb{N}\}.$$

It is easy to see that f is still an automorphism of this limit subshift (since it has the same inverse), and that every point in Y that is not in the $\{\sigma, f\}$ -orbits of x_i has at most one runner (simply because $\ell_i \rightarrow \infty$).

Finally, define $X = Y \times \{0, 1\}^{\mathbb{Z}}$ and modify the behavior of f so that it behaves as before on Y , but additionally flips the bit on the second track whenever it turns a symbol from 0 or 1 to another symbol on the Y -component. Then by the assumption that we originally made for tracks that 0 and 1 are changed to another symbol an odd number of times, we have that for all $i \in \mathbb{N}$ and $z \in \{0, 1\}^{\mathbb{Z}}$ we have $f^{q_i t_i}(x_i, z) = (x_i, z')$ where $z'_j = 1 - z_j$ for all $j \in \mathbb{Z}$.

We will now prove that the evolution of the neighborhoods of f on X has the properties we claim. More precisely, we pick a suitable sequence of times, $n_i = q_i t_i \in \mathbb{N}$, at which we look at the neighborhoods $N(f^{n_i})$. We show that due to our choice of the q_i , every interval in $N(f^{n_i})$ is of size at least $\phi(n_i)$. The reason for this is that already on the periodic points x_i generating X , we need neighborhoods of this size, since the runners move at a linear speed for a long

time. We then show that simply due to the way the words w_i were constructed, we necessarily have

$$\llbracket -\psi(n_i), \infty \llbracket , \rrbracket -\infty, \psi(n_i) \rrbracket \in N(f^{n_i}),$$

for these n_i . This is because, knowing the infinite tails, if we see no runners in those tails, we must be in a limit point, and in such points there is at most one runner, running at speed asymptotically slower than ψ . We do not have to look far to find it.

More precisely, for the first claim suppose $\llbracket a, b \llbracket \in N(f^{n_i})$. Then if $b - a + 1 < q_i \ell_i$, we see runners in $\llbracket a, b \llbracket$ in neither of the points $\sigma^{-a+1}(x_i)$ and $\sigma^{-a+1}(\chi(x_i))$. However, $f^{n_i}(\sigma^{-a+1}(x_i))_0 \neq \sigma^{-a+1}(x_i)_0$ and $f^{n_i}(\sigma^{-a+1}(\chi(x_i)))_0 \neq \sigma^{-a+1}(\chi(x_i))_0$. This contradicts $\llbracket a, b \llbracket \in N(f^{n_i})$. Since $\phi(n_i) = \phi(q_i t_i) < q_i \ell_i$, in particular every interval in $N(f^{n_i})$ is of length at least $\phi(n_i)$.

To see that $\llbracket -\psi(n_i), \infty \llbracket \in N(f^{n_i})$, note that if there are at least two runners in the $\llbracket -\psi(n_i), \infty \llbracket$ -tail of a point $x \in X$, then runners appear periodically in x , and in fact we know precisely what will be at the origin after n_i steps. It remains to consider the limit set. It is enough to show that if x is in the limit set, and the unique runner on x is in cell 0, then after n_i steps it has moved by at most $\psi(n_i)$ steps (in either direction). Namely, if this is the case, then knowing the contents of $\llbracket -\psi(n_i), \psi(n_i) \rrbracket$ allows us to determine the contents of the cell 0 after the application of f^{n_i} by simply simulating the movement of the runner as long as it does not exit $\llbracket -\psi(n_i), \psi(n_i) \rrbracket$ (and if it does, the contents of cell 0 will no longer change).

To see this, consider the movement of a runner on one of the periodic points x_k . We claim that if $j < k$, then if a runner is in cell a of $f^n(x_k)$, then in $f^{n+t_j}(x_k)$, it will be in some cell in the interval $\llbracket a - 2\ell_j, a + 2\ell_j \rrbracket$. To see this, recall that x_k is a concatenation of the words w_j separated by individual bits. Since tracks return to their original state after the runner has passed through them, in $f^n(x_k)$ a full intact copy of the track w_j appears in both of the intervals $\llbracket a, a + 2\ell_j \rrbracket$ and $\llbracket a - 2\ell_j, a \rrbracket$. It follows that exiting the interval $\llbracket a - 2\ell_j, a + 2\ell_j \rrbracket$ requires passing through at least one complete track w_j , which takes time t_j .

Since f is continuous, the same will be true in the limit: every runner in a limit point x moves at most $2\ell_j$ cells in t_j time steps. Remember that we have $\frac{1}{2}t_{k_i} < n_i = q_i t_i \leq t_{k_i}$. In t_{k_i} time steps, the head moves at most $2\ell_{k_i}$ steps, so since $n_i = q_i t_i \leq t_{k_i}$, in n_i time steps we move a distance of at most

$$2\ell_{k_i} < \psi\left(\frac{t_{i+1}}{2}\right) \leq \psi(n_i)$$

cells, as required, where the last inequality follows because ψ is nondecreasing. This concludes the proof. \square

5 Future work

As mentioned in the section on one-head machines, we do not know much about the actual speeds of distorted machines. We tend to believe that the movement

of the known examples is closer to $\log t$ than $t/\log t$, but have no rigorous proofs for any examples. It should be routine to compute the movement bound for the SMART machine, given that its trace subshift is known, and this would already clarify the situation quite a bit.

The group generated by reversible one-head machines is of prime interest and is the purpose of [21], though defined in a more natural way, which emphasizes its relation to the topological full group. In group theory, an element $g \in G$ is called *distorted* if powers of g grow sublinearly in the word norm of some finitely generated subgroup. We do not give examples of cellular automata or one-head machines that are distorted in this sense, but we do believe constructing them is possible.

Conjecture 1. The group generated by reversible one-head machines contains a distortion element.

More concretely, we find it plausible that the SMART machine is distorted. In the journal version of [21], which is in preparation, it will be shown that the subgroup of elementary one-head machines is finitely generated. Our conjecture is based on the fact that it seems plausible that a word expressing SMART in terms of these generators will not be linear, but doing this construction explicitly would presumably be a lot of work. If the conjecture is true, it follows (see Section 3.4) that automorphism groups of full shifts also have distortion elements.

References

1. Cyr, V., Franks, J., Kra, B.: The spacetime of a shift automorphism. ArXiv e-prints (October 2016)
2. Kari, J., Ollinger, N.: Periodicity and immortality in reversible computing. In: Proceedings of the 33rd international symposium on Mathematical Foundations of Computer Science. MFCS '08, Berlin, Heidelberg, Springer-Verlag (2008) 419–430
3. Cassaigne, J., Ollinger, N., Torres-Avilés, R.: A small minimal aperiodic reversible Turing machine. *Journal of Computer and System Sciences* **84** (2017) 288 – 301
4. Hedlund, G.A.: Endomorphisms and automorphisms of the shift dynamical system. *Mathematical Systems Theory* **3** (1969) 320–375
5. Cyr, V., Franks, J., Kra, B., Petite, S.: Distortion and the automorphism group of a shift. ArXiv e-prints (November 2016)
6. Tisseur, P.: Cellular automata and Lyapunov exponents. *Nonlinearity* **13**(5) (2000) 1547–1560
7. Boyle, M., Lind, D.: Expansive subdynamics. *Transactions of the AMS* **349**(1) (1997) 55–102
8. Hochman, M.: Non-expansive directions for \mathbb{Z}^2 actions. *Ergodic Theory and Dynamical Systems* **31** (1 2011) 91–112
9. Zinoviadis, C.: Hierarchy and Expansiveness in 2D Subshifts of Finite Type. PhD thesis, Turun Yliopisto, Turku (2015)
10. Guillon, P., Zinoviadis, C.: Hierarchy and expansiveness in two-dimensional subshifts of finite type. draft (March 2016)
11. Morita, K.: Universality of reversible two-counter machine. *Theoret. Comput. Sci.* **168**(2) (1996) 303–320
12. Kůrka, P.: On topological dynamics of Turing machines. *Theoretical Computer Science* **174**(1) (March 1997) 203–216
13. Gajardo, A., Guillon, P.: Zigzags in Turing machines. In Ablayev, F., Mayr, E.W., eds.: 5th International Computer Science Symposium in Russia (CSR'10). Volume 6072 of LNCS., Kazan, Russia, Springer Heidelberg (June 2010) 109–119
14. Jeandel, E.: Computability of the entropy of one-tape Turing machines. ArXiv e-prints (February 2013)
15. Hennie, F.: One-tape, off-line Turing machine computations. *Information and Control* **8** (1965) 553–578
16. Salo, V.: Subshifts with sparse projective subdynamics. ArXiv e-prints (May 2016)
17. Blondel, V.D., Cassaigne, J., Nichitiu, C.: On the presence of periodic configurations in Turing machines and in counter machines. *Theoret. Comput. Sci.* **289** (2002) 573–590
18. Adamczewski, B.: Balances for fixed points of primitive substitutions. *Theoretical Computer Science* **307**(1) (2003) 47 – 75
19. Salo, V.: A note on subgroups of automorphism groups of full shifts. *Ergodic Theory and Dynamical Systems* (2016) 1–13
20. Kim, K.H., Roush, F.W.: On the automorphism groups of subshifts. *Pure Mathematics and Applications* **1**(4) (1990) 203–230
21. Barbieri, S., Kari, J., Salo, V.: The group of reversible Turing machines. In Cook, M., Neary, T., eds.: Automata, Zurich, Switzerland, IFIP, Springer International Publishing (June 2016) 49–62