Discrete Mathematics and Theoretical Computer Science DMTCS vol. **15:3**, 2013, 41–60

Surjective cellular automata far from the Garden of Eden

Silvio Capobianco[1]† Pierre Guillon[2,3]‡ Jarkko Kari[3]§

[1] *Institute of Cybernetics at Tallinn University of Technology, Estonia*
[2] *CNRS & IML, Marseille, France*
[3] *Mathematics Department, University of Turku, Finland*

received 7th Dec. 2012, *revised 30th Sep. 2013*, *accepted 10th Oct. 2013*.

One of the first and most famous results of cellular automata theory, Moore's Garden-of-Eden theorem has been proven to hold if and only if the underlying group possesses the measure-theoretic properties suggested by von Neumann to be the obstacle to the Banach-Tarski paradox. We show that several other results from the literature, already known to characterize surjective cellular automata in dimension d, hold precisely when the Garden-of-Eden theorem does. We focus in particular on the balancedness theorem, which has been proven by Bartholdi to fail on amenable groups, and we measure the amount of such failure.

Mathematics Subject Classification 2000: 37B15, 68Q80.

Keywords: cellular automata, amenability, group theory, topological dynamics, symbolic dynamics, algorithmic randomness.

1 Introduction

Cellular automata (CA) are local descriptions of global dynamics. Given an underlying uniform graph (*e.g.*, the square grid on the plane) a CA is defined by a finite alphabet, a finite neighborhood for the nodes of the graph, and a local function that maps states of a neighborhood into states of a point. By synchronous application of the local function at all nodes, a global function on configurations is defined.

The study of global properties of CA and their relations with the local description has been a main topic of research since the field was established. Indeed, the Garden-of-Eden theorem by Moore [19] and its converse by Myhill [20], which link surjectivity of the global map of 2D CA to *pre-injectivity* (a property that may be described as the impossibility of erasing finitely many errors in finite time) also have the distinction of being the first rigorous results of cellular automata theory. Since then, several more properties were later proven to be equivalent to surjectivity for d-dimensional CA. Among them are:

†Email: `silvio@cs.ioc.ee`
‡Email: `pierre.guillon@math.cnrs.fr`
§Email: `jkari@utu.fi`

- *Balancedness* [18]: each pattern of a given shape has the same number of preimages.
- Preservation of *Martin-Löf randomness* [3]: the image of any algorithmically incompressible configuration is itself algorithmically incompressible.

With the subsequent efforts to extend the definition of CA to the more general situation of Cayley graphs of finitely generated groups, an unexpected phenomenon appeared: the Garden-of-Eden theorem actually depends on properties of the involved groups. This phenomenon dates back to Machì and Mignosi's 1993 paper [15], where counterexamples to both Moore's and Myhill's theorems on the free group on two generators are presented, but the theorems themselves are proven for groups of *subexponential growth*, a class which includes the Euclidean groups. Comparing the original papers [19] and [20], a key fact emerges, which is crucial for the proofs: in $\mathbb{Z}^d$, the size of a hypercube is a d-th power of the side, but the number of sites on its outer surface is a polynomial of degree $d-1$. In other words, it seems that, to get Moore's or Myhill's theorems for CA on a group G, we need that in G *the sphere grows more slowly than the ball*.

What is actually sufficient for the Garden-of-Eden theorem to hold is a slightly weaker property called *amenability*, which was formulated by von Neumann in an attempt to explain the *Banach-Tarski paradox*: the unit ball in the space can be decomposed into finitely many parts, and those parts reassembled so to form *two* unit balls! Informally, a group is amenable if, however given a *finite shape* for the sphere, it is always possible to find a *finite ball* whose sphere is *proportionally as small as wished*: it turns out that the Hausdorff phenomenon takes place in the space because the group of rotations of the space has a free subgroup on two generators, which precludes amenability. Ceccherini-Silberstein *et al.* [7] proved then that Moore's theorem holds for CA on any amenable group, but fails for groups that have a free subgroup on two generators. After about a decade, Bartholdi [1] completed the proof for every non-amenable group, and added preservation of the *uniform product measure* to the list of properties verified by surjective CA on all and only the amenable groups. This can also be related to characterizations by Ornstein and Weiss [22] of groups whose full shifts over distinct alphabets factor onto one another.

In this paper, we extend the range of Bartholdi's theorem by characterizing amenable groups as those where surjective CA have additional properties. We start by considering balancedness, which is the combinatorial variant of preservation the product measure: thus, amenable groups are precisely those where surjective CA are balanced. We then include several properties studied in topological dynamics: CA with any of these properties are surjective, and we show that the converse implications holds precisely for CA on amenable groups.

Theorem 1 *Let G be a group. The following are equivalent.*

1. *G is amenable.*
2. *Every surjective* CA *on G is pre-injective.*
3. *Every surjective* CA *on G preserves the uniform product measure.*
4. *Every surjective* CA *on G is balanced.*
5. *Every surjective* CA *on G is recurrent for the uniform product measure.*
6. *Every surjective* CA *on G is nonwandering.*

We then show a fact which is remarkable by its own right. Not only preservation of the uniform product measure by surjective CA characterizes amenable groups: it also fails *catastrophically* for non-amenable ones, in the sense given by the following statement.

Theorem 2 *Let G be a non-amenable group. There exist an alphabet Q, a subset U of Q^G such that $\mu_\Pi(U) = 1$, and a surjective cellular automaton $\mathcal{A}$ over G with alphabet Q such that $\mu_\Pi(F_{\mathcal{A}}^{-1}(U)) = 0$, where $F_{\mathcal{A}}$ is the global function of $\mathcal{A}$ and μ_Π the uniform product measure on Q^G.*

To prove Theorem 2, we introduce a definition of *normality* for configurations, which is modeled on the one for infinite words over a finite alphabet. Such a trick has been successfully applied before on the Euclidean groups $\mathbb{Z}^d$: however, in our more general setting, several properties do not hold, which forces us to add further conditions to ensure that the set of normal configurations (the set U in Theorem 2) has full measure. In turn, the cellular automaton $\mathcal{A}$ will be a variant of Bartholdi's counterexample, modified so that it has a spreading state.

Finally, for finitely generated groups with decidable word problem, Martin-Löf randomness can be defined: such definition depends on the measure defined on the Borel σ-algebra, which for our aims will be the product measure. Under these additional hypotheses, we show that the result by Calude *et al.* [3] about surjective CA preserving Martin-Löf randomness, holds precisely for amenable groups.

Theorem 3 *Let G be a finitely generated group with decidable word problem. Then G is amenable if and only if for every surjective* CA *$\mathcal{A}$ on G, whenever a configuration c is Martin-Löf random with respect to the product measure μ_Π, so is its image $F_{\mathcal{A}}(c)$.*

In addition, if G is not amenable, there exists a surjective CA *on G such that every Martin-Löf random configuration w.r.t. μ_Π has a nonrandom image and only nonrandom preimages. In particular, the set U in Theorem 2 can be taken as the set of Martin-Löf random configurations w.r.t. μ_Π.*

The paper is organized as follows. Section 2 provides a background. Section 3 deals with balancedness, and Section 4 with the nonwandering property. Section 5 is devoted to the proof of Theorem 2, and Section 6 to that of Theorem 3.

2 Background

Given a set X, we denote by $\mathcal{PF}(X)$ the set of all finite subsets of X.

2.1 Groups

Let G be a group. We call 1_G, or simply 1, its identity element. Given a set X, the family $\sigma = \{\sigma_g\}_{g\in G}$ of transformations of $X^G = \{c : G \to X\}$, called *translations*, defined by

$$\sigma_g(c)(z) = c^g(z) = c(gz) \ \ \forall z \in G \ \forall g \in G \tag{1}$$

is a *right action* of G on X^G, that is, $\sigma_{gh} = \sigma_h \circ \sigma_g$ for every $g, h \in G$. This is consistent with defining the product $\phi\psi$ of functions as the composition $\psi \circ \phi$. Other authors (cf. [6]) define $\sigma_g(c)(x)$ as $c(g^{-1}x)$, so that σ becomes a *left* action. However, most of the definitions and properties we deal with do not depend on the "side" of the multiplication: we will therefore stick to (1).

A set of *generators* for G is a subset $S \subseteq G$ such that for each $g \in G$ there is a word $w = w_1 \dots w_n$ on $S \cup S^{-1}$ such that $g = w_1 \cdots w_n$. The minimum length of such a word is called the *length* of g w.r.t. S, and indicated by $\|g\|_S$, or simply $\|g\|$. G is *finitely generated* (briefly, f.g.) if S can be chosen finite. A group G is *free* on a set S if it is isomorphic to the group of *reduced words* on $S \cup S^{-1}$: a word w is said to be reduced if for every $s \in S$ the pairs ss^{-1} and $s^{-1}s$ do not appear in w. For $r \geq 0$, $g \in G$ the *disk of radius* r *centered in* g is $D_r(g) = \{h \in G \mid \|g^{-1}h\| \leq r\}$. The points of $D_r(g)$ can be "reached" from the "origin" 1_G by first "walking" up to g, then making up to r steps: this is consistent with the definition of translations by (1), where to determine $c^g(z)$ we first move from 1 to g, then from g to gz. We write D_r for $D_r(1)$. We also put $U^{-r} = \{z \in G \mid D_r(z) \subseteq U\}$ and $\partial_{-r}U = U \setminus U^{-r}$.

A group G is *residually finite* (briefly, r.f.) if for every $g \neq 1$ there exists a homomorphism $\phi : G \to H$ such that H is finite and $\phi(g) \neq 1$. Equivalently, G is r.f. if the intersection of all its subgroups of finite index is trivial. It follows from the definitions that, if G is r.f. and $U \subseteq G$ is finite, then there exists $H \leq G$ such that $[G : H] < \infty$ and $U \cap H \subseteq \{1_G\}$.

Lemma 4 ([10, Lemma 2.3.2]) *Let G be a residually finite group and F a finite subset of G not containing 1_G. There exists a subgroup H of finite index in G, which does not intersect F, and such that the right cosets Hu, $u \in F$, are pairwise disjoint.*

The *stabilizer* of $c \in X^G$ is the subgroup $\mathrm{st}(c) = \{g \in G \mid c^g = c\}$: be aware, that $\mathrm{st}(c)$ might not be a normal subgroup. The configuration c is *periodic* if $[G : \mathrm{st}(c)] < \infty$; if $[G : H] < \infty$ and $H \leq \mathrm{st}(c)$ we say that c is H-*periodic*. The family of periodic configurations in X^G is indicated by $\mathrm{Per}(G, X)$.

A group G is *amenable* if it satisfies the following equivalent conditions:

1. There exists a *finitely* additive probability measure μ on G such that $\mu(gA) = \mu(A)$ for every $g \in G$, $A \subseteq G$.

2. For every $U \in \mathcal{PF}(G)$ and $\varepsilon > 0$ there exists $K \in \mathcal{PF}(G)$ such that
$$|UK \setminus K| < \varepsilon |K| \,. \tag{2}$$

3. There exists a net $\{X_i\}_{i \in I}$ of finite nonempty subsets of G such that, for every $U \in \mathcal{PF}(G)$,
$$\lim_{i \in I} \frac{|UX_i \setminus X_i|}{|X_i|} = 0 \,. \tag{3}$$

Similar definitions want μ *right*-invariant and (2) replaced by $|KU \setminus K| < \varepsilon|K|$—and similarly for (3)—or μ both left- and right-invariant and set-theoretic differences in (2) and (3) replaced by *symmetric difference* (recall that $A \triangle B = (A \setminus B) \cup (B \setminus A)$): in fact, all these definitions are equivalent. Also, a group is amenable if and only if all of its finitely generated subgroups are amenable.

Example 5 *$\mathbb{Z}^d$ is amenable: $\{X_i\}_{i \geq 1}$ with $X_i = [0, \ldots, i-1]^d$ satisfies (3) for every $U \in \mathcal{PF}(\mathbb{Z}^d)$.*

A *(left) paradoxical decomposition* of a group G is a finite set of pairs $\{(\alpha_1, A_1), \ldots, (\alpha_n, A_n)\} \subseteq G \times 2^G$ with a separator $k < n$ such that $G = \bigsqcup_{i=1}^n A_i = \bigsqcup_{i=1}^k \alpha_i A_i = \bigsqcup_{i=k+1}^n \alpha_i A_i$ (the symbol $\bigsqcup$ meaning that the union is disjoint). A group has a left paradoxical decomposition if and only if it has a *right* paradoxical decomposition, satisfying $G = \bigsqcup_{i=1}^n A_i = \bigsqcup_{i=1}^k A_i \alpha_i = \bigsqcup_{i=k+1}^n A_i \alpha_i$ instead. A group is *paradoxical* if it has a paradoxical decomposition.

Proposition 6 (Tarski alternative; cf. [6, Theorem 4.9.2]) *A group is paradoxical if and only if it is not amenable.*

A *bounded-propagation two-to-one compressing map* over a group G is a map $\phi : G \to G$ such that, for some finite *propagation set* $S \subseteq G$, $\phi(g)^{-1}g \in S$ and $|\phi^{-1}(g)| = 2$ for every $g \in G$. In particular, such a map must be surjective, and $|S| \geq 2$. By [6, Theorem 4.9.2], a group has a bounded-propagation two-to-one compressing map if and only if it is paradoxical.

Example 7 *Let $G = \mathbb{F}_2$ be the free group on two generators a, b; for $g \in G$ let $w = w(g) = w_1 w_2 \cdots w_m$ be the unique reduced word on $\{a, b, a^{-1}, b^{-1}\}$ that represents g. Define:*

- $A_1 = \{g \in G \mid w_m = a^{-1}\} \cup \{a^n \mid n \geq 0\}$,
- $A_2 = \{g \in G \mid w_m = a\} \setminus \{a^n \mid n \geq 0\}$,
- $A_3 = \{g \in G \mid w_m = b^{-1}\}$, *and*
- $A_4 = \{g \in G \mid w_m = b\}$,

so that $G = A_1 \sqcup A_2 \sqcup A_3 \sqcup A_4 = A_1 \sqcup A_2 a^{-1} = A_3 \sqcup A_4 b^{-1}$. For $g \in G$ put $\phi(g) = g$ if $g \in A_1$, $\phi(ga) = g$ if $g \in A_2 a^{-1}$, $\phi(g) = g$ if $g \in A_3$, $\phi(gb) = g$ if $g \in A_4 b^{-1}$. Then ϕ is a bounded-propagation two-to-one compressing map with $S = \{1, a, b\}$.

2.2 Cellular automata

A *cellular automaton* (briefly, CA) on a group G is a triple $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ where the *alphabet* Q is a finite set, the *neighborhood* $\mathcal{N} \subseteq G$ is finite and nonempty, and $f : Q^{\mathcal{N}} \to Q$ is a *local function*. This, in turn, induces a *global function* on the space Q^G of *configurations*, defined by

$$F_{\mathcal{A}}(c)(g) = f\left(c^g|_{\mathcal{N}}\right) = f\left(c|_{g\mathcal{N}}\right) . \tag{4}$$

Hedlund's theorem [6, Theorem 1.8.1] states that global functions of CA are exactly those functions from Q^G to itself that commute with translations and are continuous in the *prodiscrete topology*, *i.e.*, the product topology where Q is considered as a discrete space. A base for this topology is given by the *cylinders* of the form $C(E, p) = \{c : G \to Q \mid c|_E = p\}$, with E a finite subset of G and $p : E \to Q$ a *pattern*: observe that, for countable groups, this base is countable. Also, the *elementary cylinders* $C(g, q) = \{c : G \to Q \mid c(g) = q\}$ with $g \in G$ and $q \in Q$ form a subbase. The set E is called the *shape* of the pattern p. Through (4) we also consider, for every finite $E \subseteq G$, a function $f : Q^{E\mathcal{N}} \to Q^E$ between patterns, defined by $f(p)(j) = f(p|_{j\mathcal{N}})$.

As a consequence of Hedlund's theorem, CA behave well with respect to periodic configurations.

Lemma 8 *If $F : Q^G \to Q^G$ commutes with translations, then $\mathrm{st}(c) \subseteq \mathrm{st}(F(c))$ for every $c \in Q^G$. In particular, if F is bijective then $\mathrm{st}(c) = \mathrm{st}(F(c))$.*

An *occurrence* of a pattern $p : E \to Q$ in a configuration $c : G \to Q$ is an element $g \in G$ such that $c^g|_E = p$; in other words, the pattern $p_g : gE \to Q$ defined by $p_g(gz) = p(z)$ is a *copy* of p. We indicate as $\mathrm{occ}(p, c)$ the set of occurrences of the pattern p in the configuration c.

Let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ be a CA on the group G. A configuration $c : G \to Q$ is a *Garden-of-Eden* (briefly, GoE) for $\mathcal{A}$ if it has no predecessor according to $\mathcal{A}$, *i.e.*, if $c \notin F_{\mathcal{A}}(Q^G)$. A pattern $p : E \to Q$ is an *orphan* for $\mathcal{A}$ if there is no $p' : E\mathcal{N} \to Q$ such that $f(p') = p$.

The following can be seen as an example of folkloric consequence of the compactness of the configuration space.

Proposition 9 (Orphan pattern principle; cf. [6, Proposition 5.1.1]) *A cellular automaton with finite alphabet has a* GoE *configuration if and only if it has an orphan pattern.*

A configuration is *rich* (or *shift-transitive*) if it contains occurrences of every pattern. The orphan pattern principle can then be restated as follows: a CA is surjective if and only if it sends rich configurations into rich configurations.

Two configurations are *asymptotic* if they differ on at most finitely many points; a CA is *pre-injective* if distinct asymptotic configurations have distinct images.

Proposition 10 (Moore's Garden-of-Eden theorem [19]) *Let $\mathcal{A}$ be a bidimensional* CA*. If $\mathcal{A}$ is surjective then $\mathcal{A}$ is pre-injective.*

Proposition 11 (Myhill [20]) *Let $\mathcal{A}$ be a bidimensional* CA*. If $\mathcal{A}$ is pre-injective then $\mathcal{A}$ is surjective.*

Proposition 12 (Ceccherini-Silberstein, Machì and Scarabotti [7]) *Let G be an amenable group and let $\mathcal{A}$ be a* CA *on G. Then $\mathcal{A}$ is surjective if and only if it is pre-injective.*

Let $\mathcal{N} \in \mathcal{PF}(G)$, $G \leq \Gamma$, and $f : Q^{\mathcal{N}} \to Q$. The triple $\langle Q, \mathcal{N}, f \rangle$ describes both a CA $\mathcal{A}$ over G and a CA $\mathcal{A}'$ on Γ. We then say that $\mathcal{A}'$ is the CA *induced* by $\mathcal{A}$ on Γ, or that $\mathcal{A}$ is the *restriction* of $\mathcal{A}'$ to G.

Proposition 13 (cf. [5, Theorem 1.2]) *Let $\mathcal{N} \in \mathcal{PF}(G)$, $G \leq \Gamma$; let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ be a* CA *on G and let $\mathcal{A}'$ be the* CA *induced by $\mathcal{A}$ on Γ.*

1. *$\mathcal{A}$ is surjective if and only if $\mathcal{A}'$ is surjective.*
2. *$\mathcal{A}$ is pre-injective if and only if $\mathcal{A}'$ is pre-injective.*
3. *$\mathcal{A}$ is injective if and only if $\mathcal{A}'$ is injective.*

2.3 Measures

Let Σ be a σ-algebra on Q^G. If $\mu : \Sigma \to [0,1]$ is a measure on Q^G, a measurable function $F : Q^G \to Q^G$ determines a new measure $F\mu : \Sigma \to [0,1]$ defined as $F\mu(U) = \mu(F^{-1}(U))$. We say that $U \in \Sigma$ is *μ-null* if $\mu(U) = 0$; we say that μ-almost every point satisfies a property P if the set of the points which do not satisfy P is μ-null. We say that F *preserves* μ if $F\mu = \mu$. If Σ_C is the σ-algebra generated by the cylinders, by the *Carathéodory extension theorem* and the *Hahn-Kolmogorov theorem* a probability measure on Σ_C is completely determined by its value on the cylinders. The measure $\mu_\Pi : \Sigma_C \to [0,1]$ defined by $\mu_\Pi(C(E,p)) = |Q|^{-|E|}$ is called the *uniform product measure*. Observe that Σ_C coincides with the *Borel σ-algebra* generated by the open sets if and only if G is countable. Also observe that CA global functions are both Borel measurable and Σ_C-measurable.

Proposition 14 (Bartholdi's theorem [1]) *Let G be a group. The following are equivalent.*

1. *G is amenable.*
2. *Every surjective* CA *on G is pre-injective.*
3. *Every surjective* CA *on G preserves the uniform product measure μ_Π.*

Let μ be a probability measure over Q^G. We say that $F : Q^G \to Q^G$ is *μ-recurrent* if for every measurable set $A \subset Q^G$ of measure $\mu(A) > 0$ there exists $t \geq 1$ such that $\mu(A \cap F^t(A)) > 0$.

Proposition 15 (Poincaré recurrence theorem; cf. [12, Theorem 4.1.19]) *Let (X, Σ, μ) be a probability space and let $F : X \to X$ be a measurable function. If F preserves μ, then F is μ-recurrent.*

3 Balancedness

Let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ a CA on $\mathbb{Z}^d$ such that $\mathcal{N} = \{-r, \dots, r\}^d$. According to Maruoka and Kimura [18], $\mathcal{A}$ is n-*balanced* if each pattern on a hypercube of side n has $|Q|^{(n+2r)^d - n^d}$ pre-images. The authors then prove that such $\mathcal{A}$ is surjective if and only if it is n-balanced for *every* n. On the other hand, the *majority* CA on $\{0,1\}^{\mathbb{Z}}$ such that $f(c(-1), c(0), c(1)) = 0$ if and only if at most one of the arguments is 1, is 1-balanced but not 2-balanced, as 0011 and 0101 are easily checked to be the only two preimages of 01: also, it has the Garden-of-Eden pattern 01001. Such GOE is also of minimal length: for example, 0100 has the preimage 010100.

The balancedness condition is the same as saying that each pattern on a *given shape* has the *same number* of pre-images: to see how, just "patch" arbitrary shapes to "fill" a hypercube. This allows to extend the definition to CA over arbitrary groups.

Definition 16 *Let G be a group and let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ be a* CA *on G. $\mathcal{A}$ is* balanced *if for every finite nonempty $E \subseteq G$, every pattern $p : E \to Q$ has the same number of preimages:*

$$|f^{-1}(p)| = |Q|^{|E\mathcal{N}| - |E|}. \tag{5}$$

The neighborhood $\mathcal{N}$ seems to have a crucial role in Definition 16: which may make the reader suspect it to be ill-posed. However, as we will see in a moment, what looks like a property of the *presentation*, is actually a property of the *dynamics*: balancedness of $\mathcal{A}$ only depends on its global function $F_{\mathcal{A}}$, not on the choice of the neighborhood $\mathcal{N}$ or the local function f—provided $F_{\mathcal{A}}$ remains the same.

Proposition 17 *A cellular automaton is balanced if and only if it preserves the uniform product measure.*

Proof: The argument is similar to the one used in [3] for $G = \mathbb{Z}^d$. Let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ and $p : E \to Q$: then

$$\mu_\Pi \left(F_{\mathcal{A}}^{-1}(C(E,p)) \right) = \sum_{f(p')=p} \mu_\Pi \left(C(E\mathcal{N}, p') \right) = \sum_{f(p')=p} |Q|^{-|E\mathcal{N}|} .$$

As $\mathcal{A}$ is balanced if and only if the right-hand side has $|Q|^{|E\mathcal{N}|-|E|}$ summands whatever p is, and preserves μ_Π if and only if the left-hand side equals $|Q|^{-|E|}$ whatever p is, the thesis follows. □

Remark 18 *Let $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ be a* CA *on $G \leq \Gamma$ and $\mathcal{A}'$ the* CA *induced by $\mathcal{A}$ on Γ. Then $\mathcal{A}$ is balanced if and only if $\mathcal{A}'$ is balanced.*

Since the r.h.s. in (5) is always positive, no pattern is an orphan for a balanced CA. In [7], two CA on the free group on two generators are shown, one being surjective but not pre-injective, the other pre-injective but not surjective: both have an unbalanced local function. Therefore, balancedness in general groups is *strictly stronger* than surjectivity, and possibly uncorrelated with pre-injectivity.

Balancedness allows us to generalize [3, Point 1 of Theorem 4.4] to finitely generated amenable groups.

Lemma 19 (Step 1 in proof of [7, Theorem 3]) *Let G be a finitely generated amenable group, $q \geq 2$, and $n > r > 0$. For $L = D_n$ there exist $m > 0$ and $B \subseteq G$ such that B contains m disjoint copies of L and*

$$(q^{|L|} - 1)^m \cdot q^{|B| - m|L|} < q^{|B^{-r}|} . \tag{6}$$

Proposition 20 *Let G be a finitely generated amenable group and let $\mathcal{A} = \langle Q, D_r, f\rangle$, $r > 0$, be a* CA *on G. If c is not rich then $F_{\mathcal{A}}(c)$ is not rich.*

Proof: Suppose there is a pattern with support $L = D_n$, $n > r$, that does not occur in c. Choose m and B according to Lemma 19. By hypothesis, the number of patterns with support B which occur in c is at most $(q^{|L|} - 1)^m q^{|B|-m|L|}$, with $q = |Q|$; therefore, the number of patterns with support $B \setminus \partial_r B$ which occur in $F_{\mathcal{A}}(c)$ cannot exceed this number too. By Lemma 19, this is strictly less than $q^{|B|-|\partial_r B|}$, which is the total number of patterns with support $B \setminus \partial_r B$: hence, some of those patterns do not occur in $F_{\mathcal{A}}(c)$. □

The last statement of this section is a strengthening of a result by Lawton [14] (also stated in [24, Theorem 1.3]) which states that injective CA on residually finite groups are surjective.

Theorem 21 *Let G be a residually finite group and $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ an injective* CA *over G. Then $\mathcal{A}$ is balanced.*

Proof: Let E be a finite subset of G: it is not restrictive to suppose $1 \in E \cap \mathcal{N}$, so that $E, \mathcal{N} \subseteq E\mathcal{N}$. Suppose, for the sake of contradiction, that $p : E \to Q$ satisfies $|F_{\mathcal{A}}^{-1}(p)| = M > |Q|^{|E\mathcal{N}|-|E|}$. Since G is residually finite, by Lemma 4 there exists a subgroup $H \leq G$ of finite index such that $H \cap E\mathcal{N} = H \cap \mathcal{N} = \{1\}$: if J is a set of representatives of the right cosets of H such that $E\mathcal{N} \subseteq J$, then

$$|\{\pi : J \to Q \mid F_{\mathcal{A}}(\pi)|_E = p\}| = M \cdot |Q|^{[G:H]-|E\mathcal{N}|} > |Q|^{[G:H]-|E|} . \tag{7}$$

The r.h.s. in (7) is the number of H-periodic configurations that coincide with p on E. Since $\mathcal{A}$ is injective and G is r.f., by [14], $\mathcal{A}$ is reversible, and by Lemma 8, $F_{\mathcal{A}}$ sends H-periodic configurations into H-periodic configurations. But because of (7) and the pigeonhole principle, there must exist two H-periodic configurations with the same image according to $F_{\mathcal{A}}$, which contradicts injectivity of $\mathcal{A}$. □

4 The nonwandering property

Bartholdi's theorem can be expanded by adding more properties that are satisfied by every surjective CA if and only if the underlying group is amenable.

Definition 22 *Let $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ be a cellular automaton over a group G. $\mathcal{A}$ is* nonwandering *if for every nonempty open set $U \subset Q^G$ there exists $t \geq 1$ such that $F_{\mathcal{A}}^t(U) \cap U \neq \emptyset$.*

Remark 23 *If $\mathcal{A}$ is μ-recurrent for some probability measure μ with* full support*—i.e., no nonempty open set is μ-null—then $\mathcal{A}$ is nonwandering.*

Observe that, for the latter to hold, it is not necessary that every open set be measurable: it is sufficient that every open set contains a measurable open set of positive measure, which is the case for μ_Π.

We say that a state $q_0 \in Q$ is *spreading* for $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ if for every $u \in Q^{\mathcal{N}}$ such that $u_i = q_0$ for some $i \in \mathcal{N}$ we have $f(u) = q_0$.

Lemma 24 *A nonwandering nontrivial* CA *has no spreading state.*

By nontrivial, we mean that $|\mathcal{N}| > 1$ and $|Q| > 1$.

Proof: Suppose that the nontrivial CA $\mathcal{A} = \langle Q, \mathcal{N}, f \rangle$ has a spreading state q_0. Let $U = C(\mathcal{N} \cup \{1_G\}, p)$ where $p_i = q_0 \neq p_{1_G}$ for some $i \in \mathcal{N} \setminus \{1_G\}$. Then $F^t_{\mathcal{A}}(U) \cap U = \emptyset$ for every $t \geq 1$. □

It follows from the definitions and the orphan pattern principle that if a CA is nonwandering (or μ_Π-recurrent), then it is surjective. The next two statements are immediate consequences of the Poincaré recurrence theorem.

Proposition 25 *Let G be a group and let $\mathcal{A}$ be a CA on G. If $\mathcal{A}$ preserves μ_Π (equivalently, if $\mathcal{A}$ is balanced) then $\mathcal{A}$ is μ_Π-recurrent.*

Corollary 26 *Let $\mathcal{A}$ be a CA on an amenable group G. If $\mathcal{A}$ is surjective then $\mathcal{A}$ is μ_Π-recurrent.*

We might ask what the role of amenability in Corollary 26 is. The following counterexample shows that surjective CA on paradoxical groups may fail to be nonwandering.

Example 27 *Let G be a non-amenable group, ϕ a bounded-propagation two-to-one compressing map with propagation set S, $\preceq$ a total ordering of S and $Q = S \times \{0,1\} \times S \sqcup \{q_0\}$, where $q_0 \notin S \times \{0,1\} \times S$. Let $\mathcal{A} = \langle Q, S, f \rangle$ with:*

$$\begin{array}{rccl} f: & Q^S & \to & Q \\ & u & \mapsto & \left| \begin{array}{ll} q_0 & \text{if } \exists s \in S, u_s = q_0, \\ (p, \alpha, q) & \text{if } \exists!(s,t) \in S \times S, s \prec t, u_s = (s, \alpha, p), u_t = (t, 1, q), \\ q_0 & \text{otherwise.} \end{array} \right. \end{array}$$

Clearly, such a CA cannot be nonwandering, as it is nontrivial and has the spreading state q_0. In particular, it is neither μ_Π-recurrent nor balanced.

Proposition 28 *The cellular automaton $\mathcal{A}$ from Example 27 is surjective.*

Proof: Let $x \in Q^G$, $i \in G$, $j = \phi(i)$: then $i = js$ for some $s \in S$, and there exists a unique $t \in S \setminus \{s\}$ such that $\phi(jt) = j$. If $x_j = q_0$, then set $y_i = (s, 0, s)$: otherwise, we can write $x_j = (p, \alpha, q)$. If $s \prec t$, then set $y_i = (s, \alpha, p)$; otherwise set $y_i = (s, 1, q)$. This definition has the property that for every $i \in G$, $y_i \in \{\phi(i)^{-1} i\} \times \{0,1\} \times S$.

Let us prove that the configuration y is a preimage of x by the global map of the CA. Let $j \in G$ and $s, t \in S$ such that $s \prec t$, $y_{js} \in \{s\} \times \{0,1\} \times S$, and $y_{jt} \in \{t\} \times \{0,1\} \times S$. Then $s = \phi(js)^{-1} js$ and $t = \phi(jt)^{-1} jt$, and $\phi(js) = \phi(jt) = j$: hence, there exists *exactly one* such pair (s, t). If $x_j = q_0$, then the definition of y gives $y_{jt} = (t, 0, t)$, and f will apply its third subrule. If x_j is written (p, α, q), then $y_{js} = (s, \alpha, p)$ and $y_{jt} = (t, 1, q)$, and f will apply its second subrule. □

Observe that, in the proof of Proposition 28, for every configuration we construct a preimage which does not contain the state q_0. This, together with Proposition 20, leads us to the following.

Remark 29 A finitely generated *group G is paradoxical if and only if there exists a* CA *on G which takes a nonrich configuration into a rich one.*

5 Normal configurations

The results from Sections 3 and 4, together with the existing literature, show that Theorem 1 is true. We now move on to Theorem 2 and search for a suitable set $U \subseteq Q^G$ of full measure with a null preimage. To construct such a set, we introduce the concept of *normal* configuration, according to some parameters: normality shall thus be a *quantitative* concept, more precise than the nonwandering property which is a *qualitative* one.

Our definition is based on the one for normal infinite words. Let $U \subset \mathbb{N}$, and denote $P(U|n) = |\{i < n \mid i \in U\}|/n$ for $n \in \mathbb{N}$. The *lower density*, *upper density* and *density* of U are, respectively, the liminf, limsup and limit, when it exists, of $P(U|n)$ when n goes to infinity. Given an infinite word w, an *occurrence* in w of a *finite* word u is a position $i \geq 0$ such that $w_{[i:i+|u|-1]} = u$. Call $\operatorname{occ}(u, w)$ the set of occurrences of u in w. An infinite word w on the alphabet Q is said to be m*-normal*, $m \in \mathbb{N}$, if for every $u \in Q^m$ the set $\operatorname{occ}(u, w)$ has density $|Q|^{-m}$; w is *normal* if it is m-normal for every $m \in \mathbb{N}$.

The notion of m-normality admits a characterization which will be helpful in the next section.

Theorem 30 (Niven and Zuckerman; cf [21]) *Let $m \geq 1$ and let Q be a finite set. An infinite word over Q is m-normal if and only if it is 1-normal when considered as a word over Q^m.*

Let now $h : \mathbb{N} \to G$ be an injective function. For $U \subseteq G$ we define the *lower density* $\operatorname{dens\,inf}_h U$, *upper density* $\operatorname{dens\,sup}_h U$, and *density* (if it exists) $\operatorname{dens}_h U$ as those of the preimage $h^{-1}(U)$.

Note that we do *not* require that h be *bijective*. The reason for this, is that the structure of general groups is usually not as convenient as that of $\mathbb{Z}^d$, and it is not always possible to subdivide a group into "nicely shaped blocks" (such as the hypercubes of $\mathbb{Z}^d$) and see a configuration as a "coarser-grained" configuration *on the same group*. We will discuss this in further detail later on in this section.

Definition 31 *Let $h : \mathbb{N} \to G$ be an injective function. A configuration $c : G \to Q$ is* normal on support E w.r.t. h *(briefly, h-E-normal) if for every pattern $p : E \to Q$*

$$\operatorname{dens}_h \operatorname{occ}(p, c) = |Q|^{-|E|} . \tag{8}$$

For $E = \{1_G\}$ we say that c is h-1-normal. If c is h-E-normal for every $E \in \mathcal{PF}(G)$, we say that c is h-normal. We omit h if it is clear from the context.

This definition passes a basic "sanity check": normality on larger sets ensures normality on smaller sets.

Remark 32 *Let $E, F \in \mathcal{PF}(G)$ with $E \subseteq F$. If $c : G \to Q$ is F-normal, then it is also E-normal.*

Proof: Let $p : E \to Q$. Every $z \in G$ which is an occurrence of p in c, is also an occurrence of exactly one of the $|Q|^{|F|}$ patterns $\tilde{p} : F \to Q$ that extend p; vice versa, if $\tilde{p}|_E = p$, then each occurrence of $\tilde{p}$ is also an occurrence of p. Hence, whatever n is,

$$\frac{|\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}|}{n} = \sum_{\tilde{p}:F \to Q,\, \tilde{p}|_E = p} \frac{|\{i < n \mid h(i) \in \operatorname{occ}(\tilde{p}, c)\}|}{n} .$$

As the right-hand side has $|Q|^{|F|-|E|}$ summands, each converging to $|Q|^{-|F|}$ by hypothesis, the left-hand side converges to $|Q|^{-|E|}$. As p is arbitrary, c is E-normal. □

The vice versa of Lemma 32 does not hold: being h-E-normal for every proper subset E of F does not imply being h-F-normal.

Example 33 *Let $G = \mathbb{Z}$, $Q = \{0, 1\}$, $h(i) = i$ for every $i \geq 0$, $c(x) = x \mod 2$ for every $x \in \mathbb{Z}$. Then c is h-$\{0\}$-normal and h-$\{1\}$-normal but not h-$\{0, 1\}$-normal.*

Our aim is to prove that, at least under certain conditions on h and E, μ_Π-almost all configurations are h-E-normal: to do this, we need criteria for h-E-normality. A basic test is provided by

Lemma 34 *Let $|Q| > 1$, $E \in \mathcal{PF}(G)$,$c : G \to Q$. The following are equivalent.*

1. *c is h-E-normal.*
2. *For every $p : E \to Q$, $\operatorname{dens\,inf}_h \operatorname{occ}(p, c) \geq |Q|^{-|E|}$.*
3. *For every $p : E \to Q$, $\operatorname{dens\,sup}_h \operatorname{occ}(p, c) \leq |Q|^{-|E|}$.*

Proof: Clearly, Point 1 implies Points 2 and 3, and Points 2 and 3 together imply Point 1. We then only have to prove that Points 2 and 3 are equivalent: this will be easy once we observe that, for every $n > 0$,

$$\sum_{p:E\to Q} |\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}| = n \ , \tag{9}$$

which expresses the obvious fact that *every* point is an occurrence of *some* pattern with support E. Suppose that for some $\bar{p} : E \to Q$, $\delta > 0$ we have $|\{i < n \mid h(i) \in \operatorname{occ}(\bar{p}, c)\}|/n < |Q|^{-|E|} - \delta$ for infinitely many values of n: because of (9), for those values we also have

$$\sum_{p:E\to Q,\, p\neq\bar{p}} \frac{|\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}|}{n} > \left(1 - |Q|^{-|E|}\right) + \delta \ .$$

Therefore, for all such values of n, there must be some $p : E \to Q$, $p \neq \bar{p}$ with $|\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}|/n > |Q|^{-|E|} + \delta/(|Q|^{|E|} - 1)$: since the n's are infinitely many and the p's are finitely many, there must be at least one $p : E \to Q$ such that $\limsup_{n\to\infty} |\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}|/n > |Q|^{-|E|}$. The converse implication is proven similarly. □

Lemma 34 has an immediate consequence, which will have great importance later.

Lemma 35 *Let $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ be a nontrivial* CA *on G with a spreading state q_0 and s, t two distinct elements of $\mathcal{N}$. For every injective function $h : \mathbb{N} \to G$, if $c : G \to Q$ is h-$\{s, t\}$-normal, then $F_\mathcal{A}(c)$ is not h-1-normal.*

Proof: Since q_0 is spreading, $\operatorname{occ}(q_0, F_\mathcal{A}(c))$ contains $\operatorname{occ}(p, c)$ for every $\{s, t\}$-pattern p such that $p(s) = q_0$ or $p(t) = q_0$. Given c's h-$\{s, t\}$-normality, each of these $2|Q|-1$ patterns has density $1/|Q|^2$: therefore, $\operatorname{dens\,inf}_h \operatorname{occ}(q_0, F_\mathcal{A}(c)) \geq (2|Q| - 1)/|Q|^2 > 1/|Q|$ since $|Q| > 1$, and $F_\mathcal{A}(c)$ cannot be h-1-normal. □

For $p : E \to Q$, $k \geq 1$, and $h : \mathbb{N} \to G$ injective, let

$$L_{h,p,k,n} = \left\{ c : G \to Q \;\middle|\; \frac{|\{i < n \mid h(i) \in \operatorname{occ}(p, c)\}|}{n} \leq \frac{1}{|Q|^{|E|}} - \frac{1}{k} \right\} \ . \tag{10}$$

Observe that $L_{h,p,n,k}$ is a finite union of cylinders. By definition, $\operatorname{dens\,inf}_h \operatorname{occ}(p, c) < |Q|^{-|E|}$ if and only if there exists $k \geq 1$ such that $c \in L_{h,p,k,n}$ for infinitely many values of n, *i.e.*, if $c \in$

$\limsup_n L_{h,p,k,n} = \bigcap_{n\geq 1} \bigcup_{m\geq n} L_{h,p,k,m}$: this set, which we call $L_{h,p,k}$, belongs to the σ-algebra Σ_C generated by the cylinders. Then

$$L_{h,E} = \bigcup_{p\in Q^E, k\geq 1} L_{h,p,k} \tag{11}$$

is the set of all the configurations $c \in Q^G$ that are *not* h-E-normal. If each $L_{h,p,k}$ has measure 0, then—as the p's are finitely many and the k's are countably many—so has $L_{h,E}$, and almost all configurations are h-E-normal. This, in the classical case of infinite words over a $|Q|$-ary alphabet, is achieved for $E = \{0, \ldots, r-1\}$, via estimates such as the following.

Proposition 36 (Chernoff bound [8]; cf. [2, Lemma 6.56]) *Let $Y_0, \ldots, Y_{n-1}$ be* independent *nonnegative random variables; let $S_n = Y_0 + \ldots + Y_{n-1}$, and let $\mu = \mu(n)$ be the average of S_n. For every $\delta \in (0,1)$,*

$$\mathbb{P}\left(S_n < \mu \cdot (1-\delta)\right) < e^{-\frac{\mu\delta^2}{2}} . \tag{12}$$

In particular, if the Y_i's are Bernoulli trials with probability p, and $0 < \varepsilon < \min(p, 1-p)$, then for $\delta = \varepsilon/p$

$$\sum_{0\leq k<n\cdot(p-\varepsilon)} \binom{n}{k} p^k (1-p)^{n-k} < e^{-\frac{\varepsilon^2 n}{2p}} . \tag{13}$$

The Chernoff bound, together with the Borel-Cantelli lemma, allows to prove that the set of non-normal infinite words has product measure zero. However, one of the reasons why we can express m-normality of sequences as 1-normality of other sequences, is that *the interval* $\{0, \ldots, m-1\}$ *is a coset of a submonoid of* $\mathbb{N}$ *isomorphic to* $\mathbb{N}$: as any subgroup of finite index of $\mathbb{Z}^d$ is isomorphic to $\mathbb{Z}^d$, it is possible to adapt the classical argument for infinite words so that it works for d-dimensional configurations. But a subgroup of index 2 of the free group on two generators is free on *three* generators (cf. [16, Theorem 2.10]) and thus is not isomorphic to it; therefore, if we just mimic the classical argument and consider patterns with support a coset of a subgroup, we need in general to change the underlying group! Otherwise, when estimating the number of occurrences of a pattern, we have to deal with *non-independent* events, and cannot (straightforwardly) apply the Chernoff bound.

This is the key reason we mentioned earlier for our hypothesis that h may be non-surjective. In fact, if the E-shaped neighborhoods of the points of $h(\mathbb{N})$ are pairwise disjoint, then the events of the form "$h(i)$ is an occurrence of p" for $p : E \to Q$ are independent, and we *can* apply the Chernoff bound while keeping the same underlying group.

Lemma 37 *Let E be a finite subset of G and let $h : \mathbb{N} \to G$ satisfy $h(n)E \cap h(m)E = \emptyset$ for every $n \neq m$. Then $\mu_\Pi(L_{h,E}) = 0$, i.e., μ_Π-almost all $c : G \to Q$ are h-E-normal.*

Proof: As the sets $h(i)E$, $i \geq 0$, are pairwise disjoint, the Boolean random variables Y_i which take value 1 if and only if $c^{h(i)}\big|_E = p$, are independent and identically distributed according to a Bernoulli distribution of parameter $t = |Q|^{-|E|}$. If $S_n = Y_0 + \ldots + Y_{n-1}$, then

$$L_{h,p,k,n} = \{c : G \to Q \mid S_n < n \cdot |Q|^{-|E|} \cdot (1 - |Q|^{|E|}/k)\} \ :$$

as $\mu = n \cdot |Q|^{-|E|}$ is precisely the average of S_n, for $\delta = |Q|^{|E|}/k$ the Chernoff bound tells us that

$$\mu_\Pi(L_{h,p,k,n}) = \mathbb{P}\left(\{S_n < \mu \cdot (1-\delta)\}\right) < e^{-\frac{|Q|^{|E|}}{2k^2}n}$$

decreases exponentially in n for fixed k and p. By the Borel-Cantelli lemma,

$$\mu_\Pi(L_{h,p,k}) = \mu_\Pi\left(\limsup_{n\geq 1} L_{h,p,k,n}\right) = 0 \ :$$

as this holds for each of the countably many pairs (p,k) with $p : E \to Q$ and $k \geq 1$, the thesis follows. □

Observe that there is no need for the group G to be countable.

Proof of Theorem 2: Choose S, Q, and $\mathcal{A}$ as by Example 27; let $h : \mathbb{N} \to G$ be a function such that $h(n)S \cap h(m)S = \emptyset$ for $n \neq m$. Let U be the set of h-1-normal configurations: the hypotheses of Lemma 37 are also satisfied for $E = \{1_G\}$, which means that $\mu_\Pi(U) = 1$. By Lemma 35, the images via $F_\mathcal{A}$ of h-S-normal configurations are not h-1-normal: thus, the preimages of the elements of U must belong to $L_{h,S}$, which is a μ_Π-null set by Lemma 37. □

Observe that Theorem 2 holds precisely because on non-amenable groups there are surjective CA which are not balanced.

Proposition 38 *Let G be an infinite group and let $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ be a* CA *on G. The following are equivalent.*

1. *$\mathcal{A}$ is balanced.*
2. *For every injective function $h : \mathbb{N} \to G$ and $E \in \mathcal{PF}(G)$, if $c : G \to Q$ is h-$E\mathcal{N}$-normal then $F_\mathcal{A}(c)$ is h-E-normal.*

Proof: Let $c \in Q^G$, $p \in Q^E$. An arbitrary $g \in G$ is an occurrence of p in $F_\mathcal{A}(c)$ if and only if it is an occurrence in c of one of the patterns $\bar{p} : E\mathcal{N} \to Q$ such that $f(\bar{p}) = p$. Consequently, for every $n \in \mathbb{N}$,

$$\frac{|\{i < n \mid h(i) \in \mathrm{occ}(p, F_\mathcal{A}(c))\}|}{n} = \sum_{f(\bar{p})=p} \frac{|\{i < n \mid h(i) \in \mathrm{occ}(\bar{p}, c)\}|}{n} \ . \tag{14}$$

If $\mathcal{A}$ is balanced and c is h-$E\mathcal{N}$-normal, then the right-hand side of (14) is a sum of $|Q|^{|E\mathcal{N}|-|E|}$ summands, each converging to $|Q|^{-|E\mathcal{N}|}$ for $n \to \infty$: as p is arbitrary, $F_\mathcal{A}(c)$ is h-E-normal.

Suppose then that $\mathcal{A}$ is not balanced. Then there exist $E \in \mathcal{PF}(G)$ and $p : E \to Q$ such that $|f^{-1}(p)| > |Q|^{|E\mathcal{N}|-|E|}$. Let $h : \mathbb{N} \to G$ be a function such that $h(n)E\mathcal{N} \cap h(m)E\mathcal{N} = \emptyset$ for $n \neq m$: then $\mu_\Pi(L_{h,E\mathcal{N}}) = 0$ by Lemma 37, so there exists an h-$E\mathcal{N}$-normal configuration c. For such c, the right-hand side of (14) is a sum of more than $|Q|^{|E\mathcal{N}|-|E|}$ summands, each converging to $|Q|^{-|E\mathcal{N}|}$ for $n \to \infty$: hence $F_\mathcal{A}(c)$ is not h-E-normal. □

Corollary 39 *Let G be an infinite amenable group, $\mathcal{A} = \langle Q, \mathcal{N}, f\rangle$ a surjective* CA *on G and $h : \mathbb{N} \to G$ an injective function. If $c : G \to Q$ is h-normal then so is $F_\mathcal{A}(c)$.*

6 Martin-Löf random configurations

We are now left with the task of proving Theorem 3. To do this, we need to define *Martin-Löf randomness* for configurations. This requires some hypotheses on the underlying group: we must be able not only to enumerate its elements, but also to do it in a computable way.

Let G be a group, S a set of generators for G, and R a set of words on $S \cup S^{-1}$. We say that $\langle S, R \rangle$ is a *presentation* of G, and write $G = \langle S, R \rangle$, if G is isomorphic to the quotient G_S / K_R, where G_S is the free group on S (consisting of reduced words on $S \cup S^{-1}$) and K_R is the normal subgroup of G_S generated by R, *i.e.*, the intersection of all normal subgroups of G_S that contain the elements identified by the words in R. The *word problem* (briefly, w.p.) for the group $G = \langle S, R \rangle$ is the set of words on $S \cup S^{-1}$ that represent the identity element of G. Although this set depends on the choice of the presentation, its decidability does not; and although the problem is not decidable even for finitely generated groups, it is so for the Euclidean groups $\mathbb{Z}^d$, the free groups, Gromov's *hyperbolic groups* [11] which generalize free groups, and many more.

An *indexing* of a countable group G is a bijection $\phi : \mathbb{N} \to G$; we often write $G = \{g_i \mid i \geq 0\}$, to mean $g_i = \phi(i)$. An indexing is *admissible* if there exists a computable function $m : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that $g_i \cdot g_j = g_{m(i,j)}$ for every $i, j \in \mathbb{N}$. In this case, there is also a computable function $\iota : \mathbb{N} \to \mathbb{N}$ such that $g_i^{-1} = g_{\iota(i)}$ for every $i \in \mathbb{N}$: in fact, $\iota(i)$ is the only $j \in \mathbb{N}$ such that $m(i,j) = \phi^{-1}(1_G)$.

Proposition 40 (Rabin, 1960; cf. [23, Theorem 4]) *A finitely generated group has an admissible indexing if and only if it has decidable word problem.*

Proof: Assume $G = \{g_i \mid i \geq 0\}$ is an admissible indexing, *i.e.*, $g_i \cdot g_j = g_{m(i,j)}$ for every $i, j \in \mathbb{N}$ and m is computable. Let S be a set of generators for G, and $u = u_1 \dots u_\ell$ a word over $S \cup S^{-1}$; say $u_r = g_{i_r}$ for every $r \in \{1, \dots, \ell\}$. We can decide whether u and 1_G identify the same element of G by inductively computing the sequence (a_r) with $a_1 = i_1$ and $a_r = m(a_{r-1}, i_r)$ for $r = 2, \dots, \ell$; $u = 1_G$ if and only if a_ℓ is the (unique) index representing 1_G.

Suppose now that G has decidable word problem. Let S be a finite set of generators for G: define an ordering on $S \cup S^{-1}$. A computable bijection $\phi : \mathbb{N} \to G$ can be obtained by enumerating, in lexicographic order, first $D_0 = \{1_G\}$, then $D_1 \setminus D_0 = S \cup S^{-1}$, then $D_2 \setminus D_1$, and so on. Moreover, the function $m : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ given by $m(i,j) = \phi(w)$ where w is a word on $S \cup S^{-1}$ representing $g_i \cdot g_j$, is computable. $\square$

Throughout this section, G will be an infinite, finitely generated group with decidable word problem, and $\phi : \mathbb{N} \to G$ an admissible indexing: we write $G = \{g_i\}_{i \geq 0}$ to mean $g_i = \phi(i)$.

We recall the definition of Martin-Löf randomness for infinite words (cf. [17] and [2, Sections 5.4 and 6.2]). A *sequential Martin-Löf test* (briefly, M-L test) is a recursively enumerable set $U \subseteq \mathbb{N} \times Q^*$ such that the *level sets* $U_n = \{x \in Q^* \mid (n, x) \in U\}$ satisfy the following conditions:

1. For every $n \geq 1$, $U_{n+1} \subseteq U_n$.
2. For every $n \geq 1$ and $m \geq n$, $|U_n \cap Q^m| \leq |Q|^{m-n} / (|Q| - 1)$.
3. For every $n \geq 1$ and $x, y \in Q^*$, if $x \in U_n$ and $y \in xQ^*$ then $y \in U_n$.

An infinite word w *fails* a sequential M-L test U if $w \in \bigcap_{n \geq 0} U_n Q^{\mathbb{N}}$; the word w is *Martin-Löf random* if w does not fail any sequential M-L test. Observe that, according to this definition, if $\eta : \mathbb{N} \to \mathbb{N}$ is a computable bijection, then w is M-L random if and only if $w \circ \eta$ is M-L random. It is well known (cf. [17] and [2, Theorem 6.61]) that M-L random words are normal.

Thanks to an approach by Hertling and Weihrauch, it is possible to define Martin-Löf randomness of infinite words in a way that allows to introduce the concept in the more general context of configurations. The prodiscrete topology and product measure on $Q^{\mathbb{N}}$ are defined similarly as on Q^G. Given two sequences $\mathcal{U} = \{U_i\}_{i \geq 0}, \mathcal{V} = \{V_j\}_{j \geq 0}$ of open subsets of $Q^{\mathbb{N}}$, we say that $\mathcal{U}$ is *$\mathcal{V}$-computable* if there is a *recursively enumerable* set $A \subseteq \mathbb{N}$ such that

$$U_i = \bigcup_{j \in \mathbb{N} : \pi(i,j) \in A} V_j \quad \forall i \geq 0 \,, \tag{15}$$

where $\pi(i,j) = (i+j)(i+j+1)/2 + j$ is the standard primitive recursive bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. Given an ordering $Q = \{q_0, \ldots, q_{|Q|-1}\}$, let $\widetilde{B}_i = w_i Q^{\mathbb{N}}$ where w_i is the i-th element of Q^* in the *length-lexicographic order—i.e.*, w_0 is the empty word, $w_1 = q_0, \ldots, w_{|Q|} = q_{|Q|-1}, w_{|Q|+1} = q_0 q_0$, $w_{|Q|+2} = q_0 q_1$, and so on—and let $\widetilde{B}'_i = \bigcap_{j \in E(i+1)} \widetilde{B}_j$, where $n \in E(i)$ if and only if the n-th bit in the binary expansion of i is 1: then $\widetilde{B}'$ is an enumeration of a base of the prodiscrete topology of $Q^{\mathbb{N}}$. Observe that the property "$w \in \widetilde{B}'_j$" only depends on a prefix u of w which can be computed from j.

Proposition 41 (Hertling and Weihrauch; cf. [2, Theorem 6.99]) *Let $w : \mathbb{N} \to Q$ be an infinite word. The following are equivalent.*

1. *w is Martin-Löf random.*
2. *For every $\widetilde{B}'$-computable sequence $\mathcal{U} = \{U_n\}_{n \geq 0}$ of open subsets of $Q^{\mathbb{N}}$ such that $\mu_\Pi(U_n) < 2^{-n}$ we have $w \notin \bigcap_{n \geq 0} U_n$.*

We can now define Martin-Löf randomness for configurations, in analogy with the previous formalism. Given an ordering $Q = \{q_0, \ldots, q_{|Q|-1}\}$, we define a computable bijective enumeration B of the elementary cylinders as $B_{|Q|i+j} = C(g_i, q_j)$. To enumerate the cylinders, we define a computable bijection $\Psi : \mathcal{PF}(\mathbb{N}) \to \mathbb{N}$ as $\Psi(E) = \sum_{n \in E} 2^n$ (so that $\Psi(\emptyset) = 0$) and set $B'_i = \bigcap_{j \in \Psi^{-1}(i)} B_j$. Observe that the property "$c \in B_j$" only depends on the values of c on a finite subset which can be computed from j. If $\mathcal{U}$ and $\mathcal{V}$ are families of open subsets of Q^G, we say that $\mathcal{U}$ is $\mathcal{V}$-computable if there exists a r.e. set A such that (15) holds.

Definition 42 *Let G be a f.g. group with decidable word problem; let $\Sigma_C \subseteq Q^G$ be the σ-algebra generated by the cylinders (i.e., as G is countable, the Borel σ-algebra) and let $\mu : \Sigma_C \to [0,1]$ be a computable probability measure. A B'-computable family $\mathcal{U} = \{U_n\}_{n \geq 0}$ of open subsets of Q^G is a* Martin-Löf μ-test *(briefly, M-L μ-test) if $\mu(U_n) \leq 2^{-n}$ for every $n \geq 0$. A configuration $c \in Q^G$* fails *a M-L μ-test $\mathcal{U}$ if $c \in \bigcap_{n \geq 0} U_n$. c is* μ-random in the sense of Martin-Löf *(briefly, M-L μ-random) if it does not fail any M-L μ-test.*

As the set of configurations failing a given M-L μ_Π-test is μ_Π-null, and the family of M-L μ_Π-tests is countable, the set of M-L μ_Π-random configurations has full measure.

The next statement has been used by Calude et al. ([3]; cf. [2, Section 9.5]) for CA on $\mathbb{Z}^d$ and one specific admissible indexing. We need it in our more general context: however, the proof is similar.

Lemma 43 *Let $\phi : \mathbb{N} \to G$ be an admissible indexing.*

1. *The function $\bar{\phi} : Q^G \to Q^{\mathbb{N}}$ defined by $\bar{\phi}(c) = c \circ \phi$ is a homeomorphism.*
2. *For every $U \in \Sigma_C$, $\mu_\Pi(\bar{\phi}(U)) = \mu_\Pi(U)$.*
3. *A family $\mathcal{U}$ of open subsets of Q^G is B'-computable if and only if the corresponding family $\bar{\phi}(\mathcal{U})$ of open subsets of $Q^{\mathbb{N}}$ is $\widetilde{B}'$-computable.*

Corollary 44 (cf. [2, Theorem 9.10]) *Let $\phi : \mathbb{N} \to G$ be an admissible indexing. Then $c : G \to Q$ is M-L μ_Π-random if and only if $c \circ \phi : \mathbb{N} \to Q$ is M-L random.*

As a consequence of Corollary 44, the definition of Martin-Löf μ_Π-randomness does not depend on the choice of the admissible indexing. In fact, if $\phi, \psi : \mathbb{N} \to G$ are admissible indexings, then $\eta = \phi^{-1} \circ \psi : \mathbb{N} \to \mathbb{N}$ is a computable bijection such that $(c \circ \phi) \circ \eta = c \circ \psi$.

Given a pattern p, the set of configurations where p has no occurrence is an intersection of a countably infinite, computable family of cylinders U_i having equal product measure $\mu_\Pi(U_i) = m < 1$. It is then straightforward to construct a M-L μ_Π-test that every such configuration fails. We have thus

Remark 45 *Every μ_Π-random configuration is rich.*

We can now prove Theorem 3. Let us start with the “only if” direction.

Lemma 46 *Let $\mathcal{U}$ be a B'-computable sequence and $\mathcal{A}$ a CA on G. Then $F_{\mathcal{A}}^{-1}(\mathcal{U})$ is a B'-computable sequence.*

Proof: Let $\pi(i,j) = (i+j)(i+j+1)/2+j$ and let $L, K : \mathbb{N} \to \mathbb{N}$ be the two primitive recursive functions such that $\pi(L(n), K(n)) = n$ for every $n \in \mathbb{N}$. Let $A \subseteq \mathbb{N}$ be a r.e. set such that $U_i = \bigcup_{\pi(i,j) \in A} B'_j$: then

$$F_{\mathcal{A}}^{-1}(U_i) = \bigcup_{\pi(i,j) \in A} F_{\mathcal{A}}^{-1}(B'_j) \; .$$

As $\mathcal{A}$ is a CA, for every $j \in \mathbb{N}$ there exists $E_j \in \mathcal{PF}(\mathbb{N})$ such that $F_{\mathcal{A}}^{-1}(B'_j) = \bigcup_{k \in E_j} B'_k$; moreover, the function $j \mapsto E_j$ is computable because G has decidable word problem. Then

$$Z = \{n \in \mathbb{N} \mid \exists j \in \mathbb{N} : \pi(L(n), j) \in A, K(n) \in E_j\}$$

is a recursively enumerable set such that $F_{\mathcal{A}}^{-1}(U_i) = \bigcup_{\pi(i,k) \in Z} B'_k$ for every $i \geq 0$. □

Proposition 47 *Let $\mathcal{A}$ be a CA over G. If $F_{\mathcal{A}}(c)$ is μ_Π-random whenever c is, then $\mathcal{A}$ is surjective. If $\mathcal{A}$ preserves μ_Π, then $F_{\mathcal{A}}(c)$ is μ_Π-random whenever c is.*

Proof: Since μ_Π-random configurations form a set of measure 1 and contain occurrences of every pattern, the first part is immediate. For the second part, if $F_{\mathcal{A}}\mu_\Pi = \mu_\Pi$, then by Lemma 46 the preimage of a M-L μ_Π-test is still a M-L μ_Π-test: but if $F_{\mathcal{A}}(c)$ fails $\mathcal{U}$, then c fails $F_{\mathcal{A}}^{-1}(\mathcal{U})$. □

Proof of Theorem 3, sufficiency of amenability: Suppose G is an amenable group. Let $\mathcal{A}$ be a surjective CA on G with alphabet Q: by Bartholdi's theorem, $\mathcal{A}$ preserves μ_Π. Let $c : G \to Q$ be a M-L μ_Π-random configuration: by Proposition 47, $F_{\mathcal{A}}(c)$ is also M-L μ_Π-random. □

To prove the "if" part of Theorem 3[(i)], we resort to normal configurations; in doing this, we need a result which is of interest by itself. We say that $a \in Q^{\mathbb{N}}$ is M-L random *relatively* to $b \in Q^{\mathbb{N}}$ if it is M-L random when computability is considered according to Turing machines with oracle b.

Proposition 48 (van Lambalgen's theorem [13]; cf. [9, Corollary 6.9.3]) *Let a and b be two infinite words over the alphabet Q and let c be the* interleaving *of a and b, i.e., $c(2n) = a(n)$ and $c(2n+1) = b(n)$ for every $n \geq 0$. The following are equivalent.*

1. *c is M-L random.*
2. *a is M-L random, and b is M-L random relatively to a.*
3. *b is M-L random, and a is M-L random relatively to b.*

The necessity of conditions 2 and 3 is clear: if, for example, $a = b$, then c is not M-L random.

Lemma 49 *Let G be an infinite f.g. group with decidable word problem. For every nonempty $E \in \mathcal{PF}(G)$ there exists a computable injective function $h : \mathbb{N} \to G$ satisfying the following properties:*

1. *$h(\mathbb{N})$ is a recursive subset of G with infinite complement.*
2. *$h(n)E \cap h(m)E = \emptyset$ for every $n \neq m$.*
3. *For any alphabet Q, every M-L μ_Π-random configuration $c : G \to Q$ is h-E-normal.*

Proof: Let $G = \{g_i\}_{i \geq 0}$ be an admissible indexing of G. Define an injective function $\iota : \mathbb{N} \to \mathbb{N}$ by putting $\iota(0) = 0$, and $\iota(n+1)$ the smallest k such that $g_{\iota(0)}E, \ldots, g_{\iota(n)}E, g_kE$ are pairwise disjoint: then ι is computable. If $E = \{e_0, \ldots, e_{k-1}\}$, then $\tilde{h}(kn + j) = g_{\iota(2n)} \cdot e_j$ and $h(n) = g_{\iota(2n)}$ are injective, computable, and satisfy point 1, and in addition, h satisfies point 2. Taking every other value of ι ensures that the complement of the codomain is infinite: we will need this in the next step.

Let now $c : G \to Q$ be a M-L μ_Π-random configuration. Then $v(i) = c(g_i)$ is a M-L random infinite word. As the codomain of $\tilde{h}$ is recursive and its complement is infinite, there exists a computable bijection $\pi : \mathbb{N} \to \mathbb{N}$ such that $g_{\pi(2m)} = \tilde{h}(m)$ for every $m \in \mathbb{N}$. Then $v \circ \pi$ is a M-L random infinite word: by van Lambalgen's theorem, $w(i) = (v \circ \pi)(2i)$ is M-L random, thus also k-normal. By Theorem 30, for every $u \in Q^k$,

$$\lim_{n \to \infty} \frac{|\mathrm{occ}(u, w) \cap \{0, k, \ldots, (n-1)k\}|}{n} = \frac{1}{|Q|^k} ,$$

which, as $w(kn + j) = c(\tilde{h}(kn + j)) = c(h(n) \cdot e_j)$, is the same as saying that c is h-E-normal. □

Proof of Theorem 3, necessity of amenability: Let G be a non-amenable f.g. group with decidable word problem. Define S, Q and $\mathcal{A}$ as by Example 27. Construct $h : S \to G$ as by Lemma 49, with $E = S \cup \{1_G\}$. Let $c \in Q^G$: we will show that at most one between c and $F_\mathcal{A}(c)$ is M-L μ_Π-random.

Suppose that c is indeed M-L μ_Π-random. Then c is h-E-normal by the choice of h, thus h-S-normal by Lemma 32. By Lemma 35, $F_\mathcal{A}(c)$ is not h-1-normal, and cannot be M-L μ_Π-random. □

(i) The proof in [4] actually presented an error.

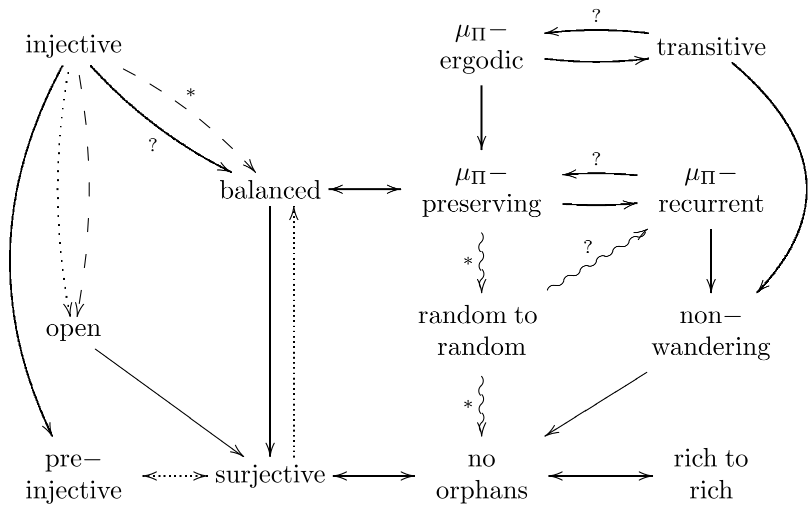

Figure 1: A diagram of implications between cellular automata properties. Full lines hold for every group; dotted lines hold for amenable groups; dashed lines hold for residually finite groups; wavy lines hold for finitely generated groups with decidable word problem. Starred implications are proven in the present paper. Implications with a question mark are conjectured. Transitivity and ergodicity have not been discussed here, but we include their implications since they are similarly conjectured equivalent for CA.

7 Conclusions

We have shown that several characterizations of surjective CA, which were known from [3] to hold on $\mathbb{Z}^d$, also hold in the more general case of amenable groups: actually, not only do they hold, but each of them characterizes amenable groups in the sense that it holds for CA on G if and only if G is amenable. This allows us to draw the graph of implications in Figure 1. In addition to this, we have determined the level to which the balancedness theorem fails on non-amenable groups: as in this case there are sets of full measure with null preimages, such failure can rightly be called catastrophic.

This is a remarkable result that sheds new light on the links between cellular automata theory and group theory. There are, however, several more questions left open. The most important of these, is surely whether Myhill's theorem as well holds only for CA on amenable groups, *i.e.*, whether pre-injectivity implies surjectivity if and only if the underlying group is amenable. The question is open and presumably very difficult (cf. the discussion in [1]) also because, contrary to the other properties examined—which always *imply* surjectivity regardless of the properties of the underlying group—pre-injectivity appears to be *independent* of it, as follows from the counterexamples in [15] and [7]. Another open problem concerns the existence of an injective CA which is not balanced: a negative answer would solve *Gottschalk's conjecture* that injective CA over arbitrary groups are surjective. Further questions arise from Remark 29, such as which cellular automata are capable of taking a nonrich configuration into a rich one, and whether this is linked to balancedness. More generally, the relationships between all properties linked here to surjectivity are, in many cases, yet to be explored.

Acknowledgements

This research was supported by the European Regional Development Fund (ERDF) through the Estonian Centre of Excellence in Theoretical Computer Science (EXCS) and the ICT National Programme project "Coinduction"; by the Estonian Research Fund (ETF) through grant nr. 7520; by the Estonian Ministry of Education and Research target-financed research theme no. 0140007s12; and by the Academy of Finland Grant 131558. We also thank the anonymous referees for their insightful comments and suggestions.

References

[1] Bartholdi, L. (2010) Gardens of Eden and amenability on cellular automata. *J. Eur. Math. Soc.* **12(1)**, 141–148.

[2] Calude, C. (2001) *Information and Randomness: An Algorithmic Perspective.* Springer Verlag.

[3] Calude, C., Hertling, P., Jürgensen, H. and Weihrauch, K. (2001) Randomness on full shift spaces. *Chaos, Solitons & Fractals* **12(3)**, 491–503.

[4] Capobianco, S., Guillon, P. and Kari, J. (2011) Garden-of-Eden-like theorems for amenable groups. In N. Fatès *et al.* (eds.), Procs. of Automata 2011, 233–242.

[5] Ceccherini-Silberstein, T. and Coornaert, M. (2009) Induction and restriction of cellular automata. *Ergod. Th. & Dynam. Sys.* **29**, 371–380.

[6] Ceccherini-Silberstein, T. and Coornaert, M. (2010) *Cellular Automata and Groups.* Springer Verlag.

[7] Ceccherini-Silberstein, T., Machì, A. and Scarabotti, F. (1999) Amenable groups and cellular automata. *Annales de l'Institut Fourier, Grenoble* **49(2)**, 673–685.

[8] Chernoff, H. (1952) A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23(4)**, 493–507.

[9] Downey, R.G. and Hirschfeldt, D. (2010) *Algorithmic Randomness and Complexity.* Springer.

[10] Fiorenzi, F. (2000) *Cellular automata and finitely generated groups.* Doctoral Thesis, Sapienza Università di Roma.

[11] Gromov, M. (1987) Hyperbolic Groups. In *Essays on Group Theory*, MSRI Publ. **8**, 75–263, Springer, New York.

[12] Katok, A. and Hasselblatt, B. (1995) *Introduction to the Modern Theory of Dynamical Systems.* Cambridge University Press.

[13] van Lambalgen, M. (1987) The axiomatization of randomness. *J. Symb. Logic* **55**, 1143–1167.

[14] Lawton, W. (1972) Note on symbolic transformation groups. *Not. Am. Math. Soc.* **19**, A375 (abstract).

[15] Machì, A. and Mignosi, F. (1993) Garden of Eden configurations for cellular automata on Cayley graph of groups. *SIAM J. Disc. Math.* **6**, 44–56.

[16] Magnus, W., Karrass, A. and Solitar, D. (1976) *Combinatorial Group Theory.* Dover.

[17] Martin-Löf, P. (1966) The Definition of Random Sequences. *Information and Control* **9**, 602–619.

[18] Maruoka, A. and Kimura, M. (1976) Condition for Injectivity of Global Maps for Tessellation Automata. *Inform. Control* **32(2)**, 158–162.

[19] Moore, E.F. (1962) Machine models of self-reproduction. *Proc. Symp. Appl. Math.* **14**, 17–33.

[20] Myhill, J. (1962) The converse of Moore's Garden-of-Eden theorem. *Proc. Amer. Mat. Soc.* **14**, 685–686.

[21] Niven, I. and Zuckerman, H.S. (1951) On the definition of normal numbers. *Pacific J. Math.* **1**, 103–109.

[22] Ornstein, D.S. and Weiss, B. (1987) Entropy and isomorphism theorems for actions of amenable groups. *J. Anal. Math.* **48**, 1–141.

[23] Rabin, M. (1960) Computable algebra, general theory and theory of computable fields. *Trans. AMS* **95**, 341-360.

[24] Weiss, B. (2000) Sofic groups and dynamical systems. *Sankhyā: Indian J. Stat.* **62**, 350–359.