Ergod. Th. & Dynam. Sys. (2006), **26**, 1177–1201
doi:10.1017/S0143385706000228

Asymptotic randomization of sofic shifts by linear cellular automata

MARCUS PIVATO and REEM YASSAWI

Department of Mathematics, Trent University, Canada 1600 West Bank Drive, Peterborough, Ontario, K9J 7B8, Canada
(e-mail: {pivato, ryassawi}@xaravve.trentu.ca)

(*Received* 31 *July* 2004 *and accepted* 6 *April* 2006)

Abstract. Let $\mathbb{M} = \mathbb{Z}^D$ be a D-dimensional lattice, and let $(\mathcal{A}, +)$ be an abelian group. $\mathcal{A}^{\mathbb{M}}$ is then a compact abelian group under componentwise addition. A continuous function $\Phi : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ is called a *linear cellular automaton* if there is a finite subset $\mathbb{F} \subset \mathbb{M}$ and non-zero coefficients $\varphi_{\mathsf{f}} \in \mathcal{Z}$ so that, for any $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$, $\Phi(\mathbf{a}) = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \sigma^{\mathsf{f}}(\mathbf{a})$. Suppose that 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 $\text{wk}^* - \lim_{\mathbb{J}\ni j\to\infty} \Phi^j \mu = \eta$, where η is the Haar measure on $\mathcal{A}^{\mathbb{M}}$, and $\mathbb{J} \subset \mathbb{N}$ has Cesàro density one. In the case when $\Phi = 1 + \sigma$ and $\mathcal{A} = (\mathbb{Z}_{/p})^s$ (p prime), we provide a condition on μ that is both necessary and sufficient. We then use this to construct zero-entropy measures which are randomized by $1 + \sigma$.

0. *Introduction*

Let $D \geq 1$, and let $\mathbb{M} := \mathbb{Z}^D$ be the D-dimensional lattice. If $\mathcal{A}$ is a (discretely topologized) finite set, then $\mathcal{A}^{\mathbb{M}}$ is compact in the Tychonoff topology. For any $\mathsf{v} \in \mathbb{M}$, let $\sigma^{\mathsf{v}} : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ be the shift map: $\sigma^{\mathsf{v}}(\mathbf{a}) := [b_{\mathsf{m}}|_{\mathsf{m}\in\mathbb{M}}]$, where $b_{\mathsf{m}} := a_{\mathsf{m}-\mathsf{v}}$, for all $\mathsf{m} \in \mathbb{M}$. 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 $\mathsf{m} \in \mathbb{M}$, $\sigma^{\mathsf{m}} \circ \Phi = \Phi \circ \sigma^{\mathsf{m}}$. 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 $\text{wk}^* - \lim_{\mathbb{J}\ni j\to\infty} \Phi^j \mu = \eta$, 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 Φ with a finite subset $\mathbb{F} \subset \mathbb{M}$ (with $\#(\mathbb{F}) \geq 2$), and non-zero coefficients $\varphi_{\mathsf{f}} \in \mathbb{Z}$ (for all $\mathsf{f} \in \mathbb{F}$) so that, for any $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$,

$$\Phi(\mathbf{a}) = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \sigma^{\mathsf{f}}(\mathbf{a}). \tag{1}$$

LCA are known to asymptotically randomize a wide variety of measures [**MHM03, MM98, MM99, Lin84, FMMN00**], including those satisfying a correlation-decay condition called *harmonic mixing* [**PY02, PY04, MMPY06**]. However, all known sufficient conditions for asymptotic randomization (and for harmonic mixing, in particular) require μ to have *full support*, i.e. $\mathsf{supp}(\mu) = \mathcal{A}^{\mathbb{M}}$.

Here we 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 let $\mathcal{A} = \mathbb{Z}_{/p}$ (p prime), and 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 let $\mathcal{A} = (\mathbb{Z}_{/p})^s$ (p prime, $s \in \mathbb{N}$) and introduce *bipartite* LCA, a broad class exemplified by the automaton $1 + \sigma$. 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 that 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 (e.g. $1 + \sigma$). Thus, the example of §2 *is* asymptotically randomized, even though it is not harmonically mixing.

In §7, we refine the results of §§3 and 4 by introducing *Lucas mixing* (a weaker condition than dispersion mixing). When $\mathcal{A} = (\mathbb{Z}_{/p})^s$, we show that a measure is asymptotically randomized by the automaton $1 + \sigma$ if and *only if* it is Lucas mixing. Finally, in §8, we use Lucas mixing to construct a class of *zero-entropy* measures which are asymptotically randomized by $1 + \sigma$, thereby refuting the conjecture that positive entropy is necessary for asymptotic randomization.

Preliminaries and notation. Throughout, $(\mathcal{A}, +)$ is an abelian group (usually $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where p is prime and $s \in \mathbb{N}$). Elements of $\mathcal{A}^{\mathbb{M}}$ are denoted by boldfaced letters (e.g. $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$), and subsets by gothic letters (e.g. $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$). Elements of $\mathbb{M}$ are sans serif (e.g. 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 $[\mathbf{b}] := \{\mathbf{c} \in \mathcal{A}^{\mathbb{M}}; \mathbf{c}_{\mathbb{U}} = \mathbf{b}\}$ be the corresponding cylinder set. In particular, if $\mathbf{a} \in \mathcal{A}^{\mathbb{M}}$, then $[\mathbf{a}_{\mathbb{U}}] := \{\mathbf{c} \in \mathcal{A}^{\mathbb{M}}; \mathbf{c}_{\mathbb{U}} = \mathbf{a}_{\mathbb{U}}\}$.

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 conditional probability measure in the cylinder set $[\mathbf{b}]$. In other words, for any $\mathfrak{X} \subset \mathcal{A}^{\mathbb{M}}$, $\mu^{(\mathbf{b})}[\mathfrak{X}] := \mu(\mathfrak{X} \cap [\mathbf{b}])/\mu[\mathbf{b}]$. 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 conditional probability measure on the $\mathbb{I}$ coordinates: for any $\mathbf{c} \in \mathcal{A}^{\mathbb{I}}$, $\mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{c}] := \mu([\mathbf{c}] \cap [\mathbf{b}])/\mu[\mathbf{b}]$.

Subshifts. A *subshift* [**Kit98, LM95**] is 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, $\mathfrak{X} := \bigcap_{\mathsf{m} \in \mathbb{M}} \sigma^{\mathsf{m}}[\mathfrak{W}]$, where $[\mathfrak{W}] := \{\mathbf{a} \in \mathcal{A}^{\mathbb{M}}; \mathbf{a}_{\mathbb{U}} \in \mathfrak{W}\}$. 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 a *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 by $\widehat{\mathcal{A}^{\mathbb{M}}}$. For any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$ there is a finite subset $\mathbb{K} \subset \mathbb{M}$, and non-trivial $\chi_{\mathsf{k}} \in \widehat{\mathcal{A}}$ for all $\mathsf{k} \in \mathbb{K}$, such 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 let $[\ell \ldots n) := \{m \in \mathbb{Z}; \ell \le m < n\}$. If $\mathbb{J} \subset \mathbb{N}$, then the *Cesàro density* of $\mathbb{J}$ is defined as

$$\mathsf{density}\,(\mathbb{J}) := \lim_{N \to \infty} \frac{1}{N} \#\,(\mathbb{J} \cap [0 \ldots N)).$$

If $\mathbb{J}, \mathbb{K} \subset \mathbb{N}$, then their *relative Cesàro density* is defined as

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

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

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

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 that $\mathbf{b}$ *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* [**Bré99, KS80**] 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 $\mathbb{B} = [-N \ldots N]$, then μ is a $\mathbb{B}$-MRF iff μ is an N-step Markov chain.

LEMMA 1.1. *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 $\mathbf{b} \in \mathcal{A}^{\mathbb{S}}$, $\#(\mathsf{supp}(\mu_0^{(\mathbf{b})})) \geq 2$.

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 $a, b \in \mathcal{A}$, $\#(\mathcal{F}(a) \cap \mathcal{P}(b)) \geq 2$.

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 the following.

LEMMA 1.2. *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 $c < 1$ such that, for all non-trivial $\chi \in \widehat{\mathcal{A}}$ and any* $\mathbf{b} \in \mathcal{A}^{\mathbb{S}}$, $|\langle \chi, \mu_0^{(\mathbf{b})} \rangle| \leq c$.

For any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$ and $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$, 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 $R \in \mathbb{N}$ such that, for any $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$,

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

The significance of this is the following [**PY02**, Theorem 12].

THEOREM A. *Let $\mathcal{A} = \mathbb{Z}_{/p}$, where p is prime. Any LCA on $\mathcal{A}^{\mathbb{M}}$ asymptotically randomizes any harmonically mixing measure.*

Most MRFs with full support are harmonically mixing [**PY04**, Theorem 15]. We now extend this.

THEOREM 1.3. *Let $\mathcal{A} = \mathbb{Z}_{/p}$, where p is prime. Any locally free MRF on $\mathcal{A}^{\mathbb{M}}$ is harmonically mixing.*

Proof. Let μ be a locally free $\mathbb{B}$-MRF. A subset $\mathbb{I} \subset \mathbb{M}$ is $\mathbb{B}$-*separated* if $(\mathsf{i} - \mathsf{j}) \notin \mathbb{B}$ for all $\mathsf{i}, \mathsf{j} \in \mathbb{I}$ with $\mathsf{i} \neq \mathsf{j}$. Let $\mathbb{K} \subset \mathbb{M}$ be finite, and let $\boldsymbol{\chi} := \bigotimes_{\mathsf{k} \in \mathbb{K}} \chi_{\mathsf{k}}$ be a character of $\mathcal{A}^{\mathbb{M}}$.

Claim 1. Let $K := \#(\mathbb{K}) = \mathsf{rank}[\boldsymbol{\chi}]$, and let $B := \max\{|\mathsf{b}_1 - \mathsf{b}_2|; \mathsf{b}_1, \mathsf{b}_2 \in \mathbb{B}\}$. There exists a $\mathbb{B}$-separated subset $\mathbb{I} \subset \mathbb{K}$ such that

$$\#(\mathbb{I}) = I \geq \frac{K}{B^D}. \tag{2}$$

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

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

for some set $\mathbb{I} \subset \mathbb{K}$. Thus, $|\mathsf{i} - \mathsf{j}| \geq B$ for any $\mathsf{i}, \mathsf{j} \in \mathbb{I}$ with $\mathsf{i} \neq \mathsf{j}$, so $(\mathsf{i} - \mathsf{j}) \notin \mathbb{B}$. Also, $\#(\widetilde{\mathbb{B}}) = B^D$, so each copy covers at most B^D points in $\mathbb{K}$. Thus, we require at least K/B^D copies to cover all of $\mathbb{K}$. In other words, $I \geq K/B^D$. $\diamond$

Thus, $\boldsymbol{\chi} = \boldsymbol{\chi}_{\mathbb{I}} \cdot \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}$, where $\boldsymbol{\chi}_{\mathbb{I}}(\mathbf{a}) := \prod_{\mathsf{i}\in\mathbb{I}} \chi_{\mathsf{i}}(a_{\mathsf{i}})$ and $\boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{a}) := \prod_{\mathsf{k}\in\mathbb{K}\setminus\mathbb{I}} \chi_{\mathsf{k}}(a_{\mathsf{k}})$.

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 $\mathbf{a} \in \mathcal{A}^{\mathbb{I}}$,

$$\mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{a}] = \prod_{\mathsf{i}\in\mathbb{I}} \mu_{\mathsf{i}}^{(\mathbf{b})}\{a_{\mathsf{i}}\}. \tag{3}$$

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

$$\begin{aligned}\langle \boldsymbol{\chi}_{\mathbb{I}}, \mu_{\mathbb{I}}^{(\mathbf{b})}\rangle &= \sum_{\mathbf{a}\in\mathcal{A}^{\mathbb{I}}} \mu_{\mathbb{I}}^{(\mathbf{b})}[\mathbf{a}] \cdot \left(\prod_{\mathsf{i}\in\mathbb{I}} \chi_{\mathsf{i}}(a_{\mathsf{i}})\right) \underset{(*)}{=} \sum_{\mathbf{a}\in\mathcal{A}^{\mathbb{I}}} \left(\prod_{\mathsf{i}\in\mathbb{I}} \mu_{\mathsf{i}}^{(\mathbf{b})}\{a_{\mathsf{i}}\} \cdot \chi_{\mathsf{i}}(a_{\mathsf{i}})\right) \\ &= \prod_{\mathsf{i}\in\mathbb{I}} \left(\sum_{a_{\mathsf{i}}\in\mathcal{A}} \mu^{(\mathbf{b})}\{a_{\mathsf{i}}\} \cdot \chi_{\mathsf{i}}(a_{\mathsf{i}})\right) = \prod_{\mathsf{i}\in\mathbb{I}} \langle \chi_{\mathsf{i}}, \mu_{\mathsf{i}}^{(\mathbf{b})}\rangle,\end{aligned}$$

where $(*)$ is by equation (3). Thus, $\langle \boldsymbol{\chi}, \mu^{(\mathbf{b})}\rangle = \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b}) \cdot \langle \boldsymbol{\chi}_{\mathbb{I}}, \mu_{\mathbb{I}}^{(\mathbf{b})}\rangle = \boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b}) \cdot \prod_{\mathsf{i}\in\mathbb{I}} \langle \chi_{\mathsf{i}}, \mu_{\mathsf{i}}^{(\mathbf{b})}\rangle$. Thus, if $I = \#(\mathbb{I})$, then

$$|\langle \boldsymbol{\chi}, \mu^{(\mathbf{b})}\rangle| = |\boldsymbol{\chi}_{\mathbb{K}\setminus\mathbb{I}}(\mathbf{b})| \cdot \prod_{\mathsf{i}\in\mathbb{I}} |\langle \chi_{\mathsf{i}}, \mu_{\mathsf{i}}^{(\mathbf{b})}\rangle| \leq 1 \cdot c^I \tag{4}$$

where the last step follows from Lemma 1.2. However, $\langle \boldsymbol{\chi}, \mu\rangle = \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot \langle \boldsymbol{\chi}, \mu^{(\mathbf{b})}\rangle$, so

$$|\langle \boldsymbol{\chi}, \mu\rangle| \leq \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot |\langle \boldsymbol{\chi}, \mu^{(\mathbf{b})}\rangle| \underset{(*)}{\leq} \sum_{\mathbf{b}\in\mathcal{A}^{\mathbb{J}}} \mu[\mathbf{b}] \cdot c^I = c^I \underset{(\dagger)}{\leq} c^{K/(B^D)} \xrightarrow[K\to\infty]{} 0.$$

Here $(*)$ is by equation (4) and $(\dagger)$ is by equation (2). □

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 will show in §§3–5 that this measure *is* asymptotically randomized by many LCA.

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

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

Let $\Phi : \mathfrak{X} \to (\mathbb{Z}_{/2})^{\mathbb{Z}}$ 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 1s between any two occurrences of 0 in $\mathbf{s}$.

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

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

LEMMA 2.1. *For all $i, j \in \mathbb{Z}_{/3}$, either $\Phi(\mathfrak{X}^N_{i,j}) \subset \mathfrak{E}_N$ or $\Phi(\mathfrak{X}^N_{i,j}) \subset \mathfrak{O}_N$. In particular,*

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

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

However, 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 $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2) \in \mathcal{M}[\mathbb{Z}_{/3}]$. Let $\nu := \Phi\mu \in \mathcal{M}[\mathfrak{S}]$, so that if $\mathfrak{U} \subset \mathfrak{S}$ is measurable, then $\nu[\mathfrak{U}] := \mu[\Phi^{-1}(\mathfrak{U})]$.

For all $N \in \mathbb{N}$, define a character $\boldsymbol{\chi}_N$ by $\boldsymbol{\chi}_N(\mathbf{x}) := \prod_{n=0}^{N} (-1)^{x_n}$ for all $\mathbf{x} \in (\mathbb{Z}_{/2})^{\mathbb{Z}}$. Then Lemma 2.1 implies

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

However, μ is mixing, so $\lim_{N\to\infty} \mu(\mathfrak{X}^N_{i,j}) = \rho_i \cdot \rho_j$. 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

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

with Perron measure $\boldsymbol{\rho} = (\frac{2}{5}, \frac{1}{5}, \frac{2}{5})$, then $\lim_{N\to\infty} \langle \boldsymbol{\chi}_N, \nu \rangle \neq 0$. Clearly, $\mathsf{rank}\,[\boldsymbol{\chi}_N] = N$, so that $\lim_{N\to\infty} \mathsf{rank}\,[\boldsymbol{\chi}_N] = \infty$. 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 define the concepts of *dispersion mixing* (DM) (for measures) and *dispersion* (for automata) which together yield asymptotic randomization. In §4 we will show that many LCA are dispersive. In §§5 and 6 we will show that many measures (including the even shift measure ν from §2) are DM.

Let Φ be an LCA as in (1). The advantage of this 'polynomial' notation is that composition of two LCA corresponds to multiplication of their respective polynomials. For example, suppose $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where $p \in \mathbb{N}$ is prime, and $s \in \mathbb{N}$. Suppose $\mathbb{M} = \mathbb{Z}$ and $\Phi = 1 + \sigma$; that is, $\Phi(\mathbf{a})_0 = a_0 + a_1 \pmod p$. Then the binomial theorem implies that

$$\text{for any } N \in \mathbb{N}, \quad \Phi^N = \sum_{n=0}^{N} \begin{bmatrix} N \\ n \end{bmatrix}_p \sigma^n, \quad \text{where} \quad \begin{bmatrix} N \\ n \end{bmatrix}_p := \binom{N}{n} \bmod p. \qquad (5)$$

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

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

If $\mathbb{F}, \mathbb{G} \subset \mathbb{M}$, and $\Phi = \sum_{\mathsf{f}\in\mathbb{F}} \varphi_{\mathsf{f}} \cdot \sigma^{\mathsf{f}}$ and $\Gamma = \sum_{\mathsf{g}\in\mathbb{G}} \gamma_{\mathsf{g}} \cdot \sigma^{\mathsf{g}}$ 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 that $\boldsymbol{\chi}$ and $\boldsymbol{\xi}$ are *S-separated* if $\mathbb{K}$ and $\mathbb{X}$ are S-separated.

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

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

For example, if

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

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

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

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

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

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

We say that μ is DM if, for every $\epsilon > 0$, there exist $S, R > 0$ such that, for any character $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{M}}}$, $(\mathsf{rank}_S(\boldsymbol{\chi}) > R) \Longrightarrow (|\langle \boldsymbol{\chi}, \mu\rangle| < \epsilon)$. Note that DM 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 such that $\lim_{\mathbb{J}\ni j\to\infty} \mathsf{rank}_S(\boldsymbol{\chi} \circ \Phi^j) = \infty$. We have the following.

THEOREM 3.1. *Let $\mathcal{A}$ be any finite abelian group. If $\Phi : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ is a dispersive LCA and $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{M}})$ is DM, then Φ asymptotically randomizes μ.* □

Theorem 3.1 is an immediate consequence of an easily verified lemma.

LEMMA 3.2. *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 $\mathsf{density}(\mathbb{J}) = 1$, such that $\lim_{\mathbb{J}\ni j\to\infty} |\langle \boldsymbol{\chi} \circ \Phi^j, \mu\rangle| = 0$.*

Proof. See the proof of Theorem 12 in **[PY02]**. □

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 $\Gamma = \sum_{\mathsf{g}\in\mathbb{G}} \gamma_{\mathsf{g}} \cdot \sigma^{\mathsf{g}}$ is an LCA, 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 $\mathbb{R}^n$):

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

We say Γ is *centred* if $|\mathsf{centre}(\Gamma)| < 1$. For any prime $p \in \mathbb{N}$, let

$$K_p := \min\left\{\frac{1}{2}, \frac{4p-7}{4p+4}\right\}.$$

Thus, $K_2 = \frac{1}{12}$, $K_3 = \frac{5}{16}$, and $K_p = \frac{1}{2}$, for $p \geq 5$.

Let $\mathcal{A} := (\mathbb{Z}_{/p})^s$ (where p is prime and $s \in \mathbb{N}$). If $\Phi : \mathcal{A}^{\mathbb{M}} \longrightarrow \mathcal{A}^{\mathbb{M}}$ is an LCA, then we say Φ is *bipartite* if $\Phi = 1 + \Gamma \circ \sigma^{\mathsf{f}}$, where Γ is centred and $\mathsf{diam}[\Gamma] \leq K_p \cdot |\mathsf{f}|$. For example:

$$\Phi = 1 + \sigma^{\mathsf{f}} \quad \text{is bipartite for any non-zero } \mathsf{f} \in \mathbb{M} \text{ and any prime } p \in \mathbb{N};$$
$$\Phi = 1 + \sigma^{12} + \sigma^{13} = 1 + (1+\sigma) \circ \sigma^{12} \quad \text{is bipartite for any prime } p \in \mathbb{N};$$
$$\Phi = 1 + \sigma^{14} + \sigma^{19} = 1 + (\sigma^{-2} + \sigma^{3}) \circ \sigma^{16} \quad \text{is bipartite for any prime } p \geq 3;$$
$$\Phi = 1 + \sigma^{2} + \sigma^{3} = 1 + (1+\sigma) \circ \sigma^{2} \quad \text{is bipartite for any prime } p \geq 5.$$

Our goal in this section is to prove the following theorem.

THEOREM 4.1. *Let $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where p prime and $s \in \mathbb{N}$. If Φ is bipartite, then Φ is dispersive.* □

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

LEMMA 4.2. (Lucas's theorem) *We have the following.*

(a) $\left[{N \atop n}\right]_p = \prod_{i=0}^{\infty} \left[{N^{(i)} \atop n^{(i)}}\right]_p$, *where we define* $\left[{N^{(i)} \atop n^{(i)}}\right]_p := 0$ *if* $n^{(i)} > N^{(i)}$, *and* $\left[{0 \atop 0}\right]_p := 1$.

(b) *Thus,* $\left[{N \atop n}\right]_p \neq 0$ *if and only if* $n \in \mathbb{L}(N)$.

For example, suppose $\mathbb{M} = \mathbb{Z}$ and $\Phi = 1 + \sigma$. If we interpret (5) in the light of Lemma 4.2, we get

$$\Phi^N = \sum_{n \in \mathbb{L}(N)} \begin{bmatrix} N \\ n \end{bmatrix}_p \sigma^n.$$

LEMMA 4.3. *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

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

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

LEMMA 4.4. *Let Φ be an LCA, and let $S > 0$.*

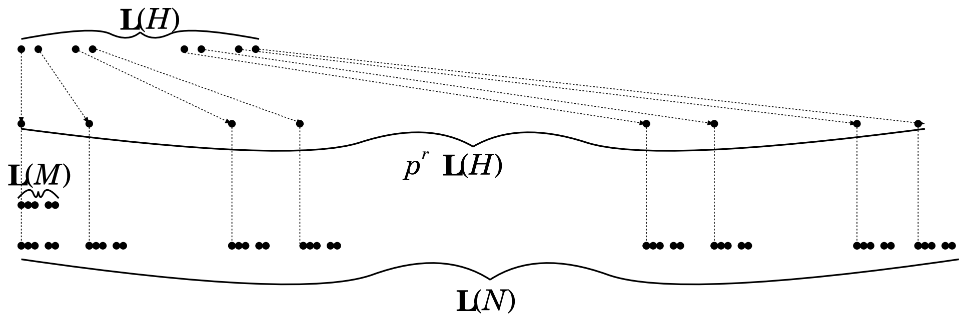

FIGURE 1. Lemma 4.3.

(a) *If* χ *is a character, and* $S_0 = S + \mathsf{diam}[\chi]$, *then* $\mathsf{rank}_S(\chi \circ \Phi) \geq \mathsf{rank}_{S_0}(\Phi)$.
(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 4.5. *The LCA* Φ *is dispersive if and only if, for any* $S_0 > 0$, *there is a subset* $\mathbb{J} \subset \mathbb{N}$ *of density 1 such that* $\lim_{\mathbb{J} \ni j \to \infty} \mathsf{rank}_{S_0}(\Phi^j) = \infty$.

To prove Theorem 4.1, we use Lemma 4.3 to verify the condition of Corollary 4.5. For any $S_0 > 0$, define

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

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 4.6. *We have*

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

Proof. Suppose that $N = M_N + p^{r_N} H_N$, 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 4.7. *We have* $\mathsf{density}\,(\mathbb{J}(S_0)) = 1$.

Proof. Let $\mathbb{I} := [pS_0 \ldots \infty]$. Then $\mathbb{I}$ is a set of density one, and Lemma 4.6 implies that

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

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

LEMMA 4.8. *If* $N \in \mathbb{J}(S_0)$, *and* $N = M + p^r H$, *then* $\Phi^N = \Phi^M \circ \Theta^H$, *where* $\Theta = \Phi^{(p^r)}$.

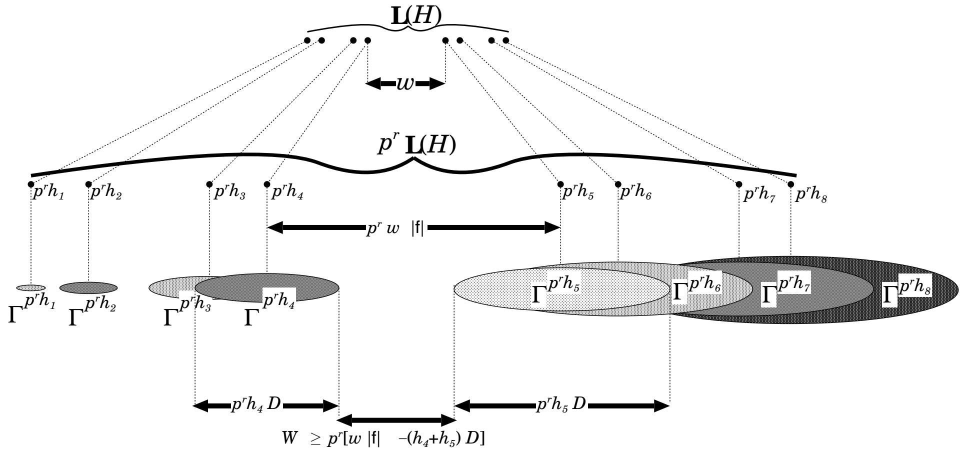

FIGURE 2. Claim 1 of Theorem 4.1.

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

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

Here (L) is by the Lucas theorem and (‡) is by Lemma 4.3(b), (†) is because $\Phi^M = \sum_{m \in \mathbb{L}(M)} \left[\begin{smallmatrix} M \\ m \end{smallmatrix}\right]_p (\Gamma \circ \sigma^{\mathsf{f}})^m$. Finally, $(\star)$ is because $\Theta = (1 + \Gamma \circ \sigma^{\mathsf{f}})^{p^r} \underset{(L)}{=} 1 + (\Gamma \circ \sigma^{\mathsf{f}})^{p^r}$. Thus,

$$\Theta^H \underset{(L)}{=} \sum_{h \in \mathbb{L}(H)} \begin{bmatrix} H \\ h \end{bmatrix}_p (\Gamma \circ \sigma^{\mathsf{f}})^{p^r h}.$$

□

Proof of Theorem 4.1. It suffices to verify the condition of Corollary 4.5. So, let $S_1 := S_0 + \mathsf{diam}[\Phi^M]$. Then

$$\mathsf{rank}_{S_0}(\Phi^N) \underset{(*)}{=} \mathsf{rank}_{S_0}(\Phi^M \circ \Theta^H) \underset{(\dagger)}{\geq} \mathsf{rank}_{S_1}(\Theta^H) \tag{6}$$

where $(*)$ is by Lemma 4.8 and (†) is by Lemma 4.4(b).

Thus, we want to show that $\mathsf{rank}_{S_1}(\Theta^H) \xrightarrow[H \to \infty]{} \infty$ for H in a set of density 1. To do this, we 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:

$$\textbf{(i)}\ h_1 \geq p \cdot h_0 \quad \text{and} \quad \textbf{(ii)}\ [h_0 \ldots h_1) \cap \mathbb{L}(H) = \emptyset.$$

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 $(\Gamma \circ \sigma^{\mathsf{f}})^{p^r h_0}$ and $(\Gamma \circ \sigma^{\mathsf{f}})^{p^r h_1}$ are S_1-separated.

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

$$\begin{aligned} W &:= p^r w|\mathsf{f}| - (\mathsf{diam}[\Gamma^{p_r h_0}] + \mathsf{diam}[\Gamma^{p_r h_1}]) = p^r w|\mathsf{f}| - (p^r h_0 D + p^r h_1 D) \\ &\geq p^r \cdot (w|\mathsf{f}| - D \cdot (h_1 + h_0)) \end{aligned} \tag{7}$$

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

$$\begin{aligned} \mathsf{diam}[\Phi^M] &\leq M \cdot |\mathsf{f}| + 2 \cdot \max_{m \in \mathbb{L}(M)} \mathsf{diam}[\Gamma^m] = M \cdot |\mathsf{f}| + 2M \cdot D \\ &= M \cdot (|\mathsf{f}| + 2D) \leq p^{r-1} \cdot (|\mathsf{f}| + 2D). \end{aligned} \tag{8}$$

Thus,

$$\begin{aligned} W - \mathsf{diam}[\Phi^M] &\underset{(*)}{\geq} p^r \cdot (w \cdot |\mathsf{f}| - D \cdot (h_1 + h_0)) - p^{r-1} \cdot (|\mathsf{f}| + 2D) \\ &= p^{r-1} \cdot (pw \cdot |\mathsf{f}| - pD \cdot (h_1 + h_0) - |\mathsf{f}| - 2D) \\ &\underset{(\dagger)}{\geq} S_0 \cdot (pw \cdot |\mathsf{f}| - pD \cdot (h_1 + h_0) - |\mathsf{f}| - 2D) \end{aligned}$$

where $(*)$ is by (7) and (8), and $(\dagger)$ is because $S_0 < p^{r-1}$. Thus, it suffices to show that

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

To see this, observe that

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

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

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

$(\star)$ is because

$$K_p \leq \frac{4p - 7}{4p + 4} = \frac{p - 7/4}{p + 1},$$

thus, $(p+1)K_p \leq p - \frac{7}{4} = p - 1 - \frac{3}{4}$, so, $\frac{3}{4} \leq (p-1) - (p+1)K_p = (1 - K_p)p - (1 + K_p)$, and

($\diamond$) is because $p \geq 2$, so $p^2 \geq 4$. It follows that $W - \mathsf{diam}[\Phi^M] \geq S_0$, so that $W \geq S_1$. $\diamond$

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

$$\mathsf{rank}_{S_1}(\Theta^H) \geq \mathsf{rank}\,[H]. \tag{9}$$

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

Suppose that $i < k$. 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 that $p = 2$ 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

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

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

Let $\#\mathbf{ZB}(H) :=$ #of zero-blocks in the p-ary expansion of H. Then Claim 2 implies that

$$\mathsf{rank}\,[H] \geq \#\mathbf{ZB}(H). \tag{10}$$

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

Claim 3. We claim $\mathsf{density}\,(\mathbb{H}) = 1$.

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

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

Then $\mathbb{H}' \subset \mathbb{H}$. The weak law of large numbers implies $\mathsf{density}\,(\mathbb{H}') = 1$. $\diamond$

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

Claim 4. We claim $\mathsf{density}\,(\mathbb{J}) = 1$.

Proof. We have $\mathbb{J} = \mathbb{J}_1 \cap \mathbb{J}_2$, where

$$\mathbb{J}_1 := \{N \in \mathbb{J}(S_0);\ N = M_N + p^{r_N} H_N, \text{ where } H_N \in \mathbb{H}\} \text{ and}$$
$$\mathbb{J}_2 := \{N \in \mathbb{J}(S_0);\ N = M_N + p^{r_N} H_N, \text{ where } r_N \leq \tfrac{1}{2}\log_p(N)\}.$$

Now, $\mathsf{density}\,(\mathbb{J}_1) = 1$ by Lemma 4.7 and Claim 3. To see that $\mathsf{density}\,(\mathbb{J}_2) = 1$, note that

$$\mathbb{J}(S_0) \setminus \mathbb{J}_2 \subset \{N \in \mathbb{N};\ N^{(r)} \neq 0 \text{ for all } r \in (\log_p(S_0) \ldots \tfrac{1}{2}\log_p(N)]\},$$

which is a set of density zero. $\diamond$

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

$$\log_p(H_N) \geq \log_p(N) - r_N \geq \log_p(N) - \tfrac{1}{2}\log_p(N) = \tfrac{1}{2}\log_p(N). \tag{11}$$

Thus,

$$\begin{aligned} \mathsf{rank}_{S_0}(\Phi^N) \underset{(\heartsuit)}{\geq} \mathsf{rank}_{S_1}(\Theta^{H_N}) \underset{(\diamondsuit)}{\geq} \mathsf{rank}\,[H_N] \underset{(\clubsuit)}{\geq} \#\mathbf{ZB}(H_N) \\ \underset{(*)}{\geq} \frac{1}{p^3}\log_p(H_N) \underset{(\spadesuit)}{\geq} \frac{1}{2p^3}\log_p(N). \end{aligned}$$

Here, $(\heartsuit)$ is by equation (6), $(\diamondsuit)$ is by equation (9), $(\clubsuit)$ is by equation (10), $(\spadesuit)$ is by equation (11), and $(*)$ is because $H \in \mathbb{H}$ by hypothesis. Thus,

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

□

5. *Uniform mixing and DM*

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 $\mathfrak{L} \subset \mathcal{A}^{(-\infty\ldots 0]}$ and $\mathfrak{R} \subset \mathcal{A}^{[0\ldots\infty)}$, and any $m > M$,

$$\mu[\sigma^m(\mathfrak{L}) \cap \mathfrak{R}] \underset{\epsilon}{\sim} \mu[\mathfrak{L}] \cdot \mu[\mathfrak{R}] \tag{12}$$

(here '$x \underset{\epsilon}{\sim} y$' means that $|x - y| < \epsilon$).

Example 5.1.

(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 $\psi : \mathcal{B}^{[-\ell\ldots r]} \longrightarrow \mathcal{A}$, then replace the M in (12) 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 the following theorem.

THEOREM 5.2. *Let $\mathcal{A}$ be a finite abelian group. If $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ is uniformly mixing and harmonically bounded, then μ is dispersion mixing.*

We will then apply Theorem 5.2 to get the following.

COROLLARY 5.3. *Let $\mathcal{A} = \mathbb{Z}_{/p}$, where p is 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 5.4. *Let $\mathcal{A} = (\mathbb{Z}_{/p})^s$, where p is prime and $s \in \mathbb{N}$. If $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ and $h(\mu, \sigma) > (s-1) \cdot \log_2(p)$, then μ is harmonically bounded.*

Proof. Suppose that 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 $\mathcal{I} := \text{image}(\boldsymbol{\chi}) \subset \mathbb{T}^1$, 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 $\alpha > 0$ such that, for any probability measure $\nu \in \mathcal{M}(\mathcal{I})$ with $|\sum_{i \in \mathcal{I}} i \cdot \nu\{i\}| > 1 - \alpha$, there is some $i_0 \in \mathcal{I}$ with $\nu\{i_0\} > 1 - \beta$. $\diamond$

Suppose that $\boldsymbol{\chi} = \bigotimes_{k \in \mathbb{K}} \chi_k$, where $\mathbb{K} \subset [0 \ldots K]$ 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 $\mathbf{b} \in \mathcal{A}^{[0 \ldots K)}$, let $\mu_K^{(\mathbf{b})}$ be the conditional measure on the Kth coordinate, and let $\nu_K^{(\mathbf{b})} := \chi_K(\mu_K^{(\mathbf{b})}) \in \mathcal{M}(\mathcal{I})$ be the projected measure on $\mathcal{I}$.

Claim 2. For any $\gamma > 0$, there exists $\beta > 0$ such that, if there exists $i_0 \in \mathcal{I}$ with $\nu\{i_0\} > 1 - \beta$, then there is a subset $\mathfrak{B} \subset \mathcal{A}^{[0 \ldots K)}$ with $\mu[\mathfrak{B}] > 1 - \gamma$, such 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 $\#(\mathcal{P}_{\mathbf{b}}) \leq p^{s-1}$ for all $\mathbf{b} \in \mathcal{A}^{[0 \ldots K)}$.) $\diamond$

For any measure $\rho \in \mathcal{M}(\mathcal{A})$, define $H(\rho) := -\sum_{a \in \mathcal{A}} \rho\{a\} \log_2(\rho\{a\})$. Recall (e.g. [**Pet89**, Proposition 5.2.12]) that the σ-entropy of μ can be computed

$$h(\mu, \sigma) = \lim_{N \to \infty} \sum_{\mathbf{b} \in \mathcal{A}^{[0 \ldots N)}} \mu[\mathbf{b}] \cdot H(\mu_N^{(\mathbf{b})}). \tag{13}$$

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

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

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 $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ with $|\langle \boldsymbol{\chi}, \mu \rangle| > 1 - \alpha$. It then follows from Claims 1–4 that $\sum_{\mathbf{b} \in \mathcal{A}^{[0 \ldots K)}} \mu[\mathbf{b}] \cdot H(\mu_N^{(\mathbf{b})}) < (s-1) \cdot \log_2(p) - \epsilon$. However, the limit in (13) is a decreasing limit, so we conclude that $h(\mu, \sigma) < (s-1) \cdot \log_2(p) - \epsilon$. Since this is true for any $\epsilon > 0$, we conclude that $h(\mu, \sigma) \leq (s-1) \cdot \log_2(p)$, contradicting our hypothesis. $\square$

COROLLARY 5.5. *If* $\mathcal{A} = \mathbb{Z}_{/p}$ (where p prime) *and* $h(\mu, \sigma) > 0$, *then* μ *is harmonically bounded.*

Say μ is *uniformly multiply mixing* if, for any $\epsilon > 0$, there is some $S > 0$ such 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

$$\mu\left(\bigcap_{r=0}^{R} \mathfrak{U}_r\right) \underset{R\epsilon}{\approx} \prod_{r=0}^{R} \mu(\mathfrak{U}_r). \tag{14}$$

LEMMA 5.6. *If* $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ *is uniformly mixing, then* μ *is uniformly multiply mixing.*

Proof (By induction on R). The case $R = 1$ is just uniform mixing. Suppose (14) 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 5.7. *Suppose that $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ 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 $r \in [0 \ldots R]$, $\boldsymbol{\chi}_r : \mathcal{A}^{\mathbb{K}_r} \longrightarrow \mathbb{C}$ are characters, and $\boldsymbol{\chi} = \prod_{r=0}^{R} \boldsymbol{\chi}_r$, then $\langle \boldsymbol{\chi}, \mu \rangle \underset{\epsilon/2}{\widetilde{\quad}} \prod_{r=0}^{R} \langle \boldsymbol{\chi}_r, \mu \rangle$.*

Proof of Theorem 5.2. Let $\epsilon > 0$. We want to find $S > 0$ and $R > 0$ such that, if $\boldsymbol{\chi}$ is any character, and $\mathsf{rank}_S(\boldsymbol{\chi}) > R$, then $|\langle \boldsymbol{\chi}, \mu \rangle| < \epsilon$. 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 5.7. Suppose $\mathsf{rank}_S(\boldsymbol{\chi}) > R$, and let $\boldsymbol{\chi} := \bigotimes_{r=0}^{R} \boldsymbol{\chi}_r$, where $\boldsymbol{\chi}_r : \mathcal{A}^{\mathbb{K}_r} \longrightarrow \mathbb{C}$ are characters, and $\mathbb{K}_0, \ldots, \mathbb{K}_R \subset \mathbb{Z}$ are S-separated. Then Lemma 5.7 implies that

$$\langle \boldsymbol{\chi}, \mu \rangle \underset{\epsilon/2}{\widetilde{\quad}} \prod_{r=0}^{R} \langle \boldsymbol{\chi}_r, \mu \rangle. \tag{15}$$

By harmonic boundedness, we know $|\langle \boldsymbol{\chi}_r, \mu \rangle| < C$ for all $r \in [0 \ldots R]$. Thus, (15) implies

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

□

Proof of Corollary 5.3. From Examples 5.1(a) and (b), we know μ is uniformly mixing. Any mixing quasi-Markov measure has non-zero entropy, so Corollary 5.5 says that μ is harmonically bounded. Theorem 5.2 says μ is dispersion mixing. Theorem 3.1 says μ is asymptotically randomized by any dispersive CA.

6. *Markov words*

If $m, n \in \mathbb{Z}$, and $m \leq n$, let $\mathcal{A}^{[m\ldots n)}$ be the set of all *words* of the form $\mathbf{a} = [a_m, a_{m+1}, \ldots, a_{n-1}]$. Let $\mathcal{A}^* := \bigcup_{-\infty<m<n<\infty} \mathcal{A}^{[m\ldots n)}$ be the set of all finite words. Elements of $\mathcal{A}^*$ are denoted by boldfaced letters (e.g. $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$), and subsets by gothic letters (e.g. $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$). 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 $\mathbf{v} \in \mathcal{A}^{[-V\ldots V)}$, we say that $\mathbf{v}$ is a *Markov word* for μ if (in the terminology of §1), $\mathbf{v}$ isolates $(-\infty \ldots -V)$ from $[V \ldots \infty)$.

Example 6.1.

(a) If μ is an N-step Markov shift, and $N \leq 2V$, then every $\mathbf{v} \in \mathcal{A}^{[-V\ldots V)}$ 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 6.2. *If $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ 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 $M_\epsilon(\mathbf{a}, \mathbf{b}) < \infty$ such that, for all $m > M_\epsilon(\mathbf{a}, \mathbf{b})$, $\mu(\sigma^m[\mathbf{a}] \cap [\mathbf{b}]) \underset{\epsilon}{\sim} \mu[\mathbf{a}] \cdot \mu[\mathbf{b}]$. 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

$$\begin{aligned}\mu(\sigma^m[\mathbf{uvw}] \cap [\mathbf{u}'\mathbf{v}\mathbf{w}']) &= \mu(\mathbf{u} \dashleftarrow \mathbf{v}) \cdot \mu(\sigma^m[\mathbf{vw}] \cap [\mathbf{u}'\mathbf{v}]) \cdot \mu(\mathbf{v} \dashrightarrow \mathbf{w}') &(16)\\ &\underset{\epsilon}{\sim} \mu(\mathbf{u} \dashleftarrow \mathbf{v}) \cdot \mu[\mathbf{vw}] \cdot \mu[\mathbf{u}'\mathbf{v}] \cdot \mu(\mathbf{v} \dashrightarrow \mathbf{w}') &(17)\\ &= \mu[\mathbf{uvw}] \cdot \mu[\mathbf{u}'\mathbf{v}\mathbf{w}']. &(18)\end{aligned}$$

Equations (16) and (18) are because $\mathbf{v}$ is a Markov word; (17) is because $m > M_\epsilon(\mathbf{vw}, \mathbf{u}'\mathbf{v})$. ◇

If $\mathbf{a} \in \mathcal{A}^*$, we say that $\mathbf{v}$ *occurs* in $\mathbf{a}$ if $\mathbf{a}|_{[n-V...n+V)} = \mathbf{v}$ for some n.

Claim 2. There is some $N > 0$ such that $\mu\{\mathbf{a} \in \mathcal{A}^{[0...N]}; \mathbf{v} \text{ occurs in } \mathbf{a}\} > 1 - \epsilon$.

Proof. By ergodicity, find N such that

$$\mu\left(\bigcup_{n=0}^{N} \sigma^n[\mathbf{v}]\right) > 1 - \epsilon.$$

◇

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

$$\mu(\mathcal{A}^*_{\mathbf{v}}) > 1 - \epsilon \text{ and } \mu({}_{\mathbf{v}}\mathcal{A}^*) > 1 - \epsilon. \tag{19}$$

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

$$\begin{aligned}\mathcal{A}^*_{\mathbf{v}} &= \{\mathbf{uvw}; \mathbf{u} \in \mathcal{A}^* \text{ and } \mathbf{w} \in \mathcal{A}^{<N}\} \quad \text{and}\\ {}_{\mathbf{v}}\mathcal{A}^* &= \{\mathbf{u}'\mathbf{v}\mathbf{w}'; \mathbf{u}' \in \mathcal{A}^{<N} \text{ and } \mathbf{w}' \in \mathcal{A}^*\}.\end{aligned} \tag{20}$$

Define

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

where $(*)$ is by (20) and $(\dagger)$ is by Claim 1. Likewise, define

$$M_2 := \max_{\mathbf{a}\in\mathcal{A}^*_{\mathbf{v}}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\mathbf{b}) = \max_{\mathbf{w}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{vw},\mathbf{b}),$$

$$M_3 := \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in_{\mathbf{v}}\mathcal{A}^*} M_\epsilon(\mathbf{a},\mathbf{b}) = \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{u}'\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\mathbf{u}'\mathbf{v}), \quad \text{and}$$

$$M_4 := \max_{\mathbf{a}\in\mathcal{A}^{<N}} \max_{\mathbf{b}\in\mathcal{A}^{<N}} M_\epsilon(\mathbf{a},\mathbf{b}).$$

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 $\mathbf{a},\mathbf{b}\in\mathcal{A}^*$, $M_\epsilon(\mathbf{a},\mathbf{b}) < M$.

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 (19) implies that $\mu[\mathbf{a}] < \epsilon$. Hence, for any $m\in\mathbb{N}$, $\mu(\sigma^m[\mathbf{a}]\cap\mathbf{b}) < \epsilon$ and $\mu[\mathbf{a}]\cdot\mu[\mathbf{b}] < \epsilon$. Thus, $\mu(\sigma^m[\mathbf{a}]\cap\mathbf{b}) \underset{\epsilon}{\sim} \mu[\mathbf{a}]\cdot\mu[\mathbf{b}]$ automatically. Hence, $M_\epsilon(\mathbf{a},\mathbf{b}) = 0 < M$.

Likewise, if $\mathbf{b}\notin\mathcal{A}^{<N}\cup_{\mathbf{v}}\mathcal{A}^*$, then $M_\epsilon(\mathbf{a},\mathbf{b}) = 0 < M$. ◇

Thus, μ is uniformly mixing. □

COROLLARY 6.3. *If μ is harmonically bounded, mixing and has a Markov word, then μ is asymptotically randomized by $\Phi = 1+\sigma$.*

Proof. Combine Proposition 6.2 with Theorems 3.1 and 5.2. □

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 $\Phi := 1+\sigma$. We will introduce a condition on μ which is weaker than DM, and which is both sufficient and *necessary* for asymptotic randomization.

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

$$\langle\langle\boldsymbol{\chi}\rangle\rangle := p^r \quad \text{where } r := \lceil \log_p |[\boldsymbol{\chi}]| \rceil.$$

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

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

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 $L := \#(\mathbb{L}(h))$, 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 non-trivial character $\boldsymbol{\chi}\in\widehat{\mathcal{A}^{\mathbb{Z}}}$, there is a subset $\mathbb{H}\subset\mathbb{N}$ of Cesàro density one such that $\lim_{\mathbb{H}\ni h\to\infty}\langle\boldsymbol{\chi}\circ\Phi^{h\cdot\langle\langle\boldsymbol{\chi}\rangle\rangle},\mu\rangle = 0$.

Our goal in this section is to prove the following.

THEOREM 7.1. *We have*

$$(\Phi = 1+\sigma \textit{ asymptotically randomizes } \mu) \iff (\mu \textit{ is Lucas mixing}).$$

It is relatively easy to see that the following holds.

LEMMA 7.2. *If μ is DM, then μ is Lucas mixing.*

Thus, the '$\Longleftarrow$' direction of Theorem'7.1 is an extension of Theorem 3.1, in the case $\Phi = 1 + \sigma$. The '$\Longrightarrow$' direction makes this the strongest possible extension for this LCA.

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 4.7 that $\mathsf{density}\,(\widetilde{\mathbb{J}}) = 1$. For any $m \in \mathbb{N}$, let $\boldsymbol{\chi}^m := \boldsymbol{\chi} \circ \Phi^m$.

LEMMA 7.3. *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 4.8 to observe that $\Phi^j = \Phi^m \circ \Phi^{h \cdot (p^r)}$. Thus,

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

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

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

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 7.1. We use Lemma 3.2.

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

$$\widetilde{\mathbb{J}}_m := \{m + p^{r(m)} h;\ h \in \mathbb{N}\}. \tag{21}$$

It follows that

$$\widetilde{\mathbb{J}} = \bigcup_{m \in \mathbb{N}} \widetilde{\mathbb{J}}_m. \tag{22}$$

If $j = m + p^{r(m)} h$ is an element of $\widetilde{\mathbb{J}}_m$, then Lemma 7.3 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 $\lim_{\widetilde{\mathbb{H}}_m \ni h \to \infty} \langle \boldsymbol{\chi}^m \circ \Phi^{h \cdot \langle\langle \boldsymbol{\chi}^m \rangle\rangle}, \mu \rangle = 0$. Define

$$\mathbb{H}_m := \left\{ h \in \widetilde{\mathbb{H}}_m;\ |\langle \boldsymbol{\chi}^m \circ \Phi^{h \cdot \langle\langle \boldsymbol{\chi}^m \rangle\rangle}, \mu \rangle| \leq \frac{1}{m} \right\},$$

$$\mathbb{J}_m := \{m + p^{r(m)} h;\ h \in \mathbb{H}_m\}, \quad \text{and} \tag{23}$$

$$\mathbb{J} := \bigcup_{m \in \mathbb{N}} \mathbb{J}_m. \tag{24}$$

Claim 1. We claim that $\mathsf{density}\,(\mathbb{J}) = 1$.

Proof. For any $m \in \mathbb{N}$, there is some K such that $\mathbb{H}_m = \widetilde{\mathbb{H}}_m \cap [K \ldots \infty)$. Thus, $\mathsf{rel\,density}[\mathbb{H}_m / \widetilde{\mathbb{H}}_m] = 1$. Thus, $\mathsf{density}\,(\mathbb{H}_m) = \mathsf{density}\,(\widetilde{\mathbb{H}}_m) = 1$. Compare (21) and (23) to see that $\mathsf{rel\,density}[\mathbb{J}_m / \widetilde{\mathbb{J}}_m] = 1$. Then compare (22) and (24) to see that $\mathsf{rel\,density}[\mathbb{J} / \widetilde{\mathbb{J}}] = 1$. Thus, $\mathsf{density}\,(\mathbb{J}) = \mathsf{density}\,(\widetilde{\mathbb{J}}) = 1$. ◇

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

Proof. Fix $\epsilon > 0$. Let M be large enough that $1/M < \epsilon$. For all $m \in \mathbb{N}$ with $m < M$, find H_m such that, if $h \in \widetilde{\mathbb{H}}_m$ and $h > H_m$, then $|\langle \chi^m \circ \Phi^{h \cdot \langle\langle \chi^m \rangle\rangle}, \mu \rangle| < \epsilon$. 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 $|\langle \chi \circ \Phi^j, \mu \rangle| = |\langle \chi^m \circ \Phi^{h \cdot \langle\langle \chi^m \rangle\rangle}, \mu \rangle| < \epsilon$.

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 $|\langle \chi \circ \Phi^j, \mu \rangle| < \epsilon$ by construction of J, or $j \in \mathbb{J}_m$ for some $m > M$, in which case

$$|\langle \chi \circ \Phi^j, \mu \rangle| \underset{(*)}{<} \frac{1}{m} < \frac{1}{M} \underset{(\dagger)}{<} \epsilon.$$

Here, $(*)$ follows by the definition of $\mathbb{H}_m$, and $(\dagger)$ follows by the definition of M. ◇

Lemma 3.2 and Claims 1 and 2 imply that Φ asymptotically randomizes μ.

'$\Longrightarrow$' Suppose that μ was not weakly harmonically mixing. Thus, there is some $\chi \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ and some subset $\mathbb{H} \subset \mathbb{N}$ of density $\delta > 0$ such that $\limsup_{\mathbb{H} \ni h \to \infty} |\langle \chi \circ \Phi^{h \cdot \langle\langle \chi \rangle\rangle}, \mu \rangle| > 0$. However, $\chi \circ \Phi^{h \cdot \langle\langle \chi \rangle\rangle} = \chi \circ \Phi^{p^r \cdot h}$ (where $r = \lceil \log_p |[\chi]| \rceil$). Hence, if $\mathbb{J} := p^r \cdot \mathbb{H}$, then $\mathsf{density}(\mathbb{J}) = p^{-r} \cdot \delta > 0$, and $\limsup_{\mathbb{J} \ni j \to \infty} |\langle \chi \circ \Phi^j, \mu \rangle| = \limsup_{\mathbb{H} \ni h \to \infty} |\langle \chi \circ \Phi^{h \cdot \langle\langle \chi \rangle\rangle}, \mu \rangle| > 0$. However, then Lemma 3.2 implies that Φ cannot randomize μ. □

8. *Randomization of zero-entropy measures*

Of the probability measures which are asymptotically randomized by LCA, every known example has positive entropy. However, we will show that positive entropy is *not* necessary, by constructing a class of zero-entropy measures which are Lucas mixing and, thus (by Theorem 7.1), randomized by $\Phi = 1 + \sigma$.

For both efficiency and lucidity, we will employ probabilistic language. Let $(\Omega, \mathcal{B}, \rho)$ be an abstract probability space (called the *sample space*). If $(\mathbf{X}, \mathcal{X})$ is any measurable space, then an ($\mathbf{X}$-valued) *random variable* is a measurable function $f : \Omega \longrightarrow \mathbf{X}$. In particular, a *random sequence* is a measurable function $\mathbf{a} : \Omega \longrightarrow \mathcal{A}^{\mathbb{Z}}$. By convention, we suppress the argument of random variables. Thus, if $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are random sequences, then the equation '$\mathbf{a} + \mathbf{b} = \mathbf{c}$' means '$\mathbf{a}(\omega) + \mathbf{b}(\omega) = \mathbf{c}(\omega)$, for ρ-almost all $\omega \in \Omega$'.

If $f : \Omega \longrightarrow \mathbf{X}$ is a random variable, and $\mathbf{U} \subset \mathbf{X}$, then '$\mathsf{Prob}[f \in \mathbf{U}]$' denotes $\rho[f^{-1}(\mathbf{U})]$. If $g : \Omega \longrightarrow \mathbf{Y}$ is another random variable, then f and g are *independent* if, for any measurable $\mathbf{U} \subset \mathbf{X}$ and $\mathbf{V} \subset \mathbf{Y}$, $\mathsf{Prob}[f \in \mathbf{U}$ and $g \in \mathbf{V}] = \mathsf{Prob}[f \in \mathbf{U}] \cdot \mathsf{Prob}[g \in \mathbf{V}]$ i.e. $\rho[f^{-1}(\mathbf{U}) \cap g^{-1}(\mathbf{V})] = \rho[f^{-1}(\mathbf{U})] \cdot \rho[g^{-1}(\mathbf{V})]$. The *distribution* of f is the probability measure $\mu := f(\rho)$ on $(\mathbf{X}, \mathcal{X})$; we then say that f is a *μ-random variable*. Thus, every random variable determines a probability measure on its range. However, given a measure μ, we can construct infinitely many independent μ-random variables.

Let $\mathcal{A} := \mathbb{Z}_{/2}$ and $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$, and consider a μ-random sequence $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$. We say μ has *independent random dyadic increments* (IRDI) if, for any $n \in \mathbb{N}$, and all $m \in [1 \ldots 2^n]$, $a_{m+2^n} = a_m + d_m^n$, where $d_1^n, \ldots, d_{2^n}^n$ are independent $\mathcal{A}$-valued random variables. If $d_1^n, \ldots, d_{2^n}^n$ have distributions $\delta_1^n, \ldots, \delta_{2^n}^n$, then μ has *lower decay rate* $\alpha \in (0, 1)$ if there is some $L > 0$ such that, for all $n \ge L$, and all $m \in [1 \ldots 2^n]$, $\alpha^n \le \delta_m^n\{1\}$.

PROPOSITION 8.1. *If μ has IRDI with lower decay rate $\alpha > 1/\sqrt{2}$, then μ is Lucas mixing.*

Proof. Let $\boldsymbol{\chi} \in \widehat{\mathcal{A}^{\mathbb{Z}}}$ be a non-trivial character. We seek $\mathbb{H} \subset \mathbb{N}$ with $\mathsf{density}(\mathbb{H}) = 1$, such that $\lim_{\mathbb{H}\ni h\to\infty} \langle \boldsymbol{\chi} \circ \Phi^{h\cdot\langle\langle\boldsymbol{\chi}\rangle\rangle}, \mu\rangle = 0$.

If $n \in \mathbb{N}$, let $I = I(n) := \lceil \log_2(n) \rceil$, and suppose that n has binary expansion $\{n^{(i)}\}_{i=0}^{I}$. Let $\mathbb{I}(n) := \{j \in [0 \ldots I];\ n^{(j)} = 1\}$. Let $\epsilon > 0$ be small, and define

$$\mathbb{H} := \{h \in \mathbb{N};\ \#(\mathbb{I}(h)) \geq \tfrac{1}{2} I(h) - \epsilon\}.$$

Then $\mathsf{density}(\mathbb{H}) = 1$. Suppose that $n \in \mathbb{H}$ is large; let $\mathbb{I} := \mathbb{I}(n)$ and $I := I(n)$. Assume that I is large (in particular, $I > L$).

Now, $\alpha > 1/\sqrt{2}$, so find β such that $1/\alpha < \beta < \sqrt{2}$. Define

$$M := \#(\mathbb{I}) - 1 \geq \tfrac{1}{2} I - \epsilon - 1 \underset{(*)}{>} \log_2(\beta) I, \tag{25}$$

where $(*)$ is because $\log_2(\beta) < \frac{1}{2}$ and I is large, while ϵ is small.

Suppose that $\mathbb{I} = \{i_1 < i_2 < \cdots < i_{M+1} = I\}$. Let $\boldsymbol{\xi}_0 := \boldsymbol{\chi}$, and for each $m \in [0 \cdots M]$, define $\boldsymbol{\xi}_{m+1} := \boldsymbol{\xi}_m \otimes (\boldsymbol{\xi}_m \circ \sigma^{L_i})$, where $L_i := 2^{i_m} \cdot \langle\langle \chi \rangle\rangle$. Thus, $\boldsymbol{\chi} \circ \Phi^{n\cdot\langle\langle\boldsymbol{\chi}\rangle\rangle} = \boldsymbol{\xi}_{M+1}$.

Let $r := \mathsf{rank}[\boldsymbol{\chi}]$. Then for all $m \in [1 \ldots M+1]$, $\mathsf{rank}[\boldsymbol{\xi}_m] = 2^m \cdot r$. In particular, define

$$R := \mathsf{rank}[\boldsymbol{\xi}_M] = 2^M \cdot r \underset{(*)}{>} (*)\beta^I \cdot r \tag{26}$$

where $(*)$ is by equation (25). Thus, $\boldsymbol{\xi}_M = \bigotimes_{x\in\mathbb{X}} \xi_x$, where $\mathbb{X} \subset \mathbb{Z}$ is a subset with $\#(\mathbb{X}) = R$. Thus, if $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$ is a μ-random sequence, then

$$\begin{aligned} \boldsymbol{\xi}_{M+1}(\mathbf{a}) &= \boldsymbol{\xi}_M(\mathbf{a}) \cdot (\boldsymbol{\xi}_M \circ \sigma^{2^I}(\mathbf{a})) = \prod_{x\in\mathbb{X}} \xi_x(a_x) \cdot \xi_x(a_{x+2^I}) \\ &= \prod_{x\in\mathbb{X}} \xi_x(a_x + a_{x+2^I}) = \prod_{x\in\mathbb{X}} \xi_x(d_x^I), \end{aligned} \tag{27}$$

where $\{d_x^I\}_{x\in\mathbb{X}}$ are IRDI. If d_x^I has distribution δ_x^I, then

$$\mathbb{E}_{\delta_x^I}[\xi_x(d_x^I)] = \delta_x^I\{0\} - \delta_x^I\{1\} = 1 - 2\delta_x^I\{1\} \underset{(*)}{\leq} 1 - 2\cdot\alpha^I = \frac{2\alpha^{-I} - 1}{2\alpha^{-I}}. \tag{28}$$

Here, $(*)$ is because μ has lower decay rate α, so $\delta_x^I\{1\} \geq \alpha^I$ (assuming $I \geq L$).

Thus,

$$\langle \mu, \boldsymbol{\chi} \circ \Phi^n \rangle \underset{(\ddagger)}{=} \mathbb{E}_\mu\left[\prod_{x\in\mathbb{X}} \xi_x(d_x^I)\right] \underset{(*)}{=} \prod_{x\in\mathbb{X}} \mathbb{E}_{\delta_x^I}[\xi_x(d_x^I)] \underset{(\dagger)}{\leq} \left(\frac{2\alpha^{-I} - 1}{2\alpha^{-I}}\right)^R.$$

Here, $(\ddagger)$ is by equation (27), $(*)$ is because $\{d_x^I\}_{x\in\mathbb{X}}$ are independent, and $(\dagger)$ is by equation (28) and because $\#(\mathbb{X}) = R$. Thus,

$$\begin{aligned} \log|\langle \mu, \boldsymbol{\chi} \circ \Phi^n \rangle| &\leq R \cdot [\log(2\alpha^{-I} - 1) - \log(2\alpha^{-I})] \underset{(*)}{\leq} -R \cdot \log'(2\alpha^{-I}) \\ &= \frac{-R}{2\alpha^{-I}} \underset{(\dagger)}{<} \frac{-\beta^I r}{2\alpha^{-I}} = -\frac{r}{2}(\alpha\beta)^I. \end{aligned}$$

where, $(*)$ is because log is a decreasing function, and $(\dagger)$ is by equation (26).

However, $\beta > 1/\alpha$, so $\alpha\beta > 1$. Thus, $\lim_{\mathbb{H}\ni h\to\infty} \log|\langle \mu, \boldsymbol{\chi} \circ \Phi^{h\cdot\langle\langle\boldsymbol{\chi}\rangle\rangle}\rangle| = -r/2 \lim_{I\to\infty}(\alpha\beta)^I = -\infty$. Hence, $\lim_{\mathbb{H}\ni h\to\infty} |\langle \mu, \boldsymbol{\chi} \circ \Phi^{h\cdot\langle\langle\boldsymbol{\chi}\rangle\rangle}\rangle| = 0$. $\square$

Suppose that $\mu \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ has IRDI; for any $n \in \mathbb{N}$, and all $m \in [1 \ldots 2^n]$, let $\delta_1^n, \ldots, \delta_{2^n}^n$ be the dyadic increment distributions, as before. Then μ has *upper decay rate* $\alpha \in (0, 1)$ if there are constants $L_1, K > 0$ such that, for all $n \geq L_1$, and all $m \in [1 \ldots 2^n]$, $\delta_m^n\{1\} \leq K \cdot \alpha^n$.

PROPOSITION 8.2. *If μ has IRDI with upper decay rate $\alpha < 1$, then $h(\mu) = 0$.*

Proof. Let $L_1, K > 0$ be as above. Assume without loss of generality that $K > 4$. Let $L_2 := (-\log_2(K) - 1)/\log_2(\alpha)$. Let $L := \max\{L_1, L_2\}$.

For any $n \in \mathbb{N}$, and $m \in [1 \ldots 2^n]$, let δ_m^n be as above. The entropy of δ_m^n is defined as

$$H(\delta_m^n) := -\delta_m^n\{0\} \log_2(\delta_m^n\{0\}) - \delta_m^n\{1\} \log_2(\delta_m^n\{1\}). \tag{29}$$

Claim 1. There exists $c_1 > 0$ such that, if $n > L$ and $m \in [1 \ldots 2^n]$, then $H(\delta_m^n) < c_1 n \cdot \alpha^n$.

Proof. We have $\alpha < 1$, so $\log_2(\alpha) < 0$. Thus, if $n \geq L_2$, then $n \log_2(\alpha) \leq L_2 \log_2(\alpha)$. Thus,

$$\begin{aligned} \log_2(K\alpha^n) &= \log_2(K) + n \log_2(\alpha) \leq \log_2(K) + L_2 \log_2(\alpha) \\ &= \log_2(K) - \log_2(K) - 1 = -1. \end{aligned} \tag{30}$$

Thus, $\delta_m^n\{1\} \underset{(*)}{\leq} K\alpha^n \underset{(\dagger)}{\leq} \frac{1}{2}$, where $(*)$ is because $n \geq L_1$. and $(\dagger)$ is by equation (30).

However, if $\delta_m^n\{1\} < \frac{1}{2}$ in (29), then $H(\delta_m^n)$ decreases as $\delta_m^n\{1\}$ decreases. Hence,

$$\begin{aligned} H(\delta_m^n) &\leq -K\alpha^n \log_2(K\alpha^n) - (1 - K\alpha^n) \log_2(1 - K\alpha^n) \\ &< K\alpha^n \underbrace{(nA - k)}_{(*)} + (1 - K\alpha^n) \cdot \underbrace{2K\alpha^n}_{(\dagger)} = K(nA + 2 - k - 2K\alpha^n) \cdot \alpha^n \\ &\underset{(\ddagger)}{<} KnA \cdot \alpha^n \underset{(\diamond)}{<} c_1 n \cdot \alpha^n. \end{aligned}$$

Here $(*)$ is the substitution $k := \log_2(K)$ and $A := -\log_2(\alpha)$; $(\dagger)$ is because, if ϵ is small, then $\log(1 - \epsilon) \approx -\epsilon$, thus, $-\log(1 - \epsilon) < 2\epsilon$; $(\ddagger)$ is because $2 - k - 2K\alpha^n < 0$ because $k > 2$ because we assume $K > 4$; $(\diamond)$ is where $c_1 := KA > 0$. $\diamond$

Let $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$ be a μ-random sequence, and fix $n > L$. To compute the conditional entropy $H(\mathbf{a}|_{(2^n \ldots 2^{n+1}]}|\mathbf{a}|_{[1 \ldots 2^n]})$, recall that, for all $m \in [1 \ldots 2^n]$, $a_{2^n+m} = a_m + d_m^n$. Thus,

$$\begin{aligned} H(\mathbf{a}|_{(2^n \ldots 2^{n+1}]}|\mathbf{a}|_{[1 \ldots 2^n]}) = H(d_1^n, d_2^n, \ldots, d_{2^n}^n) &\underset{(*)}{=} \sum_{m=1}^{2^n} H(\delta_m^n) \\ &\underset{(\dagger)}{<} 2^n \cdot c_1 n \alpha^n = c_1 n \cdot (2\alpha)^n \end{aligned} \tag{31}$$

where $(*)$ is because $d_1^n, d_2^n, \ldots, d_{2^n}^n$ are independent random variables with distributions

$\delta_1^n, \ldots, \delta_{2^n}^n$, and (†) is by Claim 1. Thus, for any $N > L$,

$$H(\mathbf{a}|_{[1\ldots 2^N]}|\mathbf{a}|_{[1\ldots 2^L]}) = \sum_{n=L}^{N-1} H(\mathbf{a}|_{(2^n\ldots 2^{n+1}]}|\mathbf{a}|_{[1\ldots 2^n]}) \underset{(*)}{<} \sum_{n=L}^{N-1} c_1 n \cdot (2\alpha)^n$$
$$< c_1 N \cdot (2\alpha)^L \sum_{n=0}^{N-L-1} (2\alpha)^n = c_1 N \cdot (2\alpha)^L \frac{(2\alpha)^{N-L} - 1}{2\alpha - 1}$$
$$\leq c_2 N \cdot (2\alpha)^N, \tag{32}$$

where $(*)$ is by equation (31), and where $c_2 \approx c_1/2\alpha - 1 > 0$ is another constant. Thus, if $H_0 := H(\mathbf{a}|_{[1\ldots 2^L]})$, then

$$H(\mathbf{a}|_{[1\ldots 2^N]}) = H(\mathbf{a}|_{[1\ldots 2^N]}|\mathbf{a}|_{[1\ldots 2^L]}) + H_0 \underset{(*)}{\leq} c_2 N \cdot (2\alpha)^N + H_0, \tag{33}$$

where $(*)$ is by equation‘(32). Thus,

$$h(\mu) = \lim_{M\to\infty} \frac{1}{M} H(\mathbf{a}|_{[1\ldots M]}) = \lim_{N\to\infty} \frac{1}{2^N} H(\mathbf{a}|_{[1\ldots 2^N]})$$
$$\underset{(*)}{\leq} \lim_{N\to\infty} \frac{c_2 N \cdot (2\alpha)^N + H_0}{2^N} \leq c_2 \lim_{N\to\infty} N\alpha^N \underset{0}{\overset{(\dagger)}{}},$$

where $(*)$ is by equation (33), and where (†) is because $|\alpha| < 1$. □

It remains to actually construct a measure with IRDI. Let $0 < \alpha < 1$. For any $n \in \mathbb{N}$, let ρ_n be the probability distribution on $\mathcal{A} = \mathbb{Z}_{/2}$ such that

$$\rho_n\{1\} = \alpha^n \quad \text{and} \quad \rho_n\{0\} = 1 - \alpha^n. \tag{34}$$

For each $n \in \mathbb{N}$, we will construct a random sequence $\mathbf{a}^n \in \mathcal{A}^{\mathbb{Z}}$ as follows. First, define $\mathbf{a}^0 := [\ldots 0000 \ldots]$. Now, suppose, inductively, that we have $\mathbf{a}^n$. Let $r_0^n, r_1^n, \ldots, r_{2^n-1}^n$ be a set of 2^n independent $\mathcal{A}$-valued, ρ_n-random variables. Let $\mathbf{r}^n \in \mathcal{A}^{\mathbb{Z}}$ be the random, 2^{n+1}-periodic sequence

$$\mathbf{r}^n := \left[\ldots, \underbrace{\overset{\text{zeroth coordinate}}{\overset{\downarrow}{0}}, 0, \ldots, 0}_{2^n}, r_0^n, r_1^n, \ldots, r_{2^n-1}^n, \underbrace{0, 0, \ldots, 0}_{2^n}, r_0^n, r_1^n, \ldots, r_{2^n-1}^n, \ldots\right],$$

and inductively define $\mathbf{a}^{n+1} := \mathbf{a}^n + \mathbf{r}^n$.

Let $\mu_n \in \mathcal{M}(\mathcal{A}^{\mathbb{Z}})$ be the distribution of $\mathbf{a}^n$, and let $\widetilde{\mu}_n := (1/2^n) \sum_{i=1}^{2^n} \sigma^i(\mu_n)$ be the stationary average of μ_n. Finally, let $\mu := \text{wk}^* - \lim_{n\to\infty} \widetilde{\mu}_n$.

Let μ_∞ be the probability distribution of the random sequence $\mathbf{a}^\infty := \sum_{n=1}^\infty \mathbf{r}^n$ (see Figure 3). Then $\mu_\infty = \text{wk}^* - \lim_{n\to\infty} \mu_n$, and loosely speaking, μ is the ‘σ-ergodic average’ of μ_∞. Thus, if $\mathbf{a}$ is a μ-random sequence, we can think of $\mathbf{a}$ as obtained by shifting $\mathbf{a}^\infty$ by a random amount. The following lemma describes the structure of $\mathbf{a}^\infty$.

LEMMA 8.3. *Let $M \in \mathbb{N}$ have binary expansion $M = \sum_{n=0}^\infty m_n 2^n$. For all $n \geq 0$, let $M_n := \sum_{i=0}^{n-1} m_i 2^i$. Then $a_M^\infty = \sum_{n=0}^\infty m_n \cdot r_{M_n}^n$.* □

For example, suppose that $M := 13 = 1+4+8$; then $m_0 = m_2 = m_3 = 1$ and $m_1 = 0$. Hence, $M_0 = 0$, $M_1 = M_2 = 1$, and $M_3 = 5$. Thus, $a_{13}^\infty = r_0^0 + r_1^2 + r_5^3$ (see Figure 3).

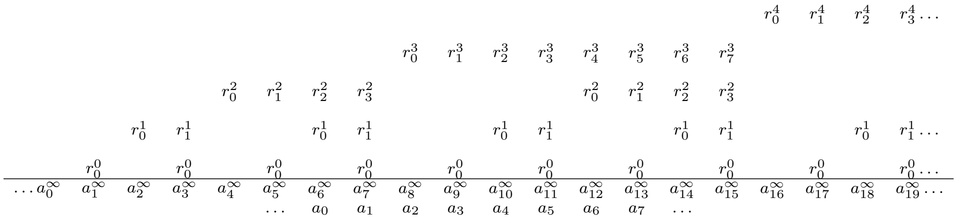

FIGURE 3. The construction of random sequence $\mathbf{a}^\infty$; the approximation of $\mathbf{a}$ as a random translate of $\mathbf{a}^\infty$.

Think of $\mathbf{a}^\infty$ as being generated by a process of 'duplication with error'. Let $\mathbf{w}^0 := [0]$ be a word of length 1. Suppose, inductively, that we have $\mathbf{w}^n = [w_1 w_2 \ldots w_{2^n-1}]$. Let $\widetilde{\mathbf{w}}^n := [\widetilde{w}_1\widetilde{w}_2 \ldots \widetilde{w}_{2^n-1}]$ be an 'imperfect copy' of $\mathbf{w}^n$: for each $m \in [0\ldots 2^n)$, $\widetilde{w}_m := w_m + r_m^n$, where $r_0^n, r_1^n \ldots, r_{2^n-1}^n$ are the independent ρ_n-distributed variables from before, which act as 'copying errors'. Let $\mathbf{w}^{n+1} := \mathbf{w}^n\widetilde{\mathbf{w}}^n$. Then $\mathbf{a}^\infty$ is the limit of $\mathbf{w}^n$ as $n\to\infty$.

PROPOSITION 8.4. *We propose that μ has IRDI, with upper and lower decay rate α.*

Proof. Let $\mathbf{a} \in \mathcal{A}^{\mathbb{Z}}$ be a μ-random sequence, and fix $N \in \mathbb{N}$. By construction, there is some $k \in \mathbb{Z}$ such that $\mathbf{a}$ looks like $\sigma^k(\mathbf{a}^\infty)$ in a neighbourhood around 0. To be precise,

$$\text{for all } m \in [0\ldots 2^{N+1}), \qquad a_m = a_{k+m}^\infty. \tag{35}$$

For example, in Figure 3, let $N = 2$, so that $2^N = 4$; suppose $k = 6$. Thus, $[a_0, a_1, \ldots, a_7] = [a_6^\infty, a_7^\infty, \ldots, a_{13}^\infty]$. Thus, $d_0^2 = a_4 - a_0 = a_{10}^\infty - a_6^\infty = r_2^3 - r_2^2 = r_2^3 + r_2^2$. We have the following more general claim.

Claim 1. Let $m \in [0\ldots 2^N)$.

(a) There is a set $S(m) := \{(n_0, m_0), (n_1, m_1), \ldots, (n_J, m_J)\}$ (for some $J \geq 0$), where $N = n_0 \leq n_1 \leq \cdots \leq n_J$, and where $m_j \in [0\ldots 2^{n_j})$ for all $j \in [0\ldots J]$, such that $d_m^N = r_{m_0}^{n_0} + r_{m_1}^{n_1} + \cdots + r_{m_J}^{n_J}$.

(b) If $m' \in [0\ldots 2^N)$, and $m' \neq m$, then $S(m') \cap S(m) = \emptyset$. □

Proof. Let $M := k + m$ and let $\widetilde{M} := k + m + 2^N$. If $M = \sum_{n=0}^\infty m_n 2^n$ and $\widetilde{M} = \sum_{n=0}^\infty \widetilde{m}_n 2^n$, then Lemma 8.3 says that

$$a_M^\infty = \sum_{n=0}^\infty m_n \cdot r_{M_n}^n \quad \text{and} \quad a_{\widetilde{M}}^\infty = \sum_{n=0}^\infty \widetilde{m}_n \cdot r_{\widetilde{M}_n}^n. \tag{36}$$

Let $N_1 \geq N$ be the smallest element of $[N\ldots\infty)$ such that $m_{N_1} = 0$. Hence, $m_n = 1$ for all $n \in [N\ldots N_1)$, and $m_{N_1} = 0$. Note that $\widetilde{M} = M + 2^N$, so binary expansions of M and $\widetilde{M}$ are related as follows:

(A) $m_n = \widetilde{m}_n$ for all $n \in [0\ldots N)$;

(B) thus, $\widetilde{M}_n = M_n$ for all $n \in [0\ldots N]$;

(C) if $m_N = 0$ then $\widetilde{m}_N = 1$. If $m_N = 1$, then $\widetilde{m}_N = 0$;

(D) $\widetilde{m}_n = 0$ for all $n \in [N\ldots N_1)$ (possibly an empty set), and $\widetilde{m}_{N_1} = 1$;

(E) $m_n = \widetilde{m}_n$ for all $n > N_1$.

Thus,

$$\begin{aligned} d_m^N &= a_{m+2^N} - a_m \underset{(*)}{=} a_{k+m+2^N}^{\infty} - a_{k+m}^{\infty} = a_{\widetilde{M}}^{\infty} + a_M^{\infty} \pmod 2 \\ &\underset{(\dagger)}{=} \sum_{n=0}^{\infty} (\widetilde{m}_n \cdot r_{\widetilde{M}_n}^n + m_n \cdot r_{M_n}^n) \underset{(\mathrm{ab})}{=} \sum_{n=N}^{\infty} (\widetilde{m}_n \cdot r_{\widetilde{M}_n}^n + m_n \cdot r_{M_n}^n) \\ &= \underbrace{r_{M_N}^N}_{(\mathrm{bc})} + \sum_{n=N+1}^{N_1-1} m_n \underbrace{r_{M_n}^n}_{(\mathrm{d})} + \underbrace{r_{\widetilde{M}_{N_1}}^{N_1}}_{(\mathrm{d})} + \sum_{n=N_1+1}^{\infty} \underbrace{m_n}_{(\mathrm{e})} \cdot (r_{\widetilde{M}_n}^n + r_{M_n}^n). \end{aligned} \tag{37}$$

Here, $(*)$ is by equation (35); $(\dagger)$ is by equation (36); (ab) is by (A) and (B); (bc) is by (B) and (C); (d) is by (D), and (e) is by (E).

Now, to see (a), let

$$S(m) := \{(n, m);\ r_m^n \text{ appears with non-zero coefficient in expression (37)}\}.$$

In particular, $r_{M_N}^N$ appears in (37), so $(n_0, m_0) := (N, M_N)$; thus, $n_0 = N$.

To see (b), suppose that $m < m'$; hence $m' = m + i$ for some $i \in [1 \dots 2^N)$.

Let $M' := M + i$ and $\widetilde{M}' := \widetilde{M} + i$. Suppose $M' = \sum_{n=0}^{\infty} m_n' 2^n$ and $\widetilde{M}' = \sum_{n=0}^{\infty} \widetilde{m}_n' 2^n$. Define M_n', $\widetilde{M}_n'$, and N_1' analogously. Then, an argument identical to (37) yields

$$d_{m'}^N = r_{M_N'}^N + \sum_{n=N+1}^{N_1'-1} m_n' r_{M_n'}^n + r_{\widetilde{M}_{N_1'}'}^{N_1'} + \sum_{n=N_1'+1}^{\infty} m_n' \cdot (r_{\widetilde{M}_n'}^n + r_{M_n'}^n). \tag{38}$$

Now, for all $n \in [N \dots \infty)$, $M_n' = M_n + i$ and $\widetilde{M}_n' = \widetilde{M}_n + i$ (because $i < 2^N$); thus, $r_{M_n'}^n = r_{M_n+i}^n \notin \{r_{M_n}^n, r_{\widetilde{M}_n}^n\}$ and $r_{\widetilde{M}_n'}^n = r_{\widetilde{M}_n+i}^n \notin \{r_{M_n}^n, r_{\widetilde{M}_n}^n\}$. Thus, every summand of (38) is distinct from every summand of (37), so $S(m') \cap S(m) = \emptyset$. $\diamond$

To see that the random variables $d_0^N, \dots, d_{2^N-1}^N$ are jointly independent, use Claim 1(a):

$$d_0^N = \sum_{(n,m)\in S(0)} r_m^n, \quad d_1^N = \sum_{(n,m)\in S(1)} r_m^n \quad , \dots, \quad d_{2^N-1}^N = \sum_{(n,m)\in S(2^N-1)} r_m^n.$$

The random variables $\{r_m^n; n \in \mathbb{N}, m \in [1 \dots 2^N]\}$ are independent, and Claim 1(b) says that $S(0), S(1) \dots, S(2^N - 1)$ are pairwise disjoint; thus $d_0^N, \dots, d_{2^N-1}^N$ are jointly independent.

Lower decay rate. $|\alpha| < 1$, so if N is sufficiently large (e.g. $N > L := -1/\log_2(\alpha)$), then $\alpha^N < 1/2$. Suppose that $d_m^N = r_{m_0}^{n_0} + r_{m_1}^{n_1} + \cdots + r_{m_J}^{n_J}$, as in Claim 1(a). For all $j \in [0 \dots J]$, let $P_j := \mathsf{Prob}\big(\sum_{i=j}^{J} r_{m_i}^{n_i} = 1\big)$. Thus,

$$\begin{aligned} \delta_m^N\{1\} &= P_0 \underset{(\dagger)}{=} \rho_N\{0\} \cdot P_1 + \rho_N\{1\} \cdot (1 - P_1) = (1 - \alpha^N) \cdot P_1 + \alpha^N \cdot (1 - P_1) \\ &= \alpha^N + (1 - 2\alpha^N) \cdot P_1 \underset{(*)}{\geq} \alpha^N. \end{aligned}$$

Here $(\dagger)$ is because Claim 1(a) says $n_0 = N$ and $(*)$ is because $1 - 2\alpha^N > 0$, because $\alpha^N < \frac{1}{2}$.

Upper decay rate. Let $K := 1/1-\alpha$. We claim that, for any N and m, $\delta_m^N\{1\} \le K\alpha^N$. As before, let $P_j := \mathsf{Prob}(\sum_{i=j}^{J} r_{m_i}^{n_i} = 1)$. For any $j \in [1 \ldots J)$, we have

$$P_j = (1-\alpha^{n_j}) \cdot P_{j+1} + \alpha^{n_j} \cdot (1-P_{j+1}) = P_{j+1} + (1-2P_{j+1})\alpha^{n_j} \le P_{j+1} + \alpha^{n_j}, \quad (39)$$

and $P_J = \alpha^{n_J}$. Hence,

$$\delta_m^N\{1\} = P_0 \underset{(*)}{\le} \alpha^{n_0} + \alpha^{n_1} + \cdots + \alpha^{n_J} \le \sum_{i=n_0}^{\infty} \alpha^i = \frac{\alpha^{n_0}}{1-\alpha} = K\alpha^{n_0} \underset{(\dagger)}{=} K\alpha^N.$$

Here, $(*)$ is obtained by applying (39) inductively, and $(\dagger)$ is because $n_0 = N$. □

Thus, if $1/\sqrt{2} < \alpha < 1$, then μ satisfies the conditions of Propositions 8.1 and 8.2, so μ is a zero-entropy, Lucas mixing measure. Hence, $1+\sigma$ asymptotically randomizes μ.

Acknowledgements. This research was partially supported by NSERC Canada, and was also supported by the kind hospitality of the Universidad de Chile during July 2003.

REFERENCES

[Bré99] P. Brémaud. *Markov Chains: Gibbs fields, Monte Carlo Simulation, and Queues*. Springer, New York, 1999.

[FMMN00] P. A. Ferrari, A. Maass, S. Martínez and P. Ney. Cesàro mean distribution of group automata starting from measures with summable decay. *Ergod. Th. & Dynam. Sys.* **20**(6) (2000), 1657–1670.

[Kit98] B. Kitchens. *Symbolic Dynamics: One-sided, Two-sided, and Countable State Markov Shifts*. Springer, New York, 1998.

[KS80] R. Kindermann and J. L. Snell. *Markov Random Fields and their Applications*. American Mathematical Society, Providence, RI, 1980.

[Lin84] D. Lind. Applications of ergodic theory and sofic systems to cellular automata. *Physica* D **10** (1984), 36–44.

[LM95] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, New York, 1995.

[MHM03] S. Martínez, A. Maass and B. Host. Uniform Bernoulli measure in dynamics of permutative cellular automata with algebraic local rules. *Discrete Continuous Dynam. Sys.* **9**(6) (2003), 1423–1446.

[MM98] A. Maass and S. Martínez. On Cesàro limit distribution of a class of permutative cellular automata. *J. Stat. Phys.* **90**(1–2) (1998), 435–452.

[MM99] A. Maass and S. Martínez. Time averages for some classes of expansive one-dimensional cellular automata. *Cellular Automata and Complex Systems*. Eds. Er. Goles and S. Martínez. Kluwer, Dordrecht, 1999, pp. 37–54.

[MMPY06] A. Maass, S. Martínez, M. Pivato and R. Yassawi. Asymptotic randomization of subgroup shifts by linear cellular automata. *Ergod. Th. & Dynam. Sys.* **26** (2006), this issue.

[Pet89] K. Petersen. *Ergodic Theory*. Cambridge University Press, New York, 1989.

[PY02] M. Pivato and R. Yassawi. Limit measures for affine cellular automata. *Ergod. Th. & Dynam. Sys.* **22**(4) (2002), 1269–1287.

[PY04] M. Pivato and R. Yassawi. Limit measures for affine cellular automata. II. *Ergod. Th. & Dynam. Sys.* **24**(6) (2004), 1961–1980.