Asymptotic Randomization of Sofic Shifts by Linear Cellular Automata*

Marcus Pivato
Department of Mathematics
Trent University
email: pivato@xaravve.trentu.ca

Reem Yassawi
Department of Mathematics
Trent University
email: ryassawi@trentu.ca

March 12, 2019

Abstract

Let $\mathbb{M} = \mathbb{Z}^D$ be a D-dimensional lattice, and $\mathcal{A}$ an abelian group. $\mathcal{A}^{\mathbb{M}}$ is then a compact abelian group; a *linear cellular automaton* (LCA) is a topological group endomorphism $\Phi : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ that commutes with all shift maps.

Suppose μ is a probability measure on $\mathcal{A}^{\mathbb{M}}$ whose support is a subshift of finite type or sofic shift. We provide sufficient conditions (on Φ and μ) under which Φ *asymptotically randomizes* μ, meaning that $\mathbf{wk}^*\text{-}\lim_{\mathbb{J} \ni j \to \infty} \Phi^j \mu = \eta$, where η is the Haar measure on $\mathcal{A}^{\mathbb{M}}$, and $\mathbb{J} \subset \mathbb{N}$ has Cesàro density 1. In the case when $\Phi = 1 + \boldsymbol{\sigma}$, we provide a condition on μ that is both necessary and sufficient.

MSC: Primary: 37B15; Secondary: 37A50

Let $D \geq 1$, and let $\mathbb{M} = \mathbb{Z}^D$ be the D-dimensional lattice. If $\mathcal{A}$ is a (discretely topologised) finite set, then $\mathcal{A}^{\mathbb{M}}$ is compact in the Tychonoff topology. For any $\mathsf{v} \in \mathbb{M}$, let $\boldsymbol{\sigma}^{\mathsf{v}} : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ be the shift map: $\boldsymbol{\sigma}^{\mathsf{v}}(\mathbf{a}) = [b_{\mathsf{m}}|_{\mathsf{m} \in \mathbb{M}}]$, where $b_{\mathsf{m}} = a_{\mathsf{m}-\mathsf{v}}$, $\forall \mathsf{m} \in \mathbb{M}$. A **cellular automaton** (CA) is a continuous map $\Phi : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ which commutes with all shifts: for any $\mathsf{m} \in \mathbb{M}$, $\boldsymbol{\sigma}^{\mathsf{m}} \circ \Phi = \Phi \circ \boldsymbol{\sigma}^{\mathsf{m}}$. Let η be the uniform Bernoulli measure on $\mathcal{A}^{\mathbb{M}}$. If μ is another probability measure on $\mathcal{A}^{\mathbb{M}}$, we say Φ **asymptotically randomizes** μ if $\mathbf{wk}^*\text{-}\lim_{\mathbb{J} \ni j \to \infty} \Phi^j \mu = \eta$, where $\mathbb{J} \subset \mathbb{N}$ has Cesàro density one.

If $\mathcal{A}$ is a finite abelian group, then $\mathcal{A}^{\mathbb{M}}$ is a product group, and η is the Haar measure. A **linear** cellular automaton (LCA) is a CA that is also an endomorphism of $\mathcal{A}^{\mathbb{M}}$. Linear cellular

*This work was partially supported by NSERC Canada.

automata are known to asymptotically randomize a wide variety of measures [1, 2, 3, 6, 10], including a those satisfying a correlation-decay condition called *harmonic mixing* [8, 9]. However, all known sufficient conditions for asymptotic randomization (and for harmonic mixing, in particular) require μ to have *full support*, ie. $\mathsf{supp}(\mu) = \mathcal{A}^{\mathbb{M}}$.

We here investigate asymptotic randomization when $\mathsf{supp}(\mu) \subsetneq \mathcal{A}^{\mathbb{M}}$. In particular we consider the case when $\mathsf{supp}(\mu)$ is a sofic shift or subshift of finite type. In §1, we demonstrate asymptotic randomization for any Markov random field that is *locally free*, a much weaker assumption than full support. However, in §2 we show that harmonic mixing is a rather restrictive condition, by exhibiting a measure whose support is a mixing sofic shift but which is *not* harmonically mixing.

Thus, in §3, we introduce the less restrictive concept of *dispersion mixing* (for measures) and the dual concept of *dispersion* (for LCA), and state our main result: any dispersive LCA asymptotically randomizes any dispersion mixing measure. In §4, we introduce *bipartite* LCA, a broad class exemplified by the automaton $1 + \boldsymbol{\sigma}$. We then show that any bipartite LCA is dispersive.

In §5, we show that any *uniformly mixing* and *harmonically bounded* measure is dispersion mixing. In particular, in §6, we show this implies that any mixing Markov measure (supported on a subshift of finite type), and any continuous factor of a mixing Markov measure (supported on a sofic shift) is dispersion mixing, and thus, is asymptotically randomized by any dispersive LCA (eg. $1 + \boldsymbol{\sigma}$). Thus, the example of §2 *is* asymptotically randomized, even though it is not harmonically mixing.

Finally, in §7, we refine the results of §3-§4 by introducing *Lucas mixing*, (a weaker condition than dispersion mixing) and showing that a measure is asymptotically randomized by the automaton $1 + \boldsymbol{\sigma}$ if and *only if* it is Lucas mixing.

Preliminaries & Notation:

Elements of $\mathcal{A}^{\mathbb{M}}$ are denoted by boldfaced letters (eg. $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$), and subsets by gothic letters (eg. $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$). Elements of $\mathbb{M}$ are sans serif (eg. l, m, n) and subsets are $\mathbb{U}, \mathbb{V}, \mathbb{W}$.

If $\mathbb{U} \subset \mathbb{M}$ and $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$ then $\mathbf{a}_{\mathbb{U}} = [a_{\mathsf{u}}|_{\mathsf{u}\in\mathbb{U}}]$ is the 'restriction' of $\mathbf{a}$ to an element of $\mathcal{A}^{\mathbb{U}}$. For any $\mathbf{b} \in \mathcal{A}^{\mathbb{U}}$, let $[\mathbf{b}] \ = \ \left\{\mathbf{c} \in \mathcal{A}^{\mathbb{M}} \ ; \ \mathbf{c}_{\mathbb{U}} = \mathbf{b}\right\}$ be the corresponding cylinder set. In particular, if $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$, then $[\mathbf{a}_{\mathbb{U}}] \ = \ \left\{\mathbf{c} \in \mathcal{A}^{\mathbb{M}} \ ; \ \mathbf{c}_{\mathbb{U}} = \mathbf{a}_{\mathbb{U}}\right\}$.

Measures: Let $\mathcal{M}(\mathcal{A}^{\mathbb{M}})$ be the set of Borel probability measures on $\mathcal{A}^{\mathbb{M}}$. If $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$ and $\mathbb{I} \subset \mathbb{M}$, then let $\mu_{\mathbb{I}} \in \mathcal{M}(\mathcal{A}^{\mathbb{I}})$ be the marginal projection of μ onto $\mathcal{A}^{\mathbb{I}}$. If $\mathbb{J} \subset \mathbb{M}$ and $\mathbf{b} \in \mathcal{A}^{\mathbb{J}}$, then let $\mu^{(\mathbf{b})} \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$ be the condition probability measure in the cylinder set $[\mathbf{b}]$. In other words, for any $\mathfrak{X} \subset \mathcal{A}^{\mathbb{M}}$, $\quad \mu^{(\mathbf{b})}[\mathfrak{X}] \ = \ \mu\left(\mathfrak{X} \cap [\mathbf{b}]\right)/\mu\left[\mathbf{b}\right]$. In particular, if $\mathbb{I} \subset \mathbb{M}$ is finite, then $\mu_{\mathbb{I}}^{(\mathbf{b})} \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$ is the condition probability measure on the $\mathbb{I}$ coordinates: for any $\mathbf{c} \in \mathcal{A}^{\mathbb{I}}$, $\quad \mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{c}] \quad = \quad \mu\left([\mathbf{c}] \cap [\mathbf{b}]\right)/\mu\left[\mathbf{b}\right]$.

Subshifts: A **subshift** [5, 7] is a a closed, shift-invariant subset $\mathfrak{X} \subset \mathcal{A}^{\mathbb{M}}$. If $\mathbb{U} \subset \mathbb{M}$, then let $\mathfrak{X}_{\mathbb{U}} = \{\mathbf{x}_{\mathbb{U}} \,;\, \mathbf{x} \in \mathfrak{X}\}$ be all **admissible** $\mathbb{U}$**-blocks** in $\mathfrak{X}$. If $\mathbb{U} \subset \mathbb{M}$ is finite, and $\mathfrak{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\} \subset \mathcal{A}^{\mathbb{U}}$ is a collection of admissible blocks, then the induced **subshift of finite type** (SFT) is the largest subshift $\mathfrak{X} \subset \mathcal{A}^{\mathbb{M}}$ such that $\mathfrak{X}_{\mathbb{U}} = \mathfrak{W}$. In other words, $\mathfrak{X} = \bigcap_{\mathsf{m}\in\mathbb{M}} \boldsymbol{\sigma}^{\mathsf{m}}[\mathfrak{W}]$, where $[\mathfrak{W}] = \left\{\mathbf{a} \in \mathcal{A}^{\mathbb{M}} \,;\, \mathbf{a}_{\mathbb{U}} \in \mathfrak{W}\right\}$. A **sofic shift** is the image of an SFT under a block map.

In particular, if $\mathbb{M} = \mathbb{Z}$ and $\mathbb{U} = \{0, 1\}$, then $\mathfrak{X}$ is called **topological Markov shift**, and the **transition matrix** of $\mathfrak{X}$ is the matrix $\mathbf{P} = [p_{ab}]_{a,b\in\mathcal{A}}$, where $p_{ab} = 1$ if $[ab] \in \mathfrak{W}$, and $p_{ab} = 0$ if $[ab] \notin \mathfrak{W}$.

Characters: Let $\mathbb{T}^1 \subset \mathbb{C}$ be the circle group. A **character** of $\mathcal{A}^{\mathbb{M}}$ is a continuous homomorphism $\boldsymbol{\chi} : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathbb{T}^1$; the group of such characters is denoted $\widehat{\mathcal{A}^{\mathbb{M}}}$. For any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$ there is a finite subset $\mathbb{K} \subset \mathbb{M}$, and nontrivial $\chi_{\mathsf{k}} \in \widehat{\mathcal{A}}$ for all $\mathsf{k} \in \mathbb{K}$, so that, for any $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$, $\boldsymbol{\chi}(\mathbf{a}) = \prod_{\mathsf{k}\in\mathbb{K}} \chi_{\mathsf{k}}(a_{\mathsf{k}})$. We indicate this by writing: "$\boldsymbol{\chi} = \bigotimes_{\mathsf{k}\in\mathbb{K}} \chi_{\mathsf{k}}$". The **rank** of $\boldsymbol{\chi}$ is the cardinality of $\mathbb{K}$.

Cesàro Density: If $\ell, n \in \mathbb{Z}$, then $[\ell...n) = \{m \in \mathbb{Z} \,;\, \ell \leq m < n\}$. If $\mathbb{J} \subset \mathbb{N}$, then the **Cesàro density** of $\mathbb{J}$ is defined: $\mathsf{density}(\mathbb{J}) = \lim_{N\to\infty} \frac{1}{N} \mathsf{card}\left(\mathbb{J} \cap [0..N)\right)$. If $\mathbb{J}, \mathbb{K} \subset \mathbb{N}$, then their **relative Cesàro density** is defined:

$$\mathsf{rel\,density}\,[\mathbb{J}/\mathbb{K}] \quad = \quad \lim_{N\to\infty} \frac{\mathsf{card}\,(\mathbb{J} \cap [0..N))}{\mathsf{card}\,(\mathbb{K} \cap [0..N))}.$$

In particular, $\mathsf{density}(\mathbb{J}) = \mathsf{rel\,density}\,[\mathbb{J}/\mathbb{N}]$.

1 Harmonic Mixing of Markov Random Fields

Let $\mathbb{B} \subset \mathbb{M}$ be a finite subset, symmetric under multiplication by -1 (usually, $\mathbb{B} = \{-1, 0, 1\}^D$). For any $\mathbb{U} \subset \mathbb{M}$, we define

$$\mathsf{cl}(\mathbb{U}) = \{u + b \,;\, u \in \mathbb{U} \text{ and } b \in \mathbb{B}\} \qquad \text{and} \qquad \partial\mathbb{U} = \mathsf{cl}(\mathbb{U}) \setminus \mathbb{U}.$$

For example, if $\mathbb{M} = \mathbb{Z}$ and $\mathbb{B} = \{-1, 0, 1\}$, then $\partial\{0\} = \{\pm 1\}$.

Let $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$. Suppose $\mathbb{U} \subset \mathbb{M}$, and let $\mathbb{V} = \partial\mathbb{U}$ and $\mathbb{W} = \mathbb{M} \setminus \mathsf{cl}(\mathbb{U})$. If $\mathbf{b} \in \mathcal{A}^{\mathbb{V}}$, then we say $\mathbf{b}$ **isolates** $\mathbb{U}$ from $\mathbb{W}$ if the conditional measure $\mu^{(\mathbf{b})}$ is a product of $\mu^{(\mathbf{b})}_{\mathbb{U}}$ and $\mu^{(\mathbf{b})}_{\mathbb{W}}$. That is, for any $\mathfrak{U} \subset \mathcal{A}^{\mathbb{U}}$ and $\mathfrak{W} \subset \mathcal{A}^{\mathbb{W}}$, we have $\mu^{(\mathbf{b})}(\mathfrak{U} \cap \mathfrak{W}) = \mu^{(\mathbf{b})}_{\mathbb{U}}(\mathfrak{U}) \cdot \mu^{(\mathbf{b})}_{\mathbb{W}}(\mathfrak{W})$.

We say that μ is a **Markov random field** [4, 13] with **interaction range** $\mathbb{B}$ (or write, "μ is a $\mathbb{B}$-**MRF**") if, for any $\mathbb{U} \subset \mathbb{M}$ with $\mathbb{V} = \partial\mathbb{U}$ and $\mathbb{W} = \mathbb{M} \setminus \mathsf{cl}(\mathbb{U})$, any choice of $\mathbf{b} \in \mathcal{A}^{\mathbb{V}}$ isolates $\mathbb{U}$ from $\mathbb{W}$.

For example, if $\mathbb{M} = \mathbb{Z}$ and $\mathbb{B} = \{-1, 0, 1\}$, then μ is a $\mathbb{B}$-MRF iff μ is a (one-step) Markov chain. If $\mathbb{B} = [-N...N]$, then μ is a $\mathbb{B}$-MRF iff μ is an N-step Markov chain.

Lemma 1 *If μ is a Markov random field, then $\mathsf{supp}\,(\mu)$ is a subshift of finite type.* □

For example, if μ is a Markov chain on $\mathcal{A}^{\mathbb{Z}}$, then $\mathsf{supp}\,(\mu)$ is a topological Markov shift.

Let $\mathbb{B} \subset \mathbb{M}$, and let $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$ be $\mathbb{B}$-MRF. Let $\mathbb{S} = \mathbb{B} \setminus \{0\}$. For any $\mathbf{b} \in \mathcal{A}^{\mathbb{S}}$, let $\mu_0^{(\mathbf{b})} \in \mathcal{M}(\mathcal{A})$ be the conditional probability measure on the zeroth coordinate. We say that μ is **locally free** if, for any $\mathbf{b} \in \mathcal{A}^{\mathbb{S}}$, $\quad \mathsf{card}\left(\mathsf{supp}\left(\mu_0^{(\mathbf{b})}\right)\right) \geq 2$.

Example: If $D = 1$, then $\mathbb{B} = \{-1, 0, 1\}$, $\mathbb{S} = \{\pm 1\}$, and μ is a Markov chain. Thus, $\mathsf{supp}\,(\mu)$ is a topological Markov shift, with transition matrix $\mathbf{P} = [p_{ab}]_{a,b \in \mathcal{A}}$. For any $a, b \in \mathcal{A}$, write $a \rightsquigarrow b$ if $p_{ab} = 1$, and define the **follower** and **predecessor** sets

$$\mathcal{F}\,(a) \;=\; \{b \in \mathcal{A}\,;\; a \rightsquigarrow b\} \quad \text{and} \quad \mathcal{P}\,(b) \;=\; \{a \in \mathcal{A}\,;\; a \rightsquigarrow b\}.$$

It is easy to show that the following are equivalent:

1. μ is locally free.
2. Every entry of $\mathbf{P}^2$ is 2 or larger.
3. For any $a, b \in \mathcal{A}$, $\mathsf{card}\,(\mathcal{F}\,(a) \cap \mathcal{P}\,(b)) \geq 2$.

Recall that $\widehat{\mathcal{A}}$ is the dual group of $\mathcal{A}$. For any $\boldsymbol{\chi} \in \widehat{\mathcal{A}}$ and $\nu \in \mathcal{M}(\mathcal{A})$, let $\langle \chi, \nu \rangle = \sum_{a \in \mathcal{A}} \chi(a) \cdot \nu\{a\}$. It is easy to check:

Lemma 2 *Let p be prime and $\mathcal{A} = \mathbb{Z}_{/p}$. If μ is a locally free MRF on $\mathcal{A}^{\mathbb{M}}$, then there is some $c < 1$ so that, for all nontrivial $\chi \in \widehat{\mathcal{A}}$, and any $\mathbf{b} \in \mathcal{A}^{\mathbb{S}}$, $\quad \left|\left\langle \chi, \mu_0^{(\mathbf{b})} \right\rangle\right| \leq c.$* □

For any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$ and $\mu \in \mathcal{M}\left(\mathcal{A}^{\mathbb{M}}\right)$, define $\langle \boldsymbol{\chi}, \mu \rangle = \int_{\mathcal{A}^{\mathbb{M}}} \boldsymbol{\chi}(\mathbf{a})\, d\mu[\mathbf{a}]$. A measure μ is called **harmonically mixing** if, for any $\epsilon > 0$, there is some $R \in \mathbb{N}$ so that, for any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$,

$$\Big(\mathsf{rank}\,[\boldsymbol{\chi}] \;>\; R \Big) \Longrightarrow \Big(|\langle \boldsymbol{\chi}, \mu \rangle| \;<\; \epsilon \Big).$$

The significance of this is the following [8, Thm.12]:

Theorem: *Let $\mathcal{A} = \mathbb{Z}_{/p}$ (p prime). Any nontrivial LCA on $\mathcal{A}^{\mathbb{M}}$ asymptotically randomizes any harmonically mixing measure.* □

Most MRFs with full support are harmonically mixing [9, Thm.15]. We now extend this.

Theorem 3 *Let $\mathcal{A} = \mathbb{Z}_{/p}$ (p prime). If μ is a locally free MRF, then μ is harmonically mixing.*

Proof: Let $\mathbb{K} \subset \mathbb{M}$ be some finite set, and let $\boldsymbol{\chi} = \bigotimes_{k\in\mathbb{K}} \chi_k$ be a character of $\mathcal{A}^{\mathbb{M}}$.

A subset $\mathbb{I} \subset \mathbb{M}$ is **$\mathbb{B}$-separated** if $(i-j) \notin \mathbb{B}$ for all $i, j \in \mathbb{I}$ with $i \neq j$. **Claim 1:** *Let $K = \mathsf{card}(\mathbb{K}) = \mathsf{rank}[\boldsymbol{\chi}]$, and let $B = \max\{|b_1 - b_2| \ ; \ b_1, b_2 \in \mathbb{B}\}$. There exists a $\mathbb{B}$-separated subset $\mathbb{I} \subset \mathbb{K}$ such that*

$$\mathsf{card}(\mathbb{I}) \quad = \quad I \quad \geq \quad \frac{K}{B^D}. \tag{1}$$

Proof: Let $\widetilde{\mathbb{B}} = [0..B)^D$ be a box of sidelength B. Cover $\mathbb{K}$ with disjoint translated copies of $\widetilde{\mathbb{B}}$, so that

$$\mathbb{K} \quad \subset \quad \bigsqcup_{i\in\mathbb{I}} \left(\widetilde{\mathbb{B}} + i\right)$$

for some set $\mathbb{I} \subset \mathbb{K}$ Thus, $|i-j| \geq B$ for any $i, j \in \mathbb{I}$ with $i \neq j$, so $(i-j) \notin \mathbb{B}$. Also, $\mathsf{card}\left(\widetilde{\mathbb{B}}\right) = B^D$, so each copy covers at most B^D points in $\mathbb{K}$. Thus, we require at least $\frac{K}{B^D}$ copies to cover all of $\mathbb{K}$. In other words, $I \geq \frac{K}{B^D}$. . . . . ◇ `[Claim 1]`

Thus, $\boldsymbol{\chi} = \boldsymbol{\chi}_{\mathbb{I}} \cdot \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}$, where $\boldsymbol{\chi}_{\mathbb{I}}(\mathbf{a}) = \prod_{i\in\mathbb{I}} \chi_i(a_i)$, and $\boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{a}) = \prod_{k\in\mathbb{K}\setminus\mathbb{I}} \chi_k(a_k)$.

Let $\mathbb{J} = (\partial\,\mathbb{I}) \cup (\mathbb{K} \setminus \mathbb{I})$; fix $\mathbf{b} \in \mathcal{A}^{\mathbb{J}}$, and let $\mu_{\mathbb{I}}^{(\mathbf{b})} \in \mathcal{M}(\mathcal{A}^{\mathbb{I}})$ be the corresponding conditional probability measure. Since μ is a Markov random field, and the $\mathbb{I}$ coordinates are 'isolated' from one another by $\mathbb{J}$ coordinates, it follows that $\mu_{\mathbb{I}}^{(\mathbf{b})}$ is a product measure. In other words, for any $\mathbf{a} \in \mathcal{A}^{\mathbb{I}}$,

$$\mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{a}] \quad = \quad \prod_{i\in\mathbb{I}} \mu_i^{(\mathbf{b})}\{a_i\}. \tag{2}$$

Thus, the conditional expectation of $\boldsymbol{\chi}_{\mathbb{I}}$ is given:

$$\left\langle \boldsymbol{\chi}_{\mathbb{I}},\ \mu_{\mathbb{I}}^{(\mathbf{b})} \right\rangle \quad = \quad \sum_{\mathbf{a}\in\mathcal{A}^{\mathbb{I}}} \mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{a}] \cdot \left(\prod_{i\in\mathbb{I}} \chi_i(a_i)\right) \quad \overline{\overline{\scriptstyle \text{by}(2)}} \quad \sum_{\mathbf{a}\in\mathcal{A}^{\mathbb{I}}} \left(\prod_{i\in\mathbb{I}} \mu_i^{(\mathbf{b})}\{a_i\} \cdot \chi_i(a_i)\right)$$

$$= \quad \prod_{i\in\mathbb{I}} \left(\sum_{a_i\in\mathcal{A}} \mu^{(\mathbf{b})}\{a_i\} \cdot \chi_i(a_i)\right) \quad = \quad \prod_{i\in\mathbb{I}} \left\langle \chi_i,\ \mu_i^{(\mathbf{b})} \right\rangle.$$

Thus, $\left\langle \boldsymbol{\chi},\ \mu^{(\mathbf{b})} \right\rangle = \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b}) \cdot \left\langle \boldsymbol{\chi}_{\mathbb{I}},\ \mu^{(\mathbf{b})}_{\mathbb{I}} \right\rangle = \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b}) \cdot \prod_{i\in\mathbb{I}} \left\langle \chi_i,\ \mu^{(\mathbf{b})}_i \right\rangle$; hence, if $I = \mathsf{card}\,(\mathbb{I})$, then

$$\left|\left\langle \boldsymbol{\chi},\ \mu^{(\mathbf{b})} \right\rangle\right| \quad = \quad \left|\boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b})\right| \cdot \prod_{i\in\mathbb{I}} \left|\left\langle \chi_i,\ \mu^{(\mathbf{b})}_i \right\rangle\right| \quad \leq \quad 1 \cdot c^I \tag{3}$$

where the last step follows from Lemma 2. But $\langle \boldsymbol{\chi}, \mu \rangle = \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot \left\langle \boldsymbol{\chi},\ \mu^{(\mathbf{b})} \right\rangle$, so

$$|\langle \boldsymbol{\chi}, \mu \rangle| \ \leq \ \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot \left|\left\langle \boldsymbol{\chi},\ \mu^{(\mathbf{b})} \right\rangle\right| \underset{\text{by}(3)}{\leq} \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot c^I \ = \ c^I \underset{\text{by}(1)}{\leq} c^{K/(B^D)} \xrightarrow[K\to\infty]{} 0. \qquad \Box$$

2 The Even Shift is Not Harmonically Mixing

We will now construct a measure ν, supported on a sofic shift, which is *not* harmonically mixing. Nonetheless, we'll show in §3-§5 that this measure *is* asymptotically randomized by many LCA.

Let $\mathfrak{X} \subset \left(\mathbb{Z}_{/3}\right)^{\mathbb{Z}}$ be the subshift of finite type defined by the transition matrix

$$\mathbf{A} \quad = \quad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \text{where, } \forall i, j \in \mathbb{Z}_{/3}, \quad a_{ij} = \begin{cases} 1 & \text{if } j \rightsquigarrow i \text{ is allowed} \\ 0 & \text{if } j \rightsquigarrow i \text{ is not allowed} \end{cases}$$

Let $\Phi : \mathfrak{X} \to \left(\mathbb{Z}_{/2}\right)^{\mathbb{Z}}$ be the factor map of radius 0 which sends 0 into 0 and both 1 and 2 to 1. Then $\mathfrak{S} := \Phi(\mathfrak{X})$ is Weiss's *Even Sofic Shift*: if $\mathbf{s} \in \mathfrak{S}$, then there are an even number of 1's between any two occurrences of 0 in $\mathbf{s}$.

For any $N \in \mathbb{N}$, and $i, j \in \mathbb{Z}_{/3}$, let $\mathfrak{X}^N_{ij} = \{\mathbf{x} \in \mathfrak{X}\ ;\ x_0 = i,\ x_N = j\}$, and let:

$$\mathfrak{E}_N \ := \ \left\{ \mathbf{s} \in \mathfrak{S}\ ;\ \sum_{n=0}^{N} s_n \text{ is even} \right\}, \text{ and } \mathfrak{O}_N \ := \ \left\{ \mathbf{s} \in \mathfrak{S}\ ;\ \sum_{n=0}^{N} s_n \text{ is odd} \right\}.$$

Lemma 4 *$\forall i, j \in \mathbb{Z}_{/3}$, either $\Phi\left(\mathfrak{X}^N_{i,j}\right) \subset \mathfrak{E}_N$ or $\Phi\left(\mathfrak{X}^N_{i,j}\right) \subset \mathfrak{O}_N$. In particular,*

$$\begin{aligned} \Phi\left(\mathfrak{X}^N_{0,0} \sqcup \mathfrak{X}^N_{1,2} \sqcup \mathfrak{X}^N_{2,1} \sqcup \mathfrak{X}^N_{0,2} \sqcup \mathfrak{X}^N_{1,0}\right) \quad &= \quad \mathfrak{E}_N, \\ \textit{and} \ \ \Phi\left(\mathfrak{X}^N_{1,1} \sqcup \mathfrak{X}^N_{0,1} \sqcup \mathfrak{X}^N_{2,0} \sqcup \mathfrak{X}^N_{2,2}\right) \quad &= \quad \mathfrak{O}_N. \end{aligned}$$

Proof: Let $\mathbf{x} \in \mathfrak{X}_{ij}^N$, and $\mathbf{s} = \Phi(\mathbf{x})$. Note that, if $k < k^*$ are any two values so that $x_k = 0 = x_{k^*}$, then $\sum_{n=k}^{k^*} s_n$ is even. In particular, let k be the first element of $[0...N]$ where $x_k = 0$, and let k^* be the last element of $[0...N]$ where $x_{k^*} = 0$. Thus, $\sum_{n=k}^{k^*} s_n \equiv 0$ (mod 2), so that $\sum_{n=0}^{N} s_n \equiv \sum_{n=0}^{k-1} s_n + \sum_{n=k^*+1}^{N} s_n \pmod 2$.

But since $x_{k-1} \neq 0 \neq x_{k^*+1}$ by construction, the definition of $\mathfrak{X}$ forces $x_{k-1} = 2$ and $x_{k^*+1} = 1$. Thus the parity of $\sum_{n=0}^{k-1} s_n$ depends only on the value of $x_0 = i$. Similarly the parity of $\sum_{n=k^*+1}^{N} s_n$ depends only on $x_N = j$. □

Let $\mu \in \mathcal{M}[\mathfrak{X}]$ be a mixing Markov measure on $\mathfrak{X}$, with transition matrix $\mathbf{P}$ and Perron measure $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2) \in \mathcal{M}\left[\mathbb{Z}_{/3}\right]$. Let $\nu = \Phi\mu \in \mathcal{M}[\mathfrak{S}]$, so that if $\mathfrak{U} \subset \mathfrak{S}$ is measurable, then $\nu[\mathfrak{U}] := \mu\left[\Phi^{-1}(\mathfrak{U})\right]$

For all $N \in \mathbb{N}$, define character $\boldsymbol{\chi}_N$ by $\boldsymbol{\chi}_N(\mathbf{x}) = \prod_{n=0}^{N}(-1)^{x_n}$ for all $\mathbf{x} \in \left(\mathbb{Z}_{/2}\right)^{\mathbb{Z}}$. Then Lemma 4 implies:

$$\begin{aligned}\langle \boldsymbol{\chi}_N, \nu\rangle &= \nu(\mathfrak{E}_N) - \nu(\mathfrak{O}_N) \\ &= \mu\left(\mathfrak{X}_{0,0}^N \sqcup \mathfrak{X}_{1,2}^N \sqcup \mathfrak{X}_{2,1}^N \sqcup \mathfrak{X}_{0,2}^N \sqcup \mathfrak{X}_{1,0}^N\right) \;-\; \nu\left(\mathfrak{X}_{1,1}^N \sqcup \mathfrak{X}_{0,1}^N \sqcup \mathfrak{X}_{2,0}^N \sqcup \mathfrak{X}_{2,2}^N\right).\end{aligned}$$

But μ is mixing, so $\lim_{N\to\infty} \mu(\mathfrak{X}_{i,j}^N) = \rho_i \cdot \rho_j$. Thus, $\lim_{N\to\infty} \langle \boldsymbol{\chi}_N, \nu\rangle = \rho_0^2 + 2\rho_1\rho_2 - \rho_1^2 - \rho_2^2$. So for example if

$$\mathbf{P} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$$

with Perron measure $\boldsymbol{\rho} = \left(\frac{2}{5}, \frac{1}{5}, \frac{2}{5}\right)$ then $\lim_{N\to\infty} \langle \boldsymbol{\chi}_N, \nu\rangle \neq 0$. But clearly, $\mathsf{rank}\left[\boldsymbol{\chi}_N\right] = N$, so that $\lim_{N\to\infty} \mathsf{rank}\left[\boldsymbol{\chi}_N\right] = \infty$. Thus ν is not harmonically mixing.

3 Dispersion Mixing

The example from §2 suggests the need for an asymptotic randomization condition on measures that is less restrictive than harmonic mixing. In this section, we'll define the concepts of *dispersion mixing* (for measures) and *dispersion* (for automata) which together yield asymptotic randomization. In §4 we'll show that many LCA are dispersive. In §5 and §6 we'll show that many measures (including the Even Shift measure ν from §2) are dispersion mixing. Throughout this section, as well as §4-§6, let $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where $p \in \mathbb{N}$ is prime, and $s \in \mathbb{N}$.

Let Φ be a linear CA on $\mathcal{A}^{\mathbb{M}}$. Suppose there is a finite subset $\mathbb{F} \subset \mathbb{M}$, and coefficients $\varphi_{\mathsf{f}} \in [0..p)$ so that, for any $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$, $\Phi(\mathbf{a}) = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \boldsymbol{\sigma}^{\mathsf{f}}(\mathbf{a})$. We then write Φ as a 'polynomial of shifts': $\Phi = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \boldsymbol{\sigma}^{\mathsf{f}}$. If $\mathcal{A} = \mathbb{Z}_{/p}$, then any LCA has this form. The advantage of this 'polynomial' notation is that composition of two LCA corresponds to multiplication of their respective polynomials.

For example, suppose $\mathbb{M} = \mathbb{Z}$ and $\Phi = 1 + \boldsymbol{\sigma}$; that is, $\Phi(\mathbf{a})_0 = a_0 + a_1 \pmod{p}$. Then for any $N \in \mathbb{N}$,

$$\Phi^N = \sum_{n=0}^{N} \begin{bmatrix} N \\ n \end{bmatrix}_p \boldsymbol{\sigma}^n, \qquad \text{where} \quad \begin{bmatrix} N \\ n \end{bmatrix}_p = \binom{N}{n} \bmod p. \tag{4}$$

Let $S > 0$, and let $\mathbb{K}, \mathbb{J} \subset \mathbb{M}$ be subsets. We say that $\mathbb{K}$ and $\mathbb{J}$ are S-**separated** if

$$\min\{|\mathsf{k} - \mathsf{j}| \; ; \; \mathsf{k} \in \mathbb{K} \text{ and } \mathsf{j} \in \mathbb{J}\} \quad \geq \quad S$$

If $\mathbb{F}, \mathbb{G} \subset \mathbb{M}$, and $\Phi = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \boldsymbol{\sigma}^{\mathsf{f}}$ and $\Gamma = \sum_{\mathsf{g}\in\mathbb{G}} \gamma_{\mathsf{g}} \cdot \boldsymbol{\sigma}^{\mathsf{g}}$ are two LCA, then we say Φ and Γ are S-**separated** if $\mathbb{F}$ and $\mathbb{G}$ are S-separated. Likewise, if $\mathbb{K}, \mathbb{X} \subset \mathbb{M}$, and $\boldsymbol{\chi} = \bigotimes_{\mathsf{k}\in\mathbb{K}} \chi_{\mathsf{k}}$ and $\boldsymbol{\xi} = \bigotimes_{\mathsf{x}\in\mathbb{X}} \xi_{\mathsf{x}}$ are two characters, then we say $\boldsymbol{\chi}$ and $\boldsymbol{\xi}$ are S-**separated** if $\mathbb{K}$ and $\mathbb{X}$ are S-separated.

If $\Phi = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \boldsymbol{\sigma}^{\mathsf{f}}$ is an LCA, then let $\mathsf{rank}_S(\Phi)$ be the maximum number of S-separated LCA which can be summed to yield Φ:

$$\mathsf{rank}_S(\Phi) \quad = \quad \max\Big\{R \; ; \; \exists\, \Phi_1, \ldots, \Phi_R \text{ mutually } S\text{-separated, so that } \Phi = \Phi_1 + \cdots + \Phi_R\Big\}.$$

For example, if

$$\Phi \quad = \quad 1 \; + \; \boldsymbol{\sigma}^5 + \boldsymbol{\sigma}^6 \; + \; \boldsymbol{\sigma}^{11} + \boldsymbol{\sigma}^{12} + \boldsymbol{\sigma}^{13},$$

then $\mathsf{rank}_4(\Phi) = 3$, because $\Phi = \Phi_1 + \Phi_2 + \Phi_3$, where

$$\Phi_1 = 1, \quad \Phi_2 = \boldsymbol{\sigma}^5 + \boldsymbol{\sigma}^6, \text{ and } \Phi_3 = \boldsymbol{\sigma}^{11} + \boldsymbol{\sigma}^{12} + \boldsymbol{\sigma}^{13}.$$

On the other hand, clearly, $\mathsf{rank}_1(\Phi) = 6$, while $\mathsf{rank}_7(\Phi) = 1$.

Likewise, if $\boldsymbol{\chi} = \bigotimes_{\mathsf{k}\in\mathbb{K}} \chi_{\mathsf{k}}$ is a character, and $S > 0$, then we define

$$\mathsf{rank}_S(\chi) \quad = \quad \max\left\{R\ ;\ \exists\, \boldsymbol{\chi}_1, \ldots, \boldsymbol{\chi}_R \text{ mutually } S\text{-separated, so that } \boldsymbol{\chi} = \boldsymbol{\chi}_1 \otimes \cdots \otimes \boldsymbol{\chi}_R\right\}.$$

(In the notation of §1, $\mathsf{rank}[\boldsymbol{\chi}] = \mathsf{rank}_1(\boldsymbol{\chi})$.)

We say that μ is **dispersion mixing** (DM) if, for every $\epsilon > 0$, there exist $S, R > 0$ so that, for any character $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$,

$$\left(\ \mathsf{rank}_S(\boldsymbol{\chi}) \ > \ R\ \right) \Longrightarrow \left(\ |\langle \boldsymbol{\chi}, \mu\rangle| \ < \ \epsilon\ \right).$$

(Observe that dispersion mixing is less restrictive than harmonic mixing.)

If Φ is an LCA and $\boldsymbol{\chi}$ is a character, then $\boldsymbol{\chi} \circ \Phi$ is also a character. We say that Φ is **dispersive** if, for any $S > 0$, and any character $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$, there is a subset $\mathbb{J} \subset \mathbb{N}$ of density 1 so that $\lim_{\mathbb{J}\ni j\to\infty} \mathsf{rank}_S\left(\boldsymbol{\chi} \circ \Phi^j\right) = \infty$. It follows:

Theorem 5 *If Φ is dispersive and μ is DM, then Φ asymptotically randomizes μ.* □

Theorem 5 is an immediate consequence of an easily verified lemma:

Lemma 6 *Φ asymptotically randomizes μ if and only if, for all $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$, there is a subset $\mathbb{J} \subset \mathbb{N}$ with $\mathsf{density}(\mathbb{J}) = 1$, so that $\lim_{\mathbb{J}\ni j\to\infty} \left|\left\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\right\rangle\right| = 0$.*

Proof: See the proof of Theorem 12 in [8] □

4 Dispersion and Bipartite CA

If $\mathsf{m} = (m_1, m_2, \ldots, m_D) \in \mathbb{M}$, then let $|\mathsf{m}| = |m_1| + |m_2| + \cdots + |m_D|$. If $\Gamma = \sum_{\mathsf{g}\in\mathbb{G}} \gamma_{\mathsf{g}} \cdot \boldsymbol{\sigma}^{\mathsf{g}}$ is a linear cellular automaton, then define $\mathsf{diam}[\Gamma] = \max\{|\mathsf{g} - \mathsf{h}|\ ;\ \mathsf{g}, \mathsf{h} \in \mathbb{G}\}$.

The **centre** of Γ is the centroid of $\mathbb{G}$ (as a subset of $\mathbb{R}^n$):

$$\mathsf{centre}(\Gamma) \quad = \quad \frac{1}{\mathsf{card}(\mathbb{G})} \sum_{\mathsf{g}\in\mathbb{G}} \mathsf{g}.$$

We say Γ is *centred* if $|\mathsf{centre}(\Gamma)| < 1$. Let

$$K_p \quad = \quad \min\left\{\frac{1}{2},\ \frac{4p-7}{4p+4}\right\}. \quad \text{Thus, } K_2 = \frac{1}{12}, \quad K_3 = \frac{5}{16}, \quad \text{and} \quad K_p = \frac{1}{2}, \text{ for } p \geq 5.$$

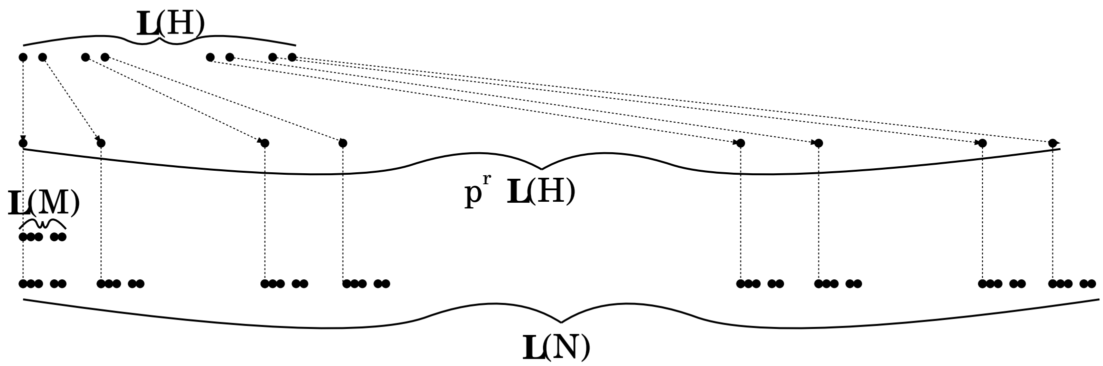

Figure 1: Lemma 9.

If Φ is an LCA, then we say Φ is **bipartite** if $\Phi = 1 + \Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}$, where Γ is centred and $\mathsf{diam}[\Gamma] \leq K_p \cdot |\mathsf{f}|$. For example:

$$\begin{array}{lllll}
\Phi & = & 1+\boldsymbol{\sigma}^{\mathsf{f}} & & \text{is bipartite for any nonzero } \mathsf{f} \in \mathbb{M} \text{ and any prime } p \in \mathbb{N}. \\
\Phi & = & 1+\boldsymbol{\sigma}^{12}+\boldsymbol{\sigma}^{13} & = \; 1+(1+\boldsymbol{\sigma})\circ\boldsymbol{\sigma}^{12} & \text{is bipartite for any prime } p \in \mathbb{N}. \\
\Phi & = & 1+\boldsymbol{\sigma}^{14}+\boldsymbol{\sigma}^{19} & = \; 1+(\boldsymbol{\sigma}^{-2}+\boldsymbol{\sigma}^{3})\circ\boldsymbol{\sigma}^{16} & \text{is bipartite for any prime } p \geq 3. \\
\Phi & = & 1+\boldsymbol{\sigma}^{2}+\boldsymbol{\sigma}^{3} & = \; 1+(1+\boldsymbol{\sigma})\circ\boldsymbol{\sigma}^{2} & \text{is bipartite for any prime } p \geq 5.
\end{array}$$

Our goal in this section is to prove:

Theorem 7 *If Φ is bipartite then Φ is dispersive.* □

For any $N \in \mathbb{N}$, let $\left[N^{(i)}|_{i=0}^{\infty}\right]$ denote the ***p*-ary expansion** of N, so that $N = \sum_{i=0}^{\infty} N^{(i)} p^i$. Let $\mathbb{L}(N) = \left\{n \in [0..N] \; ; \; n^{(i)} \leq N^{(i)}, \text{for all } i \in \mathbb{N}\right\}$.

Lemma 8 (Lucas Theorem)

(a) $\begin{bmatrix} N \\ n \end{bmatrix}_p = \prod_{i=0}^{\infty} \begin{bmatrix} N^{(i)} \\ n^{(i)} \end{bmatrix}_p$, *where we define* $\begin{bmatrix} N^{(i)} \\ n^{(i)} \end{bmatrix}_p = 0$ *if* $n^{(i)} > N^{(i)}$, *and* $\begin{bmatrix} 0 \\ 0 \end{bmatrix}_p = 1$.

(b) *Thus,* $\begin{bmatrix} N \\ n \end{bmatrix}_p \neq 0$ *iff* $n \in \mathbb{L}(N)$. □

For example, suppose $\mathbb{M} = \mathbb{Z}$ and $\Phi = 1 + \boldsymbol{\sigma}$. If we interpret equation (4) in the light of Lemma 8, we get: $\Phi^N = \sum_{n \in \mathbb{L}(N)} \begin{bmatrix} N \\ n \end{bmatrix}_p \boldsymbol{\sigma}^n$.

Lemma 9 *Let* $r,\ H\ \in\ \mathbb{N}$.

(a) *If* $M < p^r$, *and* $N = M + p^r \cdot H$, *then* $\mathbb{L}\,(N)\ =\ \mathbb{L}\,(M) + p^r \cdot \mathbb{L}\,(H)$ *(see Figure 1).*

(b) *If* $m \in \mathbb{L}\,(M)$, $h \in \mathbb{L}\,(H)$, *and* $n = m + p^r \cdot h$, *then* $\begin{bmatrix} N \\ n \end{bmatrix}_p\ =\ \begin{bmatrix} M \\ m \end{bmatrix}_p \cdot \begin{bmatrix} H \\ h \end{bmatrix}_p$. □

For example, suppose $p = 2$ and $N\ =\ 53\ =\ 5 + 48\ =\ 5 + 2^4 \cdot 3$. Then $M = 5$, $r = 4$, and $H = 3$, and

$$\begin{aligned} \mathbb{L}\,(53) &= \mathbb{L}\,(5) + 2^4 \cdot \mathbb{L}\,(3) \quad = \quad \{0,1,\ 4,5\} \quad + \quad 16 \cdot \{0,1,2,3\} \\ &= \{0,1,\ 4,5, \quad 16,17,\ 20,21, \quad 32,33,\ 37,38, \quad 48,49,\ 52,53\}. \end{aligned}$$

If $\boldsymbol{\chi}\ =\ \bigotimes_{\mathsf{k}\in\mathbb{K}} \chi_{\mathsf{k}}$ is a character, then define $\mathsf{diam}\,[\boldsymbol{\chi}]\ =\ \max\,\{|\mathsf{k}-\mathsf{j}|\ ;\ \mathsf{k},\mathsf{j} \in \mathbb{K}\}$. It follows:

Lemma 10 *Let* Φ *be an LCA, and let* $S > 0$.

(a) *If* $\boldsymbol{\chi}$ *is a character, and* $S_0 = S + \mathsf{diam}\,[\boldsymbol{\chi}]$, *then* $\mathsf{rank}_S\,(\boldsymbol{\chi} \circ \Phi)\ \geq\ \mathsf{rank}_{S_0}\,(\Phi)$.

(b) *If* Γ *is an LCA, and* $S_0 = S + \mathsf{diam}\,[\Gamma]$, *then* $\mathsf{rank}_S\,(\Gamma \circ \Phi)\ \geq\ \mathsf{rank}_{S_0}\,(\Phi)$. □

Corollary 11 Φ *is dispersive if and only if, for any* $S_0 > 0$, *there is a subset* $\mathbb{J} \subset \mathbb{N}$ *of density 1 so that* $\lim\limits_{\mathbb{J}\ni j\to\infty} \mathsf{rank}_{S_0}\left(\Phi^j\right)\ =\ \infty$. □

To prove Theorem 7, we'll use Lemma 9 to verify the condition of Corollary 11. For any $S_0 > 0$, define

$$\mathbb{J}(S_0) \quad = \quad \left\{N \in \mathbb{N}\ ;\ N\ =\ M_N + p^{r_N} H_N,\ \text{for some } H_N, r_N > 0 \text{ such that } M_N,\ S_0 < p^{r_N - 1}\right\}.$$

For example, if $p = 2$ and $S_0 = 7$, then $53 \in \mathbb{J}(7)$, because $53 = 5 + 2^4 \cdot 3$, so that $M_{53} = 5$, $r_{53} = 4$, and $H_{53} = 3$. Thus, $2^{r_{53}-1} = 2^3 = 8$, and $7 < 8$ and $5 < 8$.

Note that $53\ =\ 2^0 + 2^2 + 2^4 + 2^5$; thus, $53^{(3)} = 0$. This is exactly why $53 \in \mathbb{J}(7)$:

Lemma 12 $\mathbb{J}(S_0)\ =\ \left\{N \in \mathbb{N}\ ;\ N \geq p \cdot S_0,\ \textit{and}\ \ N^{(r)} = 0 \textit{ for some } r \in \left(\log_p(S_0) .... \log_p(N)\right]\right\}$.

Proof: Suppose $N\ =\ M_N + p^{r_N} H_N$, for some $H_N, r_N > 0$ and $M_N \geq 0$, such that $M_N,\ S_0 < p^{r_N - 1}$. Let $r = r_N - 1$; then $N^{(r)} = 0$ and $\log_p(S_0) < r < \log_p(N)$.

Conversely, suppose $N^{(r)} = 0$, where $\log_p(S_0) < r < \log_p(N)$. Let $r_N = r + 1$; then $S_0\ <\ p^r\ =\ p^{r_N - 1}$. Let $M_N\ =\ \sum_{i=0}^{r-1} N^{(i)} p^i$; then $M_N\ <\ p^r\ =\ p^{r_N - 1}$ also. Now let $H_N\ =\ \sum_{i=r_N}^{\infty} N^{(i)} p^{i - r_N}$; then $N\ =\ M_N + p^{r_N} H_N$. □

Lemma 13 $\mathsf{density}(\mathbb{J}(S_0)) = 1$.

Proof: Let $\mathbb{I} = [p \cdot S_0 ... \infty]$. Then $\mathbb{I}$ is a set of density one, and Lemma 12 implies that

$$\mathbb{I} \setminus \mathbb{J}(S_0) \quad = \quad \left\{ N \in \mathbb{I} \, ; \, N^{(r)} \neq 0 \text{ for all } r \in \left(\log_p(S_0) .... \log_p(N)\right] \right\},$$

which is a set of density zero. It follows that $\mathsf{density}(\mathbb{J}(S_0)) = \mathsf{density}(\mathbb{I}) = 1$. ∎

Lemma 14 *If $N \in \mathbb{J}(S_0)$, and $N = M + p^r H$, then $\Phi^N = \Phi^M \circ \Theta^H$, where $\Theta = \Phi^{(p^r)}$.*

Proof: Recall that $\Phi = 1 + \Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}$. Thus,

$$\begin{aligned}
\Phi^N &\underset{\text{(L)}}{=\!=} \sum_{n \in \mathbb{L}(N)} \begin{bmatrix} N \\ n \end{bmatrix}_p \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^n \underset{(9\text{b})}{=\!=\!=} \sum_{m \in \mathbb{L}(M)} \sum_{h \in \mathbb{L}(H)} \begin{bmatrix} H \\ h \end{bmatrix}_p \begin{bmatrix} M \\ m \end{bmatrix}_p \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{(m+p^r h)} \\
&= \sum_{h \in \mathbb{L}(H)} \begin{bmatrix} H \\ h \end{bmatrix}_p \left(\sum_{m \in \mathbb{L}(M)} \begin{bmatrix} M \\ m \end{bmatrix}_p \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^m \right) \circ \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{hp^r} \\
&\underset{(\dagger)}{=\!=} \sum_{h \in \mathbb{L}(H)} \begin{bmatrix} H \\ h \end{bmatrix}_p \Phi^M \circ \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h} \underset{(\star)}{=\!=} \Phi^M \circ \Theta^H.
\end{aligned}$$

(L) is by Lucas Theorem. **(9b)** is by Lemma 9(b). (†) $\Phi^M = \sum_{m \in \mathbb{L}(M)} \begin{bmatrix} M \\ m \end{bmatrix}_p \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^m$.

($\star$) $\Theta = \left(1 + \Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r} \underset{\text{(L)}}{=\!=} 1 + (\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}})^{p^r}$. Thus, $\Theta^H \underset{\text{(L)}}{=\!=} \sum_{h \in \mathbb{L}(H)} \begin{bmatrix} H \\ h \end{bmatrix}_p \left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h}$. ∎

Proof of Theorem 7: It suffices to verify the condition of Corollary 11. So, let $S_1 = S_0 + \mathsf{diam}\left[\Phi^M\right]$. Then

$$\mathsf{rank}_{S_0}\left(\Phi^N\right) \underset{\text{Lem.}14}{=\!=\!=} \mathsf{rank}_{S_0}\left(\Phi^M \circ \Theta^H\right) \underset{\text{Lem.}10(\text{b})}{\geq} \mathsf{rank}_{S_1}\left(\Theta^H\right). \tag{5}$$

Thus, we want to show that $\mathsf{rank}_{S_1}\left(\Theta^H\right) \xrightarrow[H\to\infty]{} \infty$ for H in a set of density 1. To do this, we'll use *gaps* in $\mathbb{L}(H)$. If $h_0, h_1 \in \mathbb{L}(H)$, we say that h_0 and h_1 **bracket a gap** if:

- $h_1 \geq p \cdot h_0$
- $[h_0 ... h_1) \cap \mathbb{L}(H) = \emptyset$.

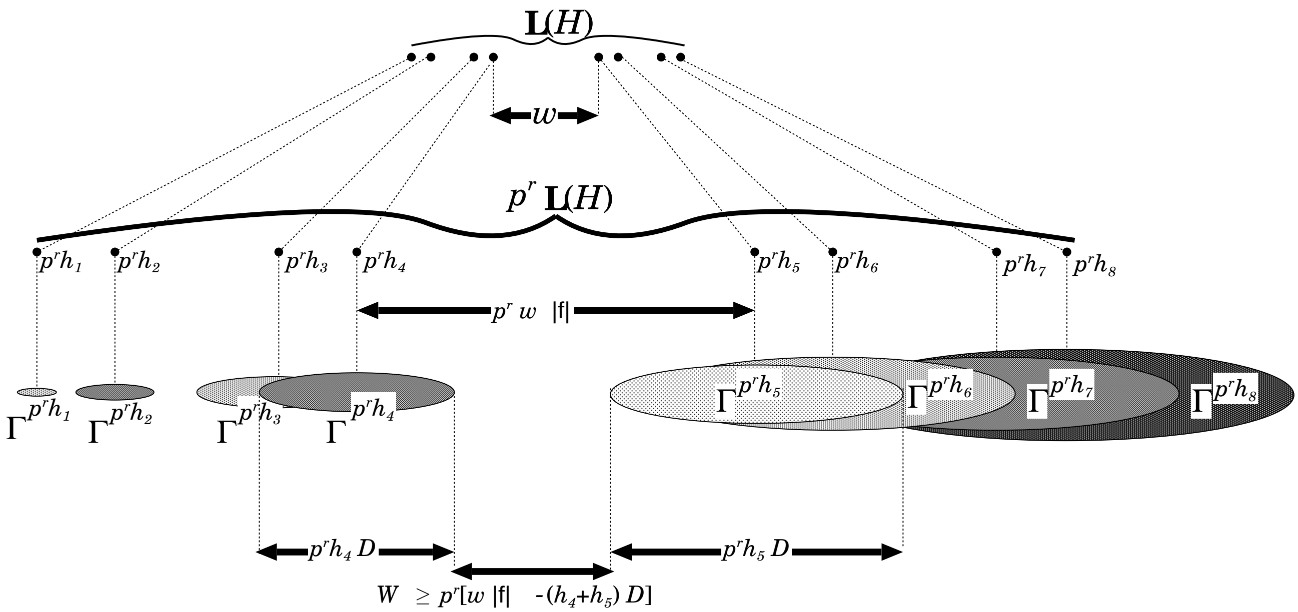

Figure 2: Claim 1. of Theorem 7.

Claim 1: *Let $h_0, h_1 \in \mathbb{L}(H)$, with $p \leq h_0 < h_1$, and suppose h_0 and h_1 bracket a gap in $\mathbb{L}(H)$. Then $\left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_0}$ and $\left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_1}$ are S_1-separated.*

Proof: Suppose $|h_0 - h_1| = w$. Then $\left(\boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_0}$ and $\left(\boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_1}$. are $(p^r \cdot w \cdot |\mathsf{f}|)$-separated. Thus, if $D = \mathsf{diam}[\Gamma]$, then $\left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_0}$ and $\left(\Gamma \circ \boldsymbol{\sigma}^{\mathsf{f}}\right)^{p^r h_1}$ are W-separated, where

$$\begin{aligned} W &= p^r w|\mathsf{f}| - \left(\mathsf{diam}\left[\Gamma^{p_r h_0}\right] + \mathsf{diam}\left[\Gamma^{p_r h_1}\right]\right) \quad = \quad p^r w|\mathsf{f}| - (p^r h_0 D + p^r h_1 D) \\ &\geq p^r \cdot \Big(w|\mathsf{f}| - D \cdot (h_1 + h_0)\Big). \end{aligned} \tag{6}$$

(see Figure 2). We want $W \geq S_1$, or, equivalently, $W - \mathsf{diam}\left[\Phi^M\right] \geq S_0$ (because $S_1 = S_0 + \mathsf{diam}\left[\Phi^M\right]$). First, note that

$$\begin{aligned} \mathsf{diam}\left[\Phi^M\right] &\leq M \cdot |\mathsf{f}| + 2 \cdot \max_{m \in \mathbb{L}(M)} \mathsf{diam}\left[\Gamma^m\right] \quad = \quad M \cdot |\mathsf{f}| + 2M \cdot D \\ &= M \cdot \Big(|\mathsf{f}| + 2D\Big) \quad \leq \quad p^{r-1} \cdot \Big(|\mathsf{f}| + 2D\Big). \end{aligned} \tag{7}$$

Thus,

$$W - \mathsf{diam}\left[\Phi^M\right] \underset{(*)}{\geq} p^r \cdot \Big(w \cdot |\mathsf{f}| - D \cdot (h_1 + h_0)\Big) - p^{r-1} \cdot \Big(|\mathsf{f}| + 2D\Big)$$

$$= \quad p^{r-1} \cdot \Big(pw \cdot |\mathsf{f}| \; - \; pD \cdot (h_1 + h_0) \; - \; |\mathsf{f}| \; - \; 2D \Big)$$

$$\underset{(\dagger)}{\geq} \quad S_0 \cdot \Big(pw \cdot |\mathsf{f}| \; - \; pD \cdot (h_1 + h_0) \; - \; |\mathsf{f}| \; - \; 2D \Big) .$$

where $(*)$ is by equations (6) and (7), and $(\dagger)$ is because $S_0 < p^{r-1}$.

Thus, it suffices to show that

$$pw \cdot |\mathsf{f}| \; - \; pD \cdot (h_1 + h_0) \; - \; |\mathsf{f}| \; - \; 2D \quad \geq \quad 1.$$

To see this, observe that

$$\begin{aligned} & pw \cdot |\mathsf{f}| \; - \; pD \cdot (h_1 + h_0) \; - \; |\mathsf{f}| \; - \; 2D \\ & \quad = \quad (pw - 1) \cdot |\mathsf{f}| - \Big[p \cdot (h_1 + h_0) - 2 \Big] \cdot D \quad \underset{(\mathrm{B})}{\geq} \quad (pw - 1) \cdot |\mathsf{f}| - \Big[p \cdot (h_1 + h_0) - 2 \Big] \cdot K_p \cdot |\mathsf{f}| \\ & \quad = \quad \Big(pw - 1 - [p \cdot (h_1 + h_0) - 2] \, K_p \Big) \cdot |\mathsf{f}| \quad \underset{(*)}{\geq} \quad p \cdot (h_1 - h_0) - 1 - [p \cdot (h_1 + h_0) - 2] \, K_p \\ & \quad = \quad p \cdot \Big((1 - K_p) \cdot h_1 - (1 + K_p) \cdot h_0 \Big) \; - \; (1 + 2 \cdot K_p) \\ & \quad \underset{(\dagger)}{\geq} \quad p \cdot \Big((1 - K_p) \cdot p - (1 + K_p) \Big) \cdot h_0 \; - \; 2 \quad \underset{(\ddagger)}{\geq} \quad p^2 \cdot \Big((1 - K_p) \cdot p - (1 + K_p) \Big) \; - \; 2 \\ & \quad \underset{(\star)}{\geq} \quad \frac{3}{4} p^2 \; - \; 2 \quad \underset{(\diamond)}{\geq} \quad 3 - 2 \quad = \quad 1. \end{aligned}$$

(B) by hypothesis that Γ is bipartite. $(*)$ because $|\mathsf{f}| \geq 1$, and $w = h_1 - h_0$.

$(\dagger)$ because $h_1 \geq p \cdot h_0$, and $K_p \leq \frac{1}{2}$. $(\ddagger)$ because $h_0 \geq p$.

$(\star)$ because $K_p \;\leq\; \frac{4p-7}{4p+4} \;=\; \frac{p - \frac{7}{4}}{p+1}$, thus, $(p+1) K_p \;\leq\; p - \frac{7}{4} \;=\; p - 1 - \frac{3}{4}$; thus, $\frac{3}{4} \;\leq\; (p-1) - (p+1) K_p \;=\; (1 - K_p) p - (1 + K_p)$.

$(\diamond)$ because $p \geq 2$, so $p^2 \geq 4$.

It follows that $W - \mathsf{diam}\left[\Phi^M\right] \geq S_0$, so that $W \geq S_1$. . . . . $\diamond$ `[Claim 1]`

Let $\mathsf{rank}\,[H] \;=\; \#$ of gaps in $\mathbb{L}\,(H)$. Then Claim 1 implies that

$$\mathsf{rank}_{S_1}\left(\Theta^H\right) \quad \geq \quad \mathsf{rank}\,[H]. \tag{8}$$

Thus, we want to show that the number of gaps is large.

Suppose $i < k$. We say that i and k **bracket a zero-block** in the p-ary expansion of H if $H^{(i-1)} \neq 0 \neq H^{(k)}$, but $H^{(j)} = 0$, for all $i \leq j < k$. For example, suppose $p = 2$ and $H = 19$. Then 3 and 5 bracket a zero block in the binary expansion ...010011.

Claim 2: *If i and k bracket a zero-block in the p-ary expansion of H, then p^i and p^j bracket a gap in $\mathbb{L}\,(H)$.*

Proof: $H^{(i)} = 0$, so the largest element in $\mathbb{L}(H)$ less than p^i is

$$h_0 \quad = \quad \sum_{j=1}^{i-1} H^{(j)} \cdot p^j \quad \leq \quad \sum_{j=1}^{i-1} (p-1) \cdot p^j \quad = \quad p^i - 1.$$

Now, $k \ = \ \min\left\{j > i \ ; \ H^{(j)} \neq 0\right\}$, so $h_1 = p^k$ is the smallest element in $\mathbb{L}(H)$ greater than p^i. Also, $h_1 \ \geq \ p^{i+1} \ > \ p \cdot (p^i - 1) \ \geq \ p \cdot h_0$. .... ◇ [Claim 2]

Let $\#\mathbf{ZB}(H) \ = \ \#$of zero-blocks in the p-ary expansion of H.

Then Claim 2 implies that that

$$\mathsf{rank}[H] \quad \geq \quad \#\mathbf{ZB}(H). \tag{9}$$

Define $\mathbb{H} \ = \ \left\{H \in \mathbb{N} \ ; \ \#\mathbf{ZB}(H) \ \geq \ \frac{1}{p^3}\log_p(H)\right\}$.

Claim 3: $\mathsf{density}(\mathbb{H}) = 1$.

Proof: Observe that $\#\mathbf{ZB}(H)$ is no less than the number of occurrences of the word "101" in the p-ary expansion of H (because 101 is a zero-block). Let

$$\mathbb{H}' \quad = \quad \left\{H \in \mathbb{N} \ ; \ (\#\text{ of occurrences of "101"}) \ \geq \ \frac{1}{p^3}\log_p(H)\right\}.$$

Then $\mathbb{H}' \subset \mathbb{H}$. The Weak Law of Large Numbers implies $\mathsf{density}(\mathbb{H}') = 1$. ◇ [Claim 3]

Define $\mathbb{J} = \ \left\{N \in \mathbb{J}(S_0) \ ; \ N = M_N + p^{r_N} H_N, \text{ where } r_N \leq \frac{1}{2}\log_p(N), \text{ and } H_N \in \mathbb{H}\right\}$.

Claim 4: $\mathsf{density}(\mathbb{J}) = 1$.

Proof: $\mathbb{J} = \mathbb{J}_1 \cap \mathbb{J}_2$, where

$$\begin{aligned} \mathbb{J}_1 \quad &= \quad \{N \in \mathbb{J}(S_0) \ ; \ N = M_N + p^{r_N} H_N, \text{ where } H_N \in \mathbb{H}\} \\ \text{and } \mathbb{J}_2 \quad &= \quad \left\{N \in \mathbb{J}(S_0) \ ; \ N = M_N + p^{r_N} H_N, \text{ where } r_N \leq \frac{1}{2}\log_p(N)\right\}. \end{aligned}$$

Now, $\mathsf{density}(\mathbb{J}_1) = 1$ by Lemma 13 and Claim 3. To see that $\mathsf{density}(\mathbb{J}_2) = 1$, note that

$$\mathbb{J}(S_0) \setminus \mathbb{J}_2 \quad \subset \quad \left\{N \in \mathbb{N} \ ; \ N^{(r)} \neq 0 \text{ for all } r \in \left(\log_p(S_0)....\frac{1}{2}\log_p(N)\right]\right\}.$$

which is a set of density zero. .... ◇ [Claim 4]

If $N = M_N + p^{r_N} H_N$ is an element of $\mathbb{J}$, then

$$\log_p(H_N) \quad \geq \quad \log_p(N) - r_N \quad \geq \quad \log_p(N) - \frac{1}{2}\log_p(N) \quad = \quad \frac{1}{2}\log_p(N). \tag{10}$$

Thus,

$$\begin{aligned} \mathsf{rank}_{S_0}\left(\Phi^N\right) \underset{\text{eqn.}(5)}{\geq} \mathsf{rank}_{S_1}\left(\Theta^{H_N}\right) \underset{\text{eqn.}(8)}{\geq} \mathsf{rank}\left[H_N\right] \underset{\text{eqn.}(9)}{\geq} \#\mathbf{ZB}\left(H_N\right) \\ \underset{(*)}{\geq} \frac{1}{p^3}\log_p(H_N) \underset{\text{eqn.}(10)}{\geq} \frac{1}{2p^3}\log_p(N). \end{aligned}$$

where $(*)$ is because $H \in \mathbb{H}$ by hypothesis.

Thus $\lim_{\mathbb{J}\ni N\to\infty} \mathsf{rank}_{S_0}\left(\Phi^N\right) \quad \geq \quad \frac{1}{2p^3} \lim_{\mathbb{J}\ni N\to\infty} \log_p(N) \; = \; \infty.$ □

5 Uniform Mixing and Dispersion Mixing

A measure $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ is **uniformly mixing** if, for any $\epsilon > 0$, there is some $M > 0$ so that, for any cylinder subsets $\mathfrak{L} \subset \mathcal{A}^{(-\infty..0]}$ and $\mathfrak{R} \subset \mathcal{A}^{[0..\infty)}$, and any $m > M$,

$$\mu\left[\boldsymbol{\sigma}^m(\mathfrak{L}) \cap \mathfrak{R}\right] \quad \underset{\epsilon}{\sim} \quad \mu\left[\mathfrak{L}\right] \cdot \mu\left[\mathfrak{R}\right] \tag{11}$$

(here, "$x \underset{\epsilon}{\sim} y$" means $|x-y| < \epsilon$.) **Example 15:**

(a) Any mixing N-step Markov chain is uniformly mixing. (See §6).

(b) If $\nu \in \mathcal{M}(\mathcal{B}^{\mathbb{Z}})$ is uniformly mixing, and $\Psi : \mathcal{B}^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ is a block map, then $\mu = \Phi(\nu)$ is also uniformly mixing. (If Ψ has local map $\psi : \mathcal{B}^{[-\ell..r]} \longrightarrow \mathcal{A}$, then replace the M in (11) with $M + \ell + r + 1$).

(c) Hence, if $\mathfrak{F} \subset \mathcal{B}^{\mathbb{Z}}$ is an SFT, and $\mathfrak{S} = \Psi(\mathfrak{F}) \subset \mathcal{A}^{\mathbb{Z}}$ a sofic shift, and $\nu \in \mathcal{M}(\mathfrak{F})$ is any mixing N-step Markov chain, then $\mu = \Phi(\nu)$ is a uniformly mixing measure on $\mathfrak{S}$. We call μ a **quasi-Markov measure**.

We say that μ is **harmonically bounded** (HB) if there is some $C < 1$ so that $|\langle \boldsymbol{\chi}, \mu \rangle| < C$ for all $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ except $\boldsymbol{\chi} = \mathbb{1}$. The goal of this section is to prove:

Theorem 16 *If μ is uniformly mixing and harmonically bounded, then μ is DM.* □

We will then apply Theorem 16 to get:

Corollary 17 *Let $\mathcal{A} = \mathbb{Z}_{/p}$ (p prime). If $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ is a mixing quasi-Markov measure, then μ is asymptotically randomized by any dispersive LCA.* □

Harmonic boundedness and entropy:

Lemma 18 *Let $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where $s \in \mathbb{N}$ and p is prime. If $h(\mu, \boldsymbol{\sigma}) \ > \ (s-1) \cdot \log_2(p)$, then μ is harmonically bounded.*

Proof: Suppose μ was not HB. Then for any $\alpha > 0$, we can find $\mathbb{1} \neq \boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ with $|\langle \boldsymbol{\chi}, \mu \rangle| \ > \ 1 - \alpha$. Let $\mathcal{I} = \mathsf{image}\,[\boldsymbol{\chi}] \subset \mathbb{T}^1$, and let $\nu = \boldsymbol{\chi}(\mu) \in \mathcal{M}(\mathcal{I})$ be the projected measure on $\mathcal{I}$. Thus, $\langle \boldsymbol{\chi}, \mu \rangle \ = \ \sum_{i \in \mathcal{I}} i \cdot \nu\{i\}$. The following four claims are easy to check.

Claim 1: *For any $\beta > 0$, there exists $\alpha > 0$ so that, for any probability measure $\nu \in \mathcal{M}(\mathcal{I})$ with $\left| \sum_{i \in \mathcal{I}} i \cdot \nu\{i\} \right| > 1 - \alpha$, there is some $i_0 \in \mathcal{I}$ with $\nu\{i_0\} > 1 - \beta$.* .... ◇

Suppose $\boldsymbol{\chi} = \bigotimes_{k \in \mathbb{K}} \chi_k$, where $\mathbb{K} \subset [0...K]$ and $K \in \mathbb{K}$. Thus, if $\boldsymbol{\xi} = \bigotimes_{k \in \mathbb{K} \setminus \{K\}} \chi_k$, then $\boldsymbol{\chi} = \boldsymbol{\xi} \otimes \chi_K$. For any $\mathbf{b} \in \mathcal{A}^{[0..K)}$, let $\mu_K^{(\mathbf{b})}$ be the conditional measure on the Kth coordinate, and let $\nu_K^{(\mathbf{b})} \ = \ \chi_K \left(\mu_K^{(\mathbf{b})} \right) \in \mathcal{M}(\mathcal{I})$ be the projected measure on $\mathcal{I}$.

Claim 2: *For any $\gamma > 0$, there exists $\beta > 0$ so that, if $\exists\, i_0 \in \mathcal{I}$ with $\nu\{i_0\} > 1 - \beta$, then there is a subset $\mathfrak{B} \subset \mathcal{A}^{[0..K)}$ with $\mu[\mathfrak{B}] > 1 - \gamma$, so that, for every $\mathbf{b} \in \mathfrak{B}$, there is some $i_{\mathbf{b}} \in \mathcal{I}$ with $\nu_K^{(\mathbf{b})}\{i_{\mathbf{b}}\} \ > \ 1 - \gamma$. Thus, if $\mathcal{P}_{\mathbf{b}} = \chi_K^{-1}\{i_{\mathbf{b}}\} \subset \mathcal{A}$, then $\mu_K^{(\mathbf{b})}[\mathcal{P}_{\mathbf{b}}] \ > \ 1 - \gamma$. (Observe that $\mathsf{card}\,(\mathcal{P}_{\mathbf{b}}) \leq p^{s-1}$ for all $\mathbf{b} \in \mathcal{A}^{[0..K)}$.)* .... ◇

For any measure $\rho \in \mathcal{M}(\mathcal{A})$, define $H(\rho) \ = \ -\sum_{a \in \mathcal{A}} \rho\{a\} \log_2 \left(\rho\{a\} \right)$. Recall (eg. [11, Prop. 5.2.12]) that the $\boldsymbol{\sigma}$-entropy of μ can be computed:

$$h(\mu, \boldsymbol{\sigma}) \quad = \quad \lim_{N \to \infty} \sum_{\mathbf{b} \in \mathcal{A}^{[0...N)}} \mu\,[\mathbf{b}] \cdot H\left(\mu_N^{(\mathbf{b})} \right) \tag{12}$$

Claim 3: *For any $\delta > 0$, there exists $\gamma_1 > 0$ so that, for any probability measure ρ on $\mathcal{A}$, if there is a subset $\mathcal{P} \subset \mathcal{A}$ with $\mathsf{card}\,(\mathcal{P}) \leq p^{s-1}$ and $\rho[\mathcal{P}] \ > \ 1 - \gamma_1$, then $H(\rho) \ < \ (s-1) \cdot \log_2(p) \ + \ \delta$.* .... ◇

Claim 4: *For any $\epsilon > 0$, and $S > 0$, there exist $\delta, \gamma_2 > 0$ so that, for any $K \in \mathbb{N}$ and probability measure μ on $\mathcal{A}^{[0..K]}$, if there is a subset $\mathfrak{B} \subset \mathcal{A}^{[0...K)}$ with $\mu[\mathfrak{B}] > 1 - \gamma_2$, such that, for all $\mathbf{b} \in \mathfrak{B}$, $H\left(\mu_K^{(\mathbf{b})} \right) < S - \delta$, then $\sum_{\mathbf{b} \in \mathcal{A}^{[0...K)}} \mu\,[\mathbf{b}] \cdot H\left(\mu_K^{(\mathbf{b})} \right) \ < \ S - \epsilon$.* ...... ◇

Now, set $S = (s-1) \cdot \log_2(p)$. For any $\epsilon > 0$, find $\delta, \gamma_2 > 0$ as in Claim 4. Then find $\gamma_1 > 0$ as in Claim 3, and let $\gamma = \min\{\gamma_1, \gamma_2\}$. Next, find β as in Claim 2 and then find α

as in Claim 1. Finally, find $\chi \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ with $|\langle \chi, \mu \rangle| \; > \; 1-\alpha$. It then follows from Claims 1-4 that $\displaystyle\sum_{\mathbf{b}\in\mathcal{A}^{[0\ldots K)}} \mu[\mathbf{b}] \cdot H\left(\mu_N^{(\mathbf{b})}\right) \; < \; (s-1)\cdot \log_2(p) - \epsilon$. But the limit in (12) is a decreasing limit, so we conclude that $h(\mu, \boldsymbol{\sigma}) \; < \; (s-1)\cdot\log_2(p) - \epsilon$. Since this is true for any $\epsilon > 0$, we conclude that $h(\mu,\boldsymbol{\sigma}) \; \leq \; (s-1)\cdot \log_2(p)$, contradicting our hypothesis. □

Corollary 19 *If $\mathcal{A} = \mathbb{Z}_{/p}$ (p prime), and $h(\mu,\boldsymbol{\sigma}) > 0$, then μ is harmonically bounded.* □

Say μ is **uniformly multiply mixing** if, for any $\epsilon > 0$, there is some $S > 0$ so that, for any $R > 0$, if $\mathbb{K}_0, \mathbb{K}_1, \ldots, \mathbb{K}_R \subset \mathbb{M}$ are finite, mutually S-separated subsets of $\mathbb{M}$, and $\mathfrak{U}_0 \subset \mathcal{A}^{\mathbb{K}_0}, \ldots, \mathfrak{U}_R \subset \mathcal{A}^{\mathbb{K}_R}$ are cylinder sets, then:

$$\mu\left(\bigcap_{r=0}^{R} \mathfrak{U}_r\right) \quad \underset{R\epsilon}{\widetilde{\quad}} \quad \prod_{r=0}^{R} \mu(\mathfrak{U}_r). \tag{13}$$

Lemma 20 *If μ is uniformly mixing, then μ is uniformly multiply mixing.*

Proof: (*by induction on R*). The case $R = 1$ is just uniform mixing. Suppose (13) is true for all $R' < R$. Find $S > 0$ so that, if $\mathbb{K}_0, \ldots, \mathbb{K}_R$ are mutually S-separated, then

$$\mu\left(\bigcap_{r=0}^{R} \mathfrak{U}_r\right) \;=\; \mu\left(\mathfrak{U}_0 \cap \bigcap_{r=1}^{R} \mathfrak{U}_r\right) \;\underset{\epsilon}{\sim}\; \mu(\mathfrak{U}_0)\cdot\mu\left(\bigcap_{r=1}^{R}\mathfrak{U}_r\right) \;\underset{(R-1)\epsilon}{\widetilde{\quad\quad}}\; \mu(\mathfrak{U}_0)\cdot\prod_{r=1}^{R}\mu(\mathfrak{U}_r),$$

where "$\underset{\epsilon}{\sim}$" comes by setting $R' = 1$, and "$\underset{(R-1)\epsilon}{\widetilde{\quad\quad}}$" comes by setting $R' = R - 1$. □

Lemma 21 *Suppose μ is uniformly multiply mixing. For any $\epsilon > 0$ and $R \in \mathbb{N}$, there is some $S > 0$ so that: if $\mathbb{K}_0, \ldots, \mathbb{K}_R \subset \mathbb{Z}$ are S-separated sets, and, for all $r \in [0..R]$, $\chi_r : \mathcal{A}^{\mathbb{K}_r} \longrightarrow \mathbb{C}$ are characters, and $\chi = \displaystyle\prod_{r=0}^{R} \chi_r$, then $\langle \chi, \mu \rangle \underset{\epsilon/2}{\widetilde{\quad}} \displaystyle\prod_{r=0}^{R} \langle \chi_r, \mu \rangle$.* □

Proof of Theorem 16: Let $\epsilon > 0$. We want to find $S > 0$ and $R > 0$ so that, if χ is any character, and $\mathsf{rank}_S(\chi) > R$, then $|\langle \chi, \mu\rangle| < \epsilon$.

Let $C < 1$ be the harmonic bound. Find $R \in \mathbb{N}$ so that $C^R < \epsilon/2$. Let $S > 0$ be as in Lemma 21. Suppose $\mathsf{rank}_S(\chi) > R$, and let $\chi = \displaystyle\bigotimes_{r=0}^{R} \chi_r$, where $\chi_r : \mathcal{A}^{\mathbb{K}_r} \longrightarrow \mathbb{C}$ are characters, and $\mathbb{K}_0, \ldots, \mathbb{K}_R \subset \mathbb{Z}$ are S-separated. Then Lemma 21 implies:

$$\langle \chi, \mu \rangle \quad \underset{\epsilon/2}{\widetilde{\quad}} \quad \prod_{r=0}^{R} \langle \chi_r, \mu \rangle. \tag{14}$$

By harmonic boundedness, we know $|\langle \boldsymbol{\chi}_r,\ \mu\rangle| < C$ for all $r \in [0..R]$. Thus, (14) implies:

$$|\langle \boldsymbol{\chi},\ \mu\rangle| \quad \underset{\epsilon/2}{\widetilde{\quad}} \quad \prod_{r=0}^{R} |\langle \boldsymbol{\chi}_r,\ \mu\rangle| \quad < \quad \prod_{r=0}^{R} C \quad = \quad C^{R+1} \quad < \quad C^{R} \quad < \quad \epsilon/2. \quad \Box$$

Proof of Corollary 17: From examples (15a) and (15b), we know μ is uniformly mixing. Any mixing quasi-Markov measure has nonzero entropy, so Corollary 19 says that μ is harmonically bounded. Theorem 16 says μ is dispersion mixing. Theorem 5 says μ is asymptotically randomized by any dispersive CA. $\Box$

6 Markov Words

If $m, n \in \mathbb{Z}$, let $\mathcal{A}^{[m...n)}$ be the set of all **words** of the form $\mathbf{a} = [a_m,\ a_{m+1}, \ldots, a_{n-1}]$. Let $\mathcal{A}^* = \bigcup_{-\infty<m<n<\infty} \mathcal{A}^{[m...n)}$ be the set of all finite words. Elements of $\mathcal{A}^*$ are denoted by boldfaced letters (eg. $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$), and subsets by gothic letters (eg. $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$). Concatenation of words is indicated by juxtaposition. Thus, if $\mathbf{a} = [a_0 \ldots a_n]$ and $\mathbf{b} = [b_0 \ldots b_m]$, then $\mathbf{ab} = [a_0 \ldots a_n b_0 \ldots b_m]$.

If $V > 0$ and $\mathbf{v} \in \mathcal{A}^{[-V..V)}$, we say that $\mathbf{v}$ is a **Markov word** for μ if (in the terminology of §1), $\mathbf{v}$ isolates $(-\infty... -V)$ from $[V...\infty)$.

Example 22:

(a) If μ is an N-step Markov shift, and $N \leq 2V$, then every $\mathbf{v} \in \mathcal{A}^{[-V..V)}$ is a Markov word.

(b) Let $\mathfrak{F} \subset \mathcal{B}^{\mathbb{Z}}$ be a subshift of finite type, let $\Psi : \mathfrak{F} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ be a block map, so that $\mathfrak{S} = \Psi(\mathfrak{F})$ is a sofic shift. Let ν be a Markov measure on $\mathfrak{F}$ and let $\mu = \Psi(\nu)$. If $s \in \mathfrak{S}_{[-V..V]}$ is a synchronizing word for Ψ, then s is a Markov word for μ.

Proposition 23 *If μ is mixing and has a Markov word, then μ is uniformly mixing.*

Proof: Fix $\epsilon > 0$. For any words $\mathbf{a}, \mathbf{b} \in \mathcal{A}^*$, the mixing of μ implies that there is some $M_\epsilon(\mathbf{a}, \mathbf{b}) < \infty$ so that, for all $m > M_\epsilon(\mathbf{a}, \mathbf{b})$, $\mu(\boldsymbol{\sigma}^m[\mathbf{a}] \cap [\mathbf{b}]) \underset{\epsilon}{\sim} \mu[\mathbf{a}] \cdot \mu[\mathbf{b}]$. Our goal is to find some $M > 0$ so that $M_\epsilon(\mathbf{a}, \mathbf{b}) < M$ for all $\mathbf{a}, \mathbf{b} \in \mathcal{A}^*$.

Let $\mathbf{v} \in \mathcal{A}^*$ be a Markov word for μ.

Claim 1: *Let $\mathbf{u}, \mathbf{w}, \mathbf{u}', \mathbf{w}' \in \mathcal{A}^*$, and consider the words $\mathbf{uvw}$ and $\mathbf{u}'\mathbf{v}\mathbf{w}'$. We have:* $M_\epsilon(\mathbf{uvw},\ \mathbf{u}'\mathbf{v}\mathbf{w}') = M_\epsilon(\mathbf{vw},\ \mathbf{u}'\mathbf{v})$.

Proof: Define transition probabilities: $\mu(\mathbf{u} \dashleftarrow \mathbf{v}) = \mu(\mathbf{uv})/\mu(\mathbf{v})$ and $\mu(\mathbf{v} \dashrightarrow \mathbf{w}) = \mu(\mathbf{vw})/\mu(\mathbf{v})$. If $m > M_\epsilon(\mathbf{vw},\ \mathbf{u}'\mathbf{v})$, then

$$\begin{aligned}
\mu\left(\boldsymbol{\sigma}^m[\mathbf{uvw}] \cap [\mathbf{u}'\mathbf{vw}']\right) &= \mu(\mathbf{u} \dashleftarrow \mathbf{v}) \cdot \mu\left(\boldsymbol{\sigma}^m[\mathbf{vw}] \cap [\mathbf{u}'\mathbf{v}]\right) \cdot \mu(\mathbf{v} \dashrightarrow \mathbf{w}') & (15)\\
&\underset{\epsilon}{\sim} \mu(\mathbf{u} \dashleftarrow \mathbf{v}) \cdot \mu[\mathbf{vw}] \cdot \mu[\mathbf{u}'\mathbf{v}] \cdot \mu(\mathbf{v} \dashrightarrow \mathbf{w}') & (16)\\
&= \mu[\mathbf{uvw}] \cdot \mu[\mathbf{u}'\mathbf{vw}']. & (17)
\end{aligned}$$

(15) and (17) are because $\mathbf{v}$ is a Markov word; (16) is because $m > M_\epsilon(\mathbf{vw}, \mathbf{u}'\mathbf{v})$. ◇ `[Claim 1]`

If $\mathbf{a} \in \mathcal{A}^*$, we say that $\mathbf{v}$ **occurs** in $\mathbf{a}$ if $\mathbf{a}\big|_{[n-V\ldots n+V)} = \mathbf{v}$ for some n. **Claim 2:** *There is some $N > 0$ so that $\mu\{\mathbf{a} \in \mathcal{A}^{[0..N]}\ ;\ \mathbf{v}$ occurs in $\mathbf{a}\} > 1 - \epsilon$.*

Proof: By ergodicity, find N so that $\mu\left(\bigcup_{n=0}^{N} \boldsymbol{\sigma}^n[\mathbf{v}]\right) > 1 - \epsilon$. .... ◇ `[Claim 2]`

Let $\mathcal{A}^*_{\mathbf{v}}$ be the set of words (of length at least N) in $\mathcal{A}^*$ with $\mathbf{v}$ occuring in the last $(N+V)$ coordinates, and let ${}_{\mathbf{v}}\mathcal{A}^*$ be the set of all words in $\mathcal{A}^*$ with $\mathbf{v}$ occuring in the first $(N+V)$ coordinates. Then Claim 2 implies that:

$$\mu(\mathcal{A}^*_{\mathbf{v}}) \quad > \quad 1-\epsilon \quad \text{and} \quad \mu({}_{\mathbf{v}}\mathcal{A}^*) \quad > \quad 1-\epsilon. \tag{18}$$

Let $\mathcal{A}^{<N} = \bigcup_{n=1}^{N} \mathcal{A}^{[0..n]}$. Then

$$\begin{aligned}
\mathcal{A}^*_{\mathbf{v}} &= \{\mathbf{uvw}\ ;\ \mathbf{u} \in \mathcal{A}^* \text{ and } \mathbf{w} \in \mathcal{A}^{<N}\}.\\
\text{and } {}_{\mathbf{v}}\mathcal{A}^* &= \{\mathbf{u}'\mathbf{vw}'\ ;\ \mathbf{u}' \in \mathcal{A}^{<N} \text{ and } \mathbf{w}' \in \mathcal{A}^*\},
\end{aligned} \tag{19}$$

Define

$$\begin{aligned}
M_1 &:= \max_{\mathbf{a}\in\mathcal{A}^*_{\mathbf{v}}} \max_{\mathbf{b}\in{}_{\mathbf{v}}\mathcal{A}^*} M_\epsilon(\mathbf{a},\mathbf{b}) \underset{(e19)}{=\!=} \max_{\substack{\mathbf{u}\in\mathcal{A}^*\\ \mathbf{w}\in\mathcal{A}^{<N}}} \max_{\substack{\mathbf{u}'\in\mathcal{A}^{<N}\\ \mathbf{w}'\in\mathcal{A}^*}} M_\epsilon(\mathbf{uvw},\ \mathbf{u}'\mathbf{vw}')\\
&\underset{(c1)}{=\!=} \max_{\mathbf{w},\mathbf{u}'\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{vw},\ \mathbf{u}'\mathbf{v}).
\end{aligned}$$

where **(e19)** is by eqn. 19 and **(c1)** is by Claim 1. Likewise, define

$$\begin{aligned}
M_2 &= \max_{\mathbf{a}\in\mathcal{A}^*_{\mathbf{v}}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\mathbf{b}) = \max_{\mathbf{w}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{vw},\ \mathbf{b}),\\
M_3 &= \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in{}_{\mathbf{v}}\mathcal{A}^*} M_\epsilon(\mathbf{a},\mathbf{b}) = \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{u}'\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\ \mathbf{u}'\mathbf{v}),\\
\text{and } M_4 &= \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\mathbf{b}).
\end{aligned}$$

Thus, $M_1, \ldots, M_4$ each maximizes a finite collection of finite values, so each is finite. Thus, $M = \max\{M_1, \ldots, M_4\}$ is finite.

Claim 3: *For any* $\mathbf{a}, \mathbf{b} \in \mathcal{A}^*$, $\quad M_\epsilon(\mathbf{a}, \mathbf{b}) < M$.

Proof: If $\mathbf{a} \in \mathcal{A}^{<N} \cup \mathcal{A}^*_{\mathbf{v}}$ and $\mathbf{b} \in \mathcal{A}^{<N} \cup {}_{\mathbf{v}}\mathcal{A}^*$, then $M_\epsilon(\mathbf{a}, \mathbf{b}) < M$ by definition.

So, suppose $\mathbf{a} \notin \mathcal{A}^{<N} \cup \mathcal{A}^*_{\mathbf{v}}$. Then eqn. (18) implies that $\mu[\mathbf{a}] < \epsilon$. Hence, for any $m \in \mathbb{N}$, $\mu(\boldsymbol{\sigma}^m[\mathbf{a}] \cap \mathbf{b}) < \epsilon$ and $\mu[\mathbf{a}] \cdot \mu[\mathbf{b}] < \epsilon$. Thus, $\mu(\boldsymbol{\sigma}^m[\mathbf{a}] \cap \mathbf{b}) \underset{\epsilon}{\sim} \mu[\mathbf{a}] \cdot \mu[\mathbf{b}]$ automatically. Hence, $M_\epsilon(\mathbf{a}, \mathbf{b}) = 0 < M$.

Likewise, if $\mathbf{b} \notin \mathcal{A}^{<N} \cup {}_{\mathbf{v}}\mathcal{A}^*$, then $M_\epsilon(\mathbf{a}, \mathbf{b}) = 0 < M$. .... ◇ `[Claim 3]`

Thus, μ is uniformly mixing. □

Corollary 24 *If μ is harmonically bounded, mixing and has a Markov word, then μ is asymptotically randomized by $\Phi = 1 + \boldsymbol{\sigma}$.*

Proof: Combine Proposition 23 with Theorems 5 and 16. □

7 Lucas Mixing

Throughout this section, let $D = 1$, so that $\mathbb{M} = \mathbb{Z}$. Let $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where $p \in \mathbb{N}$ is prime, and $s \in \mathbb{N}$. Let $\Phi = 1 + \boldsymbol{\sigma}$. We will introduce a condition on μ which is weaker than dispersion mixing, and which is both sufficient and *necessary* for asymptotic randomization.

Let $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$, and suppose $\boldsymbol{\chi} = \bigotimes_{\mathsf{k} \in \mathbb{K}} \chi_{\mathsf{k}}$. We define $|[\boldsymbol{\chi}]| = \max(\mathbb{K}) - \min(\mathbb{K})$, and define

$$\langle\!\langle \boldsymbol{\chi} \rangle\!\rangle \quad = \quad p^r, \quad \text{where} \quad r \; = \; \left\lceil \log_p |[\boldsymbol{\chi}]| \right\rceil.$$

It follows from Lucas' Theorem that $\Phi^{\langle\!\langle \boldsymbol{\chi} \rangle\!\rangle} = 1 + \boldsymbol{\sigma}^{\langle\!\langle \boldsymbol{\chi} \rangle\!\rangle}$. Thus, for any $h \in \mathbb{N}$,

$$\Phi^{h \cdot \langle\!\langle \boldsymbol{\chi} \rangle\!\rangle} \;=\; \sum_{\ell \in \mathbb{L}(h)} \begin{bmatrix} h \\ \ell \end{bmatrix}_p \boldsymbol{\sigma}^{\langle\!\langle \boldsymbol{\chi} \rangle\!\rangle \cdot \ell}, \quad \text{and thus,} \quad \boldsymbol{\chi} \circ \Phi^{h \cdot \langle\!\langle \boldsymbol{\chi} \rangle\!\rangle} \;=\; \bigotimes_{\ell \in \mathbb{L}(h)} \begin{bmatrix} h \\ \ell \end{bmatrix}_p \boldsymbol{\chi} \circ \boldsymbol{\sigma}^{\langle\!\langle \boldsymbol{\chi} \rangle\!\rangle \cdot \ell}.$$

Observe that $\mathbb{K} + p^r \ell$ and $\mathbb{K} + p^r \ell'$ are disjoint for any $\ell \neq \ell' \in \mathbb{L}(h)$. Hence, if $L = \mathsf{card}(\mathbb{L}(h))$, then $\boldsymbol{\chi} \circ \Phi^{h \cdot \langle\!\langle \boldsymbol{\chi} \rangle\!\rangle}$ is a product of L 'disjoint translates' of $\boldsymbol{\chi}$.

If μ is a measure on $\mathcal{A}^{\mathbb{Z}}$, we say that μ is **Lucas mixing** if, for any nontrivial character $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$, there is a subset $\mathbb{H} \subset \mathbb{N}$ of Cesàro density one so that

$$\lim_{\mathbb{H} \ni h \to \infty} \left\langle \boldsymbol{\chi} \circ \Phi^{h \cdot \langle\!\langle \boldsymbol{\chi} \rangle\!\rangle}, \; \mu \right\rangle \quad = \quad 0.$$

Our goal in this section is to prove:

Theorem 25 $\Big($ $\Phi = 1+\boldsymbol{\sigma}$ *asymptotically randomizes* μ $\Big) \iff \Big($ μ *is Lucas mixing* $\Big)$. □

It is relatively easy to see that:

Lemma 26 *If μ is dispersion-mixing, then μ is Lucas mixing.* □

Thus, the "$\Longleftarrow$" direction of Theorem 25 is slight generalization of Theorem 5, in the case when $\Phi = 1+\boldsymbol{\sigma}$. The "$\Longrightarrow$" direction makes this the maximum possible generalization for this automaton.

Set $S = |[\boldsymbol{\chi}]|$, and let $\widetilde{\mathbb{J}} = \mathbb{J}(S)$, where $\mathbb{J}(S)$ is defined as in §4. It follows from Lemma 13 that $\mathsf{density}\left(\widetilde{\mathbb{J}}\right) = 1$. For any $m \in \mathbb{N}$, let $\boldsymbol{\chi}^m = \boldsymbol{\chi} \circ \Phi^m$.

Lemma 27 *Let $j \in \widetilde{\mathbb{J}}$, with $j = m + p^r \cdot h$. Then $\boldsymbol{\chi} \circ \Phi^j = \boldsymbol{\chi}^m \circ \Phi^{h' \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle}$, where $h' = p^s \cdot h$ for some $s \geq 0$.*

Proof: Apply Lemma 14 to observe that $\Phi^j = \Phi^m \circ \Phi^{h \cdot (p^r)}$. Thus,

$$\boldsymbol{\chi} \circ \Phi^j \quad = \quad \boldsymbol{\chi} \circ \Phi^m \circ \Phi^{h \cdot (p^r)} \quad = \quad \boldsymbol{\chi}^m \circ \Phi^{h \cdot (p^r)}.$$

By definition, r is such that $m < p^{r-1}$ and $|[\boldsymbol{\chi}]| < p^{r-1}$. Thus,

$$|[\boldsymbol{\chi}^m]| \quad = \quad |[\boldsymbol{\chi}]| + m \quad < \quad p^{r-1} + p^{r-1} \quad \leq \quad p^r.$$

Now, let $s = r - \log_p |[\boldsymbol{\chi}^m]|$, and let $h' = p^s \cdot h$. Then $h \cdot (p^r) = h' \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle$, so that $\Phi^{h \cdot (p^r)} = \Phi^{h' \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle}$. □

Proof of Theorem 25: We will use Lemma 6.

'$\Longleftarrow$' For any $m \in \mathbb{N}$, let $r(m) = \left\lceil \log_p \left(\max\left\{ m,\ |[\boldsymbol{\chi}]| \right\} \right) \right\rceil + 1$, and define

$$\widetilde{\mathbb{J}}_m \quad = \quad \left\{ m + p^{r(m)} h \ ;\ h \in \mathbb{N} \right\}. \tag{20}$$

It follows that:

$$\widetilde{\mathbb{J}} \quad = \quad \bigcup_{m \in \mathbb{N}} \widetilde{\mathbb{J}}_m. \tag{21}$$

If $j = m + p^{r(m)} h$ is an element of $\widetilde{\mathbb{J}}_m$, then Lemma 27 says $\boldsymbol{\chi} \circ \Phi^j = \boldsymbol{\chi}^m \circ \Phi^{h' \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle}$, for some $h' \geq h$. Now, μ is Lucas mixing, so find a subset $\widetilde{\mathbb{H}}_m \subset \mathbb{N}$ of density one with $\displaystyle\lim_{\widetilde{\mathbb{H}}_m \ni h \to \infty} \left\langle \boldsymbol{\chi}^m \circ \Phi^{h \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle},\ \mu \right\rangle = 0$. Define:

$$\begin{aligned} \mathbb{H}_m &= \left\{ h \in \widetilde{\mathbb{H}}_m \ ;\ \left| \left\langle \boldsymbol{\chi}^m \circ \Phi^{h \cdot \langle\!\langle \boldsymbol{\chi}^m \rangle\!\rangle},\ \mu \right\rangle \right| \ \leq\ \frac{1}{m} \right\}, \\ \mathbb{J}_m &= \left\{ m + p^{r(m)} h \ ;\ h \in \mathbb{H}_m \right\}, \qquad (22) \\ \text{and} \quad \mathbb{J} &= \bigcup_{m \in \mathbb{N}} \mathbb{J}_m. \qquad (23) \end{aligned}$$

Claim 1: $\mathsf{density}\,(\mathbb{J}) = 1$.

Proof: For any $m \in \mathbb{N}$, there is some K so that $\mathbb{H}_m = \widetilde{\mathbb{H}}_m \cap [K...\infty)$. Thus, $\mathsf{rel\,density}\left[\mathbb{H}_m/\widetilde{\mathbb{H}}_m\right] = 1$. Thus, $\mathsf{density}(\mathbb{H}_m) = \mathsf{density}\left(\widetilde{\mathbb{H}}_m\right) = 1$. Compare (20) and (22) to see that $\mathsf{rel\,density}\left[\mathbb{J}_m/\widetilde{\mathbb{J}}_m\right] = 1$. Then compare (21) and (23) to see that $\mathsf{rel\,density}\left[\mathbb{J}/\widetilde{\mathbb{J}}\right] = 1$. Thus, $\mathsf{density}(\mathbb{J}) = \mathsf{density}\left(\widetilde{\mathbb{J}}\right) = 1$. .... ◇ `[Claim 1]`

Claim 2: $\lim_{\mathbb{J}\ni j\to\infty} \left\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\right\rangle = 0$.

Proof: Fix $\epsilon > 0$. Let M be large enough that $\frac{1}{M} < \epsilon$. For all $m \in \mathbb{N}$ with $m < M$, find H_m so that, if $h \in \widetilde{\mathbb{H}}_m$ and $h > H_m$, then $\left|\left\langle \boldsymbol{\chi}^m \circ \Phi^{h\cdot\langle\!\langle \boldsymbol{\chi}^m\rangle\!\rangle},\ \mu\right\rangle\right| < \epsilon$. Let $J_m = m + 2^{r(m)} \cdot H_m$. Thus, if $j = m + 2^{r(m)} \cdot h$ is an element of $\mathbb{J}_m$, and $j > J_m$, then we must have $h > H_m$, so that $|\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\rangle| = \left|\left\langle \boldsymbol{\chi}^m \circ \Phi^{h\cdot\langle\!\langle \boldsymbol{\chi}^m\rangle\!\rangle},\ \mu\right\rangle\right| < \epsilon$.

Now let $J = \max_{1\le m\le M} J_m$. Thus, for all $j \in \mathbb{J}$, if $j > J$, then either $j \in \mathbb{J}_m$ for some $m \le M$, in which case $|\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\rangle| < \epsilon$ by construction of J, or $j \in \mathbb{J}_m$ for some $m > M$, in which case

$$\left|\left\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\right\rangle\right| \underset{(a)}{<} \frac{1}{m} < \frac{1}{M} \underset{(b)}{<} \epsilon.$$

Here, (**a**) follows by definition of $\mathbb{H}_m$, and (**b**) follows by definition of M. ◇ `[Claim 2]`

Lemma 6 and Claims 1 and 2 imply that Φ asymptotically randomizes μ.

'$\Longrightarrow$' Suppose μ was not weakly harmonically mixing. Thus, there is some $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ and some subset $\mathbb{H} \subset \mathbb{N}$ of density $\delta > 0$ so that $\limsup_{\mathbb{H}\ni h\to\infty} \left|\left\langle \boldsymbol{\chi} \circ \Phi^{h\cdot\langle\!\langle\boldsymbol{\chi}\rangle\!\rangle},\ \mu\right\rangle\right| > 0$. But $\boldsymbol{\chi} \circ \Phi^{h\cdot\langle\!\langle\boldsymbol{\chi}\rangle\!\rangle} = \boldsymbol{\chi} \circ \Phi^{p^r\cdot h}$ (where $r = \left\lceil \log_p |[\boldsymbol{\chi}]| \right\rceil$). Hence, if $\mathbb{J} = p^r \cdot \mathbb{H}$, then $\mathsf{density}(\mathbb{J}) = p^{-r} \cdot \delta > 0$, and $\limsup_{\mathbb{J}\ni j\to\infty} \left|\left\langle \boldsymbol{\chi} \circ \Phi^j,\ \mu\right\rangle\right| = \limsup_{\mathbb{H}\ni h\to\infty} \left|\left\langle \boldsymbol{\chi} \circ \Phi^{h\cdot\langle\!\langle\boldsymbol{\chi}\rangle\!\rangle},\ \mu\right\rangle\right| > 0$. But then Lemma 6 implies that Φ cannot randomize μ. □

Conclusion

We have shown that many probability measures supported on sofic shifts are asymptotically randomized by certain linear cellular automata. This includes locally free Markov random fields on $\mathcal{A}^{\mathbb{Z}^D}$, and uniformly mixing measures on $\mathcal{A}^{\mathbb{Z}}$ having large enough entropy.

Many questions remain, however. If $\mathfrak{G} \subset \mathcal{A}^{\mathbb{Z}}$ is a subgroup shift, and $\Phi(\mathfrak{G}) \subset \mathfrak{G}$, then we expect an initial measure $\mu \in \mathcal{M}(\mathfrak{G})$ to asymptotically randomize to the Haar measure on $\mathfrak{G}$. If $\mathcal{A}$ is a *nonabelian* group, then little is known about asymptotic randomization, aside from some structure-theoretic results [12]. For arbitrary (nonlinear) permutative CA, virtually nothing is known. Is there a condition analogous to harmonic mixing or dispersion mixing for *nonlinear* permutative cellular automata?

References

[1] Alejandro Maass and Servet Martínez. On Cesàro limit distribution of a class of permutative cellular automata. *Journal of Statistical Physics*, 90(1-2):435–452, 1998.

[2] Alejandro Maass and Servet Martínez. Time averages for some classes of expansive one-dimensional cellular automata. In Eric Goles and Servet Martinez, editors, *Cellular Automata & Complex Systems*, pages 37–54. Kluwer Academic Publishers, Dordrecht, 1999.

[3] Servet Martinez Alejandro Maass, Bernard Host. Uniform Bernoulli measure in dynamics of permutative cellular automata with algebraic local rules. to appear in *Discrete & Continuous Dyn. Sys.*, 2002.

[4] Pierre Brémaud. *Markov Chains: Gibbs fields, Monte Carlo Simulation, and Queues*. Springer-Verlag, 1999.

[5] Bruce Kitchens. *Symbolic dynamics : one-sided, two-sided, and countable state Markov shifts*. Springer-Verlag, New York, 1998.

[6] D. Lind. Applications of ergodic theory and sofic systems to cellular automata. *Physica D*, 10:36–44, 1984.

[7] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, New York, 1995.

[8] Marcus Pivato and Reem Yassawi. Limit measures for affine cellular automata. *Ergodic Theory & Dynamical Systems*, 22(4):1269–1287, August 2002.

[9] Marcus Pivato and Reem Yassawi. Limit measures for affine cellular automata II. to appear in *Ergodic Theory & Dynamical Systems;* preprint available: `http://arXiv.org/abs/math.DS/0108083`, 2003.

[10] S. Martínez P. Ferrari, A. Maass and P. Ney. Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory & Dynamical Systems*, 20(6):1657–1670, 2000.

[11] Karl Petersen. *Ergodic Theory*. Cambridge University Press, New York, 1989.

[12] Marcus Pivato. Multiplicative cellular automata on nilpotent groups: Structure, entropy, and asymptotics. *Journal of Statistical Physics*, 110(1/2):247–267, January 2003.

[13] Ross Kindermann and J. Laurie Snell. *Markov Random Fields and their Applications.* American Mathematical Society, Providence, Rhode Island, 1980.