SHIFTS OF FINITE TYPE WITH NEARLY FULL ENTROPY

RONNIE PAVLOV

ABSTRACT. For any fixed alphabet A, the maximum topological entropy of a $\mathbb{Z}^d$ subshift with alphabet A is obviously $\log|A|$. We study the class of nearest neighbor $\mathbb{Z}^d$ shifts of finite type which have topological entropy very close to this maximum, and show that they have many useful properties. Specifically, we prove that for any d, there exists $\delta_d > 0$ such that for any nearest neighbor $\mathbb{Z}^d$ shift of finite type X with alphabet A for which $(\log|A|) - h(X) < \delta_d$, X has a unique measure of maximal entropy μ. We also show that any such X is a measure-theoretic universal model in the sense of [25], that $h(X)$ is a computable number, that μ is measure-theoretically isomorphic to a Bernoulli measure, and that the support of μ has topologically completely positive entropy. Though there are other sufficient conditions in the literature (see [9], [15], [22]) which guarantee a unique measure of maximal entropy for $\mathbb{Z}^d$ shifts of finite type, this is (to our knowledge) the first such condition which makes no reference to the specific adjacency rules of individual letters of the alphabet.

1. INTRODUCTION

A dynamical system consists of a space X endowed with some sort of structure, along with a G-action (T_g) on the space for some group G which preserves that structure. (For our purposes, G will always be $\mathbb{Z}^d$ for some d.) Two examples are measurable dynamics, where X is a probability space and T_g is a measurable family of measure-preserving maps, and topological dynamics, where X is a compact space and the T_g are a continuous family of homeomorphisms. In each setup, when G is an amenable group, (this class of amenable groups includes $G = \mathbb{Z}^d$) there is an invaluable notion of entropy; measure-theoretic, or metric, entropy in the setup of measurable dynamics, and topological entropy in the setup of topological dynamics. (We postpone rigorous definitions of these and other terms until Section 2.) These two notions are related by the famous Variational Principle, which says that the topological entropy of a topological dynamical system is the supremum of the measure-theoretic entropy over all (T_g)-invariant Borel probability measures it supports. If in addition the system is taken to be expansive, then this supremum is achieved for at least one measure, and any such measures are called measures of maximal entropy.

Measures of maximal entropy are important objects of study, but are sometimes difficult to find and analyze. One particular class of topological dynamical systems for which measures of maximal entropy are well-understood are the one-dimensional shifts of finite type, or SFTs. A one-dimensional shift of finite type is defined by a finite set A, called the alphabet, and a set $\mathcal{F}$ of forbidden words, or finite strings of letters from A. The shift of finite type X induced by $\mathcal{F}$ then consists of all $x \in A^{\mathbb{Z}}$

1

(biinfinite strings of letters from A) which do not contain any of the forbidden words from $\mathcal{F}$. The space $A^{\mathbb{Z}}$ is endowed with the (discrete) product topology, and X inherits the induced topology, under which it is a compact metrizable space. The dynamics of a shift of finite type are always given by the $\mathbb{Z}$-action of integer shifts on sequences in X. Any one-dimensional SFT which satisfies a mild nontriviality condition called irreducibility has a unique measure of maximal entropy called the Parry measure, which is just a Markov chain with transition probabilities which can be algorithmically computed. For more on one-dimensional shifts of finite type and their measures of maximal entropy, see [21].

However, things become more complicated when one moves to multiple dimensions. A d-dimensional SFT is defined analogously to the one-dimensional case: specify the alphabet A and set of forbidden (d-dimensional) finite configurations $\mathcal{F}$, and define a shift of finite type X induced by $\mathcal{F}$ to be the set of all $x \in A^{\mathbb{Z}^d}$ (infinite d-dimensional arrays of letters from A) which do not contain any of the configurations from $\mathcal{F}$. The dynamics are now given by the $\mathbb{Z}^d$-action of all shifts by vectors in $\mathbb{Z}^d$. The easiest class of d-dimensional SFTs to work with are the nearest neighbor SFTs; a d-dimensional SFT X is called nearest neighbor if $\mathcal{F}$ consists entirely of adjacent pairs of letters, meaning that the legal points of X are defined purely by rules about which letters can legally sit next to which letters in each direction. A useful illustrative example of a nearest neighbor SFT is the d-dimensional hard-core shift $\mathcal{H}_d$, defined by $A = \{0, 1\}$, and $\mathcal{F}$ consisting of all configurations made of adjacent pairs of 1s (in each of the d cardinal directions). Then X consists of all ways of assigning 0 and 1 to each site in $\mathbb{Z}^d$ which do not contain two adjacent 1s.

It turns out that many questions regarding d-dimensional SFTs are extremely difficult or intractable. For instance, given only the alphabet A and forbidden list $\mathcal{F}$, the question of whether or not X is even empty is algorithmically undecidable! ([4], [28]) The structure of the measures of maximal entropy of multidimensional SFTs is similarly murky; it has been shown, for instance, that not even the strongest desirable topological mixing properties imply the existence of a unique measure of maximal entropy. ([8]) Even when the measure of maximal entropy is unique, its structure is not necessarily as simple as in the one-dimensional case: it may be a Bernoulli measure (for instance in the case when the SFT is all of $A^{\mathbb{Z}^d}$), but there also exist examples where it is not even measure-theoretically weak mixing. ([11])

There are existing conditions in the literature which guarantee a unique measure of maximal entropy, but some of these require quite strong restrictions on the adjacency rules defining X. For instance, it was first shown in [22] that if a nearest neighbor d-dimensional SFT X has an alphabet mostly comprised of letters which are safe symbols, meaning that they may legally sit next to any letter of the alphabet in any direction, then X has a unique measure of maximal entropy. It was later shown in [15] that if all letters of the alphabet of a nearest neighbor d-dimensional SFT X are only "nearly safe," meaning that they can legally sit next to a large proportion of the letters in A in any direction, then again X has a unique measure of maximal entropy. Both of these conditions, though useful, have two problems. Firstly, they make reference to combinatorial information about the adjacency rules themselves, rather than more coarse topological information about the system itself. Secondly, they are not very robust conditions; if one takes an SFT satisfying one of these conditions, and then adds a single letter to A with new adjacency rules

which do not allow it to sit next to a large portion of A, then the conditions are no longer satisfied.

The main focus of this paper is to define a more robust, less combinatorial, condition on multidimensional SFTs which guarantees existence of a unique measure of maximal entropy. Roughly speaking, our condition is similar to the one from [15], but rather than requiring every single letter of the alphabet to be "nearly safe," i.e. allowed to sit next to a large proportion of the letters in A in any direction, we require only that a large proportion of the letters of the alphabet are "nearly safe" in this sense. More specifically, call a nearest neighbor d-dimensional SFT ϵ-full if there exists a subset of the alphabet of size $(1-\epsilon)|A|$ consisting of letters which each have $(1-\epsilon)|A|$ legal neighbors in each cardinal direction. Our main result is that for small enough ϵ (dependent on d), every ϵ-full nearest neighbor d-dimensional SFT X has a unique measure of maximal entropy μ. We also prove several other desirable properties for such SFTs, such as showing that μ is measure-theoretically isomorphic to a Bernoulli measure.

Somewhat surprisingly, we also show that the ϵ-fullness condition is implied by a condition which makes no mention of adjacency rules whatsoever, namely having topological entropy very close to the log of the alphabet size. Specifically, for any ϵ, there exists δ for which any nearest neighbor d-dimensional SFT with entropy at least $(\log|A|) - \delta$ is ϵ-full, which shows that all of the nice properties of ϵ-full nearest neighbor d-dimensional SFTs are shared by nearest neighbor d-dimensional SFTs with entropy close enough to the log of the alphabet size.

We now briefly summarize the layout of the rest of the paper. In Section 2, we give definitions and basic preliminary results required for our arguments. In Section 3, we state and prove our main result. In Section 4, we show that that ϵ-fullness is unrelated to any existing topological mixing conditions in the literature, i.e. it does not imply and is not implied by any of these conditions. In the final two sections, we focus only on uniqueness of measure of maximal entropy (and not the other properties implied by our main result.) Section 5 compares our condition with some other sufficient conditions for uniqueness of measure of maximal entropy from the literature, and in Section 6, we discuss the optimal value of δ_d which guarantees uniqueness of measure of maximal entropy for nearest neighbor d-dimensional SFTs with entropy at least $(\log|A|) - \delta_d$. (We strongly suspect that the values we achieve in our main result are quite far from optimal.)

2. Definitions and preliminaries

We begin with some geometric definitions for $\mathbb{Z}^d$. Throughout, (e_i) represents the standard orthonormal basis of $\mathbb{Z}^d$. We use d to denote the ℓ_∞ metric to compare points in $\mathbb{Z}^d$: $d(s,t) := \|s-t\|_\infty = \sum_i |s_i - t_i|$. For any sets $S, T \subseteq \mathbb{Z}^d$, we define $d(S,T) := \min_{s \in S, t \in T} d(s,t)$. We say that two sites $s, t \in \mathbb{Z}^d$ are **adjacent** if $d(s,t) = 1$. We also refer to adjacent sites as **neighbors**, and correspondingly define the neighbor set N_s of any $s \in \mathbb{Z}^d$ as the set of sites in $\mathbb{Z}^d$ adjacent to s.

This notion of adjacency gives $\mathbb{Z}^d$ a graph structure, and the notions of **paths** and **connected subsets** of $\mathbb{Z}^d$ are defined with this graph structure in mind. The **outer boundary** of a set $S \subseteq \mathbb{Z}^d$, written ∂S, is the set of all $t \in \mathbb{Z}^d \setminus S$ adjacent to some $s \in S$. The **inner boundary** of S, written $\underline{\partial} S$, is the set of all $s \in S$ adjacent to some $t \in \mathbb{Z}^d \setminus S$. A **closed contour surrounding** S is any set of the form ∂T for a connected set $T \subseteq \mathbb{Z}^2$ containing S.

Definition 2.1. For any finite alphabet A, the $\mathbb{Z}^d$ **full shift** over A is the set $A^{\mathbb{Z}^d}$, which is viewed as a compact topological space with the (discrete) product topology.

Definition 2.2. A **configuration** over A is a member of A^S for some finite $S \subset \mathbb{Z}^d$, which is said to have **shape** S. The set $\bigcup_{S \subset \mathbb{Z}^d, |S| < \infty} A^S$ of all configurations over A is denoted by A^*. When $d = 1$, a configuration whose shape is an interval of integers is sometimes referred to as a **word**.

Definition 2.3. For two configurations $v \in A^S$ and $w \in A^T$ with $S \cap T = \varnothing$, the **concatenation** of v and w, written vw, is the configuration on $S \cup T$ defined by $(vw)|_S = v$ and $(vw)|_T = w$.

Definition 2.4. The $\mathbb{Z}^d$**-shift action**, denoted by $\{\sigma_v\}_{v \in \mathbb{Z}^d}$, is the $\mathbb{Z}^d$-action on a full shift $A^{\mathbb{Z}^d}$ defined by $(\sigma_v x)(u) = x(u + v)$ for $u, v \in \mathbb{Z}^d$.

Definition 2.5. A $\mathbb{Z}^d$ **subshift** is a closed subset of a full shift $A^{\mathbb{Z}^d}$ which is invariant under the shift action.

Each σ_v is a homeomorphism on any $\mathbb{Z}^d$ subshift, and so any $\mathbb{Z}^d$ subshift, when paired with the $\mathbb{Z}^d$-shift action, is a topological dynamical system. An alternate definition for a $\mathbb{Z}^d$ subshift is in terms of disallowed configurations; for any set $\mathcal{F} \subset A^*$, one can define the set $X(\mathcal{F}) := \{x \in A^{\mathbb{Z}^d} : x|_S \notin \mathcal{F} \ \forall \text{ finite } S \subset \mathbb{Z}^d\}$. It is well known that any $X(\mathcal{F})$ is a $\mathbb{Z}^d$ subshift, and all $\mathbb{Z}^d$ subshifts are representable in this way. All $\mathbb{Z}^d$ subshifts are assumed to be nonempty in this paper.

Definition 2.6. A $\mathbb{Z}^d$ **shift of finite type (SFT)** is a $\mathbb{Z}^d$ subshift equal to $X(\mathcal{F})$ for some finite $\mathcal{F}$. If $\mathcal{F}$ is made up of pairs of adjacent letters, i.e. if $\mathcal{F} \subseteq \bigcup_{i=1}^{d} A^{\{0, \vec{e}_i\}}$, then X is called a **nearest neighbor (or n.n.)** $\mathbb{Z}^d$ SFT.

Definition 2.7. The **language** of a $\mathbb{Z}^d$ subshift X, denoted by $L(X)$, is the set of all configurations which appear in points of X. For any finite $S \subset \mathbb{Z}^d$, $L_S(X) := L(X) \cap A^S$, the set of configurations in the language of X with shape S.

Definition 2.8. A configuration $u \in A^S$ is **locally admissible** for a $\mathbb{Z}^d$ subshift $X = X(\mathcal{F})$ if $x|_T \notin \mathcal{F}$ for all $T \subseteq S$. In other words, U is locally admissible if it does not contain any of the forbidden configurations for X. We denote by $LA(X)$ the set of all locally admissible configurations for X, and by $LA_S(X)$ the set $LA(X) \cap A^S$.

We note that $LA(X) \subseteq L(X)$, but the reverse inclusion is not necessarily true. (In general, a configuration could be locally admissible, but attempting to complete it to all of $\mathbb{Z}^d$ always leads to a forbidden configuration.)

Definition 2.9. For any $\mathbb{Z}^d$ subshift and configuration $w \in L_S(X)$, the **cylinder set** $[w]$ is the set of all $x \in X$ with $x|_S = w$. We define the **configuration set** $\langle w \rangle$ to be the set of all configurations u in $L(X)$ with shape containing S for which $u|_S = w$. For any set C of configurations, we use the shorthand notations $[C]$ and $\langle C \rangle$ to refer to $\bigcup_{w \in C}[w]$ and $\bigcup_{w \in C}\langle w \rangle$ respectively.

In the following definition and hereafter, for any integers $m < n$, $[m, n]$ denotes the set of integers $\{m, m+1, \ldots, n\}$.

Definition 2.10. The **topological entropy** of a $\mathbb{Z}^d$ subshift X is

$$h(X) := \lim_{n_1,\ldots,n_d \to \infty} \frac{1}{\prod_{i=1}^{d} n_i} \log |L_{\prod_{i=1}^{d}[1,n_i]}(X)|.$$

We will also need several measure-theoretic definitions.

Definition 2.11. For any measures μ, ν on the same finite probability space X, the **total variational distance** between μ and ν is

$$d(\mu,\nu) := \frac{1}{2} \sum_{x \in X} |\mu(\{x\}) - \nu(\{x\})| = \max_{A \subseteq X} |\mu(A) - \nu(A)|.$$

Definition 2.12. For any μ, ν measures on probability spaces X and Y respectively, a **coupling** of μ and ν is a measure λ on $X \times Y$ whose marginals are μ and ν; i.e. $\lambda(A \times Y) = \mu(A)$ for all measurable $A \subseteq X$ and $\lambda(X \times B) = \nu(B)$ for all measurable $B \subseteq Y$. If $X = Y$, then an **optimal coupling** of μ and ν is a coupling λ which minimizes the probability $\lambda(\{(x,y) \ : \ x \neq y\})$ of disagreement.

The connection between Definitions 2.11 and 2.12 is the well-known fact that for any μ and ν on the same finite probability space, optimal couplings exist, and the probability of disagreement for an optimal coupling is equal to the total variational distance $d(\mu,\nu)$.

From now on, any measure μ on a full shift $A^{\mathbb{Z}^d}$ is assumed to be a Borel probability measure which is shift invariant, i.e. $\mu(\sigma_v C) = \mu(C)$ for any measurable C and $v \in \mathbb{Z}^d$.

Definition 2.13. For any measure μ on a full shift $A^{\mathbb{Z}^d}$, the **measure-theoretic entropy** of μ is

$$h(\mu) := \lim_{n_1,\ldots,n_d \to \infty} \frac{1}{\prod_{i=1}^{d} n_i} \sum_{w \in A^{\prod_{i=1}^{d}[1,n_i]}} -\mu([w]) \log \mu([w]),$$

where terms with $\mu([w]) = 0$ are omitted from the sum.

In Definitions 2.10 and 2.13, a subadditivity argument shows that the limits can be replaced by infimums; i.e. for any $n_1, \ldots, n_d$, $h(X) \leq \dfrac{1}{\prod_{i=1}^{d} n_i} \log |L_{\prod_{i=1}^{d}[1,n_i]}(X)|$

and $h(\mu) \leq \dfrac{1}{\prod_{i=1}^{d} n_i} \displaystyle\sum_{w \in A^{\prod_{i=1}^{d}[1,n_i]}} -\mu([w]) \log \mu([w]).$

Definition 2.14. For any $\mathbb{Z}^d$ subshift X, a **measure of maximal entropy** on X is a measure μ with support contained in X for which $h(\mu) = h(X)$.

The classical variational principle (see [23] for a proof) says that for any $\mathbb{Z}^d$ subshift X, $\sup_\mu h(\mu) = h(X)$, where the supremum, taken over all shift-invariant Borel probability measures whose support is contained in X, is achieved. Therefore, any $\mathbb{Z}^d$ subshift has at least one measure of maximal entropy. In the specific case when X is a nearest neighbor $\mathbb{Z}^d$ SFT, much is known about the conditional distributions of a measure of maximal entropy.

Definition 2.15. A measure μ on $A^{\mathbb{Z}^d}$ is called a **Markov random field** (or **MRF**) if, for any finite $S \subset \mathbb{Z}^d$, any $\eta \in A^S$, any finite $T \subset \mathbb{Z}^d \setminus S$ s.t. $\partial S \subseteq T$, and any $\delta \in A^T$ with $\mu([\delta]) \neq 0$,

$$\mu(x|_S = \eta \mid x|_{\partial S} = \delta|_{\partial S}) = \mu(x|_S = \eta \mid x|_T = \delta).$$

Informally, μ is an MRF if, for any finite $S \subset \mathbb{Z}^d$, the sites in S and the sites in $\mathbb{Z}^d \setminus (S \cup \partial S)$ are μ-conditionally independent given the sites on ∂S.

Proposition 2.16. ([9], Proposition 1.20) *For any nearest neighbor $\mathbb{Z}^d$ SFT X, all measures of maximal entropy for X are MRFs, and for any such measure μ and any finite shape $S \subseteq \mathbb{Z}^d$, the conditional distribution of μ on S given any $\delta \in L_{\partial S}(X)$ is uniform over all configurations $x \in L_S(X)$ such that the configuration y defined by $y|_S = x$ and $y|_{\partial S} = \delta$ is locally admissible in X.*

In other words, given any nearest neighbor $\mathbb{Z}^d$ SFT X, there is a unique set of conditional distributions that any measure of maximal entropy μ must match up with. However, this does not uniquely determine μ, as there could be several different measures with the same conditional distributions. For any $\delta \in L_{\partial S}(X)$ as in Proposition 2.16, we denote by Λ^δ the common uniform conditional distribution on S given δ that every measure of maximal entropy μ must have.

Next, we define some useful conditions for SFTs and measures supported on SFTs from the literature, many of which we will be able to prove for nearest neighbor SFTs which are ϵ-full for small enough ϵ.

Definition 2.17. A **measure-theoretic factor map** between two measures μ on $A^{\mathbb{Z}^d}$ and μ' on $B^{\mathbb{Z}^d}$ is a measurable function $F : A^{\mathbb{Z}^d} \to B^{\mathbb{Z}^d}$ which commutes with the shift action (i.e. $F(\sigma_v x) = \sigma_v F(x)$ for all $x \in A^{\mathbb{Z}^d}$) and for which $\mu'(C) = \mu(F^{-1}C)$ for all measurable $C \subseteq B^{\mathbb{Z}^d}$.

Definition 2.18. A **measure-theoretic isomorphism** is a measure-theoretic factor map which is bijective between sets of full measure in the domain and range.

Definition 2.19. A measure μ on $A^{\mathbb{Z}^d}$ is **ergodic** if any measurable set C which is shift-invariant, meaning $\mu(C \triangle \sigma_v C) = 0$ for all $v \in \mathbb{Z}^d$, has measure 0 or 1. Equivalently, μ is ergodic iff for any configurations u, u' over A,

$$\lim_{n \to \infty} \frac{1}{(2n+1)^d} \sum_{v \in [-n,n]^d} \mu([u] \cap \sigma_v[u']) = \mu([u])\mu([u']).$$

Definition 2.20. A measure μ on $A^{\mathbb{Z}^d}$ is **measure-theoretically strong mixing** if for any configurations u, u' over A and any sequence $v_n \in \mathbb{Z}^d$ for which $\|v_n\|_\infty \to \infty$,

$$\lim_{n \to \infty} \mu([u] \cap \sigma_{v_n}[u']) = \mu([u])\mu([u']).$$

Definition 2.21. A measure μ on $A^{\mathbb{Z}^d}$ is **Bernoulli** if it is measure-theoretically isomorphic to a measure on some $B^{\mathbb{Z}^d}$ which is independent and identically distributed over the sites of $\mathbb{Z}^d$.

There is an entire hierarchy of measure-theoretic mixing conditions, all of which are useful isomorphism invariants of measures. We will not spend much space here discussing this hierarchy, because Bernoullicity is the strongest of all of them, and we will verify that the unique measure of maximal entropy of ϵ-full nearest neighbor $\mathbb{Z}^d$ SFTs is Bernoulli for sufficiently small ϵ. For more information, see [27].

Definition 2.22. A **topological factor map** between two $\mathbb{Z}^d$ subshifts X and X' is a surjective continuous function $F : X \to X'$ which commutes with the shift action (i.e. $F(\sigma_v x) = \sigma_v F(x)$ for all $x \in A^{\mathbb{Z}^d}$).

In fact, topological factor maps on $\mathbb{Z}^d$ subshifts have a very specific form.

Definition 2.23. A **sliding block code** is a function ϕ defined on a subshift X induced by an integer $n > 0$ and a function Φ on $A^{[-n,n]^d}$ by the rule $(\phi x)(v) = \Phi(x|_{[-n,n]^d+v})$ for all $v \in \mathbb{Z}^d$.

The well-known Curtis-Hedlund-Lyndon theorem states that any topological factor map on a $\mathbb{Z}^d$ is a sliding block code. (See [21] for a proof for $d = 1$, which extends easily to the $d > 1$ case.)

Definition 2.24. A **topological conjugacy** is a bijective topological factor map.

Definition 2.25. A $\mathbb{Z}^d$ SFT X is **topologically mixing** if for any configurations $u, u' \in L(X)$, there exists n so that $[u] \cap \sigma_v[u'] \neq \varnothing$ for any $v \in \mathbb{Z}^d$ with $\|v\|_\infty > n$.

Definition 2.26. A $\mathbb{Z}^d$ SFT X is **block gluing** if there exists n so that for any configurations $u, u' \in L(X)$ with shapes rectangular prisms and any v for which u and $\sigma_v u$ are separated by distance at least n, $[u] \cap \sigma_v[u'] \neq \varnothing$.

Definition 2.27. A $\mathbb{Z}^d$ SFT X has the **uniform filling property** or **UFP** if there exists n such that for any configuration $u \in L(X)$ with shape a rectangular prism $R = \prod[a_i, b_i]$, and any point $x \in X$, there exists $y \in X$ such that $y|_R = u$, and $y|_{\mathbb{Z}^d \setminus \prod[a_i-n, b_i+n]} = x|_{\mathbb{Z}^d \setminus \prod[a_i-n, b_i+n]}$.

Again, all of these conditions are invariant under topological conjugacy. Note the subtle difference in the definitions: Definitions 2.26 and 2.27 require a uniform distance which suffices to mix between all pairs of configurations of a certain type, whereas Definition 2.25 allows this distance to depend on the configurations. In general, topological mixing is not a very strong condition for $\mathbb{Z}^d$ SFTs; usually a stronger condition involving a uniform mixing length such as block gluing or UFP is necessary to prove interesting results. (See [7] for a detailed description of a hierarchy of topological mixing conditions for $\mathbb{Z}^d$ SFTs.)

The final topological properties that we will show for ϵ-full SFTs for small ϵ do not quite fit into the topological mixing hierarchy. The first involves modeling measure-theoretic dynamical systems within a subshift.

Definition 2.28. A $\mathbb{Z}^d$ subshift X is a **measure-theoretic universal model** if for any $\mathbb{Z}^d$ ergodic measure-theoretic dynamical system $(Y, \mu, (T_v)_{v \in \mathbb{Z}^d})$, there exists a measure ν on X so that $(X, \nu, (\sigma_v)_{v \in \mathbb{Z}^d}) \cong (Y, \mu, (T_v)_{v \in \mathbb{Z}^d})$.

It was shown in [25] that any $\mathbb{Z}^d$ SFT with the UFP is a measure-theoretic universal model.

Definition 2.29. A $\mathbb{Z}^d$ SFT X has **topologically completely positive entropy** if every topological factor map from X is either a conjugacy or has image with topological entropy 0. In other words, X has no nontrivial positive entropy factors.

This is tenuously connected to mixing conditions by analogy to the measure-theoretic case: the corresponding condition for measure-theoretic factor maps is called measure-theoretic completely positive entropy, which in one dimension is also known as the K-property. The K-property is a very strong measure-theoretic mixing condition in one dimension which is implied by Bernoullicity, and which implies most other mixing properties. Unfortunately this analogy does not carry over to the topological setup; it was shown in [6] that topological completely positive entropy, even in one dimension, does not even imply topological mixing.

We also need a definition from computability theory:

Definition 2.30. A real number α is **computable in time** $f(n)$ if there exists a Turing machine which, on input n, outputs a pair (p_n, q_n) of integers such that $|\frac{p_n}{q_n} - \alpha| < \frac{1}{n}$, and if this procedure takes less than $f(n)$ operations for every n. We say that α is **computable** if it is computable in time $f(n)$ for some function $f(n)$.

Informally speaking, a real number α is computable if it is possible to give a finite description of α which allows someone to reconstruct as many digits of the decimal expansion of α as desired. For instance, e is computable since we can describe it as the sum of the reciprocals of the factorials of nonnegative numbers. All algebraic numbers are computable, but there are many more computable numbers than algebraic (though still only countably many.) For an introduction to computability theory, see [19].

The relationship between multidimensional symbolic dynamics and computability theory has been the subject of much work in recent years, but is still not completely understood. One foundational result is from [16], where it is shown that a real number is the entropy of some $\mathbb{Z}^d$ SFT for any $d > 1$ if and only if it has a property called right recursive enumerability, which is strictly weaker than computability and which we do not define here. It is also shown in [16] that if a $\mathbb{Z}^d$ SFT has the uniform filling property, then its entropy is in fact computable.

We conclude this section by finally defining ϵ-fullness of a nearest neighbor $\mathbb{Z}^d$ SFT, which we will show in the next section implies many useful properties, including uniqueness of the measure of maximal entropy.

Definition 2.31. For any $\epsilon > 0$, we say that a nearest neighbor $\mathbb{Z}^d$ SFT X with alphabet A is ϵ**-full** if A can be partitioned into sets G (good letters) and B (bad letters) with the properties that

(i) $|G| > (1 - \epsilon)|A|$

(ii) $\forall g \in G, i \in [1, d], \epsilon \in \{\pm 1\}$, the set of legal neighbors of g in the ϵe_i-direction has cardinality greater than $(1 - \epsilon)|A|$.

Surprisingly, the ϵ-fullness property is closely related to a simpler property which can be stated without any reference to adjacency rules, i.e. having entropy close to the log of the alphabet size.

Theorem 2.32. *For any $\epsilon > 0$ and d, there exists a $\delta = \delta(\epsilon, d)$ so that for a $\mathbb{Z}^d$ nearest neighbor SFT X with alphabet A,*

$$h(X) > (\log |A|) - \delta \implies X \text{ is } \epsilon - full.$$

Also, for any $\delta > 0$ and d, there exists an $\epsilon = \epsilon(\delta, d)$ so that

$$X \text{ is } \epsilon - full \implies h(X) > (\log |A|) - \delta.$$

Proof. Fix any d and $\epsilon > 0$, and suppose that X is not ϵ-full. This implies that if we define B to be the set of $b \in A$ for which there exists $i \in [1, d]$ and $\epsilon \in \{\pm 1\}$ so that there are at least $\epsilon|A|$ letters which cannot follow b in the ϵe_i-direction, then $|B| > \epsilon|A|$. (Otherwise, the complement of B could play the role of G in the definition of ϵ-full, and show that X is ϵ-full.)

Then, there exist $\epsilon \in \{\pm 1\}$, $i \in [1, d]$ and a set $B_i \subset B$ with $|B_i| > \frac{\epsilon}{2d}|A|$ so that for each $b \in B_i$, there are at least $\epsilon|A|$ letters which cannot follow b in the e_i-direction. This implies that there are at least $|B_i|\epsilon|A| > \frac{\epsilon^2}{2d}|A|^2$ words with shape

$\{0\} \cup \{\epsilon e_i\}$ which are not in $L(X)$, and so $|L_{\{0\}\cup\{\epsilon e_i\}}(X)| < |A|^2 \left(1 - \frac{\epsilon^2}{2d}\right)$. Then

$$h(X) \leq \frac{1}{2} \log |L_{\{0\}\cup\{e_i\}}(X)| < \log |A| - \log \sqrt{\frac{2d}{2d - \epsilon^2}},$$

and so taking $\delta(\epsilon, d) = \log \sqrt{\frac{2d}{2d-\epsilon^2}}$ proves the first half of the theorem.

Now fix any $\epsilon > 0$, and suppose that X is ϵ-full. Then, for any n, we bound from below the size of $L_{[1,n]^d}(X)$. Construct configurations in the following way: order the sites in $[1,n]^d$ lexicographically. Then fill the first site in $[1,n]^d$ with any G-letter. Fill the second site with any G-letter which can legally appear next to the first placed G-letter, and continue in this fashion, filling the sites in order with G-letters, and each time placing any legal choice given the letters which have already been placed. Then regardless of what has been placed before, at each step we will have $|A|(1-(d+1)\epsilon)$ choices. This is because the set of G-letters which can be legally placed at a site depends only on what has been placed at its neighbors, and by the definition of G, each neighboring letter rules out at most $|A|\epsilon$ choices. Therefore, regardless of what has been placed, there are always at least $|A|(1-d\epsilon)$ letters of A which can legally be placed at a site, and since $|G| > |A|(1-\epsilon)$, at least $|A|(1-(d+1)\epsilon)$ of these are in G. We can create $(|A|(1-(d+1)\epsilon))^{n^d}$ configurations in this way, and each one is in $L(X)$; by continuing to append G-letters, each could be extended to all of $\mathbb{Z}^d$ to create a point of X. This means that for any n,

$$|L_{[1,n]^d}(X)| > (|A|(1-(d+1)\epsilon))^{n^d},$$

which implies that $h(X) \geq \log(|A|(1-(d+1)\epsilon)) > \log |A| + \log(1-(d+1)\epsilon)$. Then taking $\epsilon(\delta, d) = \frac{1-e^\delta}{d+1}$ proves the second half of the theorem.

$\square$

We will informally refer to nearest neighbor $\mathbb{Z}^d$ SFTs with topological entropy close to $\log |A|$ as having "nearly full entropy." By Theorem 2.32, this condition is equivalent to being ϵ-full for small ϵ, and so we will often use the terms "ϵ-full for small ϵ" and "nearly full entropy" somewhat interchangeably in the sequel.

Before stating and proving our main results, here are some examples of nearest neighbor $\mathbb{Z}^d$ SFTs which are ϵ-full for small ϵ.

Example 2.33. Take X to have alphabet $A = \{0, 1, \ldots, n\}$, and the only adjacency rule is that any neighbor of a 0 must also be a 0. Then X is just the union of the full shift on $\{1, \ldots, n\}$ and a fixed point of all 0s. Clearly X is ϵ-full for $\epsilon < \frac{1}{n+1}$.

Example 2.34. Take X to have alphabet $A = \{0, 1, \ldots, n\}$, and the only adjacency rule is that a 0 can only appear above and below other 0s. Then X consists of points whose columns are either sequences on $\{1, \ldots, n\}$ (with no restrictions on which rows can appear) or all 0s. Again, clearly X is ϵ-full for $\epsilon < \frac{1}{n+1}$.

Example 2.35. Take X to be the full shift on $A = \{0, 1, \ldots, n\}$. Then trivially X is ϵ-full for any $\epsilon > 0$. For the purposes of this example though, think of A as being partitioned into $G = \{1, \ldots, n\}$ and $B = \{0\}$, which would show that X is ϵ-full for $\epsilon < \frac{1}{n+1}$.

These examples illustrate the different ways in which B-letters can coexist with G-letters, which is the unknown quantity in the description of ϵ-full SFTs. In Examples 2.33 and 2.34, the existence of a B-letter forces the existence of an infinite

component of B-letters. It turns out that in such examples, B-letters are rare in "most" configurations of X; in particular, they have zero measure for any measure of maximal entropy. In contrast, Example 2.35 clearly has a unique Bernoulli measure of maximal entropy, whose support contains all configurations (including those with B-letters). However, for large n, in "most" configurations the B-letters appear with the small frequency $\frac{1}{n}$. The dichotomy is that for measures of maximal entropy, either B-letters can only appear within infinite clusters of B-letters which appear with probability 0, or B-letters "coexist peacefully" with G-letters, and appear in most configurations with frequency less than ϵ.

3. Properties of ϵ-full/nearly full entropy SFTs

The goal of this section is to prove the following theorem, which is the main result of this paper.

Theorem 3.1. *For any d, there exists $\epsilon_d > 0$ so that any ϵ_d-full nearest neighbor $\mathbb{Z}^d$ SFT X has the following properties:*
  (A) *X has a unique measure of maximal entropy μ*
  (B) *$h(X)$ is computable in time $|A|^{(3d \log |A|n)^d (1+o(1))}$*
  (C) *X is a measure-theoretic universal model*
  (D) *μ is Bernoulli*
  (E) *The support of μ has topologically completely positive entropy.*

By Theorem 2.32, all such properties also hold for nearest neighbor $\mathbb{Z}^d$ SFTs with entropy close to the logarithm of their alphabet size:

Corollary 3.2. *For any d, there exists δ_d so that any nearest neighbor $\mathbb{Z}^d$ SFT X with alphabet A for which $h(X) > (\log |A|) - \delta_d$ has properties (A)-(E) from Theorem 3.1.*

We begin by proving some general facts about ϵ-full nearest neighbor $\mathbb{Z}^d$ SFTs for small ϵ.

Lemma 3.3. *If X is ϵ-full, then for any locally admissible configuration w with shape S with $w|_{\partial S} \in G^{\partial S}$ and any $p \in \mathbb{Z}^d \setminus S$, there exists a subset G' of G with cardinality greater than $|A|(1 - (2d+1)\epsilon)$ so that for any $g' \in G'$, the concatenation wg' is locally admissible.*

Proof. Define $N = N_p \cap S$, and note that $N \subseteq \partial S$. For any $a \in A$, as long as a can appear legally next to each of the at most $2d$ letters in $w|_N$, the concatenation wa is locally admissible. Each letter in $w|_N$ is a G-letter, and so by ϵ-fullness, for each $t \in N$, the set of letters which can appear legally next to $w(t)$ has cardinality at least $|A|(1 - \epsilon)$, and so there are at least $|A|(1 - 2d\epsilon)$ letters in A for which wa is locally admissible. Since $|G| > |A|(1 - \epsilon)$, at least $|A|(1 - (2d+1)\epsilon)$ of these letters are in G, and we are done. $\square$

Corollary 3.4. *If X is ϵ-full and $|A|(1 - (2d+1)\epsilon) \geq 1$, then any locally admissible configuration w with shape S with $w|_{\partial S}$ consisting only of G-letters is also globally admissible. In particular, w can be extended to a point of X by appending only G-letters to w.*

Proof. Suppose w is a locally admissible configuration with shape S s.t. $w|_{\underline{\partial}S}$ consists only of G-letters, and arbitrarily order the sites in $\mathbb{Z}^d \setminus S$ as s_i, $i \in \mathbb{N}$. We claim that for any n, there exists a locally admissible configuration w_n with shape $S \cup \bigcup_{i=1}^{n}\{s_i\}$ such that $w_n|_S = w$, $w_n|_{(\underline{\partial}S) \cup \bigcup_{i=1}^{n}\{s_i\}}$ consists only of G-letters, and each w_n is a subconfiguration of w_{n+1}.

The proof is by induction: the existence of w_1 is obvious by applying Lemma 3.3 to w and s_1, and for any n, if we assume the existence of w_n, the existence of w_{n+1} comes from applying Lemma 3.3 to w_n and s_{n+1}, along with the observation that clearly $\underline{\partial}(S \cup \bigcup_{i=1}^{n}\{s_i\}) \subseteq (\underline{\partial}S) \cup \bigcup_{i=1}^{n}\{s_i\}$.

Then the w_n approach a limit point $x \in G^{\mathbb{Z}^d}$, which is in X since each w_n was locally admissible. Since w was a subconfiguration of each w_n, it is a subconfiguration of x, and so $w \in L(X)$.

$\square$

The next lemma will be fundamental to almost all future arguments, and deals with the conditional measure w.r.t. a measure of maximal entropy μ of a configuration consisting only of B-letters given a boundary configuration. We would like to be able to say that all such configurations are fairly unlikely, but this depends on the boundary. For instance, in the SFT of Example 2.34, conditioning on a boundary $\delta \in A^{\partial[-n,n]^2}$ for which $\delta(0, -n) = \delta(0, n) = 0$ actually forces an entire column of 0s between them! For this reason, we for now deal only with the case where we condition on a boundary consisting only of G-letters.

Lemma 3.5. *For any $\epsilon < \frac{1}{4d+6}$ and any ϵ-full nearest neighbor $\mathbb{Z}^d$ SFT X, any connected set $C \subseteq \mathbb{Z}^d$, any set S containing ∂C and contained in $C \cup \partial C$, any $\eta \in G^S$, any set T contained in $(C \cup \partial C) \setminus S$, and any measure of maximal entropy μ on X,*
$$\mu(x|_T \in B^T \mid [\eta]) < N^{-|T|},$$
where N is $\lfloor \frac{1}{2}(\epsilon^{-1} - 4d - 4) \rfloor$.

Proof. Consider any such ϵ, X, C, S, T, and η, and define $N = \lfloor \frac{1}{2}(\epsilon^{-1} - 4d - 4) \rfloor$; since $\epsilon < \frac{1}{4d+6}$, $N \geq 1$ and the inequality we wish to prove is nontrivial. The lemma is trivial if $|A|\epsilon < 1$, since in this case B is empty by ϵ-fullness of X. We therefore assume that $|A|\epsilon \geq 1$. By ϵ-fullness of X, $|G| > |A|(1-\epsilon) \geq |A|(1-2\epsilon)+1$, and by definition of N, $1 - 2\epsilon > 2\epsilon(N + 2d + 1)$. Therefore, $|G| > 2|A|\epsilon(N + 2d + 1) + 1 \geq 2\lceil |A|\epsilon(N + 2d + 1)\rceil$. We can then partition G into two pieces, call them G_I and G_B, each of size at least $|A|\epsilon(N + 2d + 1)$, and fix any orderings on the elements of G_I, G_B, and B.

Consider any configuration $u \in L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle$, i.e. u is globally admissible with shape $C \cup \partial C$, $u|_T$ consists entirely of B-letters, and $u|_S = \eta$. Then the locations of the B-letters within u can be partitioned into maximal connected components $C_i(u)$, $1 \leq i \leq k(u)$ (say we order these lexicographically by least element), and we denote the subconfigurations of u occupying these components by $B_i(u) = u|_{C_i}$, $1 \leq i \leq k(u)$. We will now define a family of words $f(u) \subseteq L_{C \cup \partial C}(X) \cap \langle \eta \rangle$.

Begin by removing all $B_i(u)$ from u, defining a new configuration $v(u)$ with shape $(C \cup \partial C) \setminus \bigcup C_i(u)$ which consists only of G-letters. We fill the holes with shapes $C_i(u)$ in order, starting with $C_1(u)$. For each i, we order the sites in $C_i(u)$ lexicographically, and choose G-letters to fill them, one by one. We will do this in

such a way that at each step, regardless of what letters have been assigned, we have N choices of letters to use, and so the total number of configurations we define by filling all holes, $|f(u)|$, will be at least $N^{|T|}$.

Suppose that we wish to fill a site $s \in C_i(u)$, meaning that each $C_j(u)$ for $j < i$ has been filled, and all sites lexicographically less than s in $C_i(u)$ have been filled with G-letters. Then, consider all G-letters which can legally fill the site s given the letters already assigned within $C \cup \partial C$. Since all letters assigned are G-letters, by Lemma 3.3, there are at least $|A|(1 - (2d + 1)\epsilon)$ choices. If $s \in \partial C_i(u)$, then will use only letters from G_B, and if $s \in C_i(u) \setminus \partial C_i(u)$, then we will use only letters from G_I. In either case, though, since $|G_I|$ and $|G_B|$ are greater than $|A|\epsilon(N + 2d + 1)$, there are at least $|A|N\epsilon > N|B|$ possible choices. If $u(s)$ was the kth letter of B with respect to the previously defined ordering on B, then we use any of the N letters between the $((k-1)N+1)$th and kNth letters (inclusive) in either G_B or G_I with respect to the previously defined orderings on these sets. Denote by $f(u)$ the set of all configurations in $L_{C \cup \partial C}(X) \cap \langle \eta \rangle$ obtainable by using this filling algorithm to fill all of the sites of T in order. Now for each site $s \in \bigcup C_i(u)$, each configuration in $f(u)$ has a letter at s which encodes the following information: whether s was a boundary site or an interior site within its $C_i(u)$ (encoded by whether we chose a letter from G_B or G_I), and the B-letter $u(s)$ which appeared at s in u (encoded by which of the possible letters in G_B or G_I we used).

We now show that for any configurations $u \neq u'$ in $L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle$, $f(u)$ and $f(u')$ are disjoint. First, we deal with the case where $k(u) = k(u')$ and $C_i(u) = C_i(u')$ for $1 \leq i \leq k(u) = k(u')$. (Since they are equal, we just write C_i for $C_i(u) = C_i(u')$ and k for $k(u) = k(u')$.) Since $u \neq u'$, u and u' either disagree somewhere outside the union of the C_i or somewhere inside. If there is a disagreement somewhere outside, then since all configurations in $f(u)$ and $f(u')$ agree with u and u' respectively outside the union of the C_i, it is obvious that $f(u)$ and $f(u')$ are disjoint. If there is a disagreement inside the union of the C_i, then take j minimal so that there is a disagreement in C_j, and take $s \in C_j$ the minimal site lexicographically for which $u(s) \neq u'(s)$. For a contradiction, assume that there is a configuration w in $f(u) \cap f(u')$. Since all C_i are identical and since u and u' agree outside the union of the C_i, we know that exactly the same sites had been filled, with exactly the same letters, when $w(s)$ was chosen in the filling procedure defining $f(u)$ as when $w(s)$ was chosen in the filling procedure defining $f(u')$. But this is a contradiction; since $u(s) \neq u'(s)$ and the same set of letters was available to fill s in both procedures, the same letter could not possibly have been a legal choice in both procedures.

Now we deal with the case where either $k(u) \neq k'(u)$ or $k(u) = k'(u)$ and $C_i(u) \neq C_i(u')$ for some $1 \leq i \leq k(u) = k(u')$. This implies that either there exists $C_j(u)$ disjoint from all $C_i(u')$ (or the same statement with u and u' reversed), or there exist nonequal $C_j(u)$ and $C_{j'}(u')$ which have nonempty intersection (or the same statement with u and u' reversed). The first case is impossible since by definition, each $C_j(u)$ contains some site in T, and each site in T is contained in some $C_i(u')$ (and the same statement is true when u and u' are reversed). Suppose then that there exist j, j' so that $C_j(u) \neq C_{j'}(u')$ and $C_j(u) \cap C_{j'}(u') \neq \varnothing$. Then there exists s which is in the boundary of $C_j(u)$ and the interior of $C_{j'}(u')$, or vice versa. This means that when s is assigned in the filling procedures defining $f(u)$

and $f(u')$, either $w(s)$ must be from G_B in the former case and G_I in the latter, or vice versa. Either way, it ensures that $f(u) \cap f(u') = \varnothing$.

We have shown that all of the sets $f(u)$, $u \in L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle$, are disjoint. Since each is a subset of $L_{C \cup \partial C}(X) \cap \langle \eta \rangle$ and each has size at least $N^{|T|}$, we have shown that

$$|L_{C \cup \partial C}(X) \cap \langle \eta \rangle| \geq N^{|T|} |L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle|.$$

Recall that since μ is a measure of maximal entropy for X, by Proposition 2.16 it is an MRF with conditional probabilities Λ^δ. Therefore, if we define $\delta := \eta|_{\partial C}$, then

$$\mu(x|_T \in B^T \mid \langle \eta \rangle) = \frac{\mu(x|_T \in B^T, x|_S = \eta \mid \langle \delta \rangle)}{\mu(x|_S = \eta \mid \langle \delta \rangle)} = \frac{\Lambda^\delta(\langle B^T \rangle \cap \langle \eta \rangle)}{\Lambda^\delta(\langle \eta \rangle)} =$$

$$\frac{|L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle| / |L_{C \cup \partial C}(X) \cap \langle \delta \rangle|}{|L_{C \cup \partial C}(X) \cap \langle \eta \rangle| / |L_{C \cup \partial C}(X) \cap \langle \delta \rangle|} = \frac{|L_{C \cup \partial C}(X) \cap \langle B^T \rangle \cap \langle \eta \rangle|}{|L_{C \cup \partial C}(X) \cap \langle \eta \rangle|} < N^{-|T|}.$$

$\square$

We are now prepared to prove Theorem 3.1. We fix d, define $\epsilon_d = \frac{1}{160e} d^{-2} (2d)^{-8d}$, and consider any ϵ_d-full nearest neighbor $\mathbb{Z}^d$ SFT X. We usually suppress the dependence on d and just write $\epsilon = \epsilon_d$ in the sequel, and note for future reference that $\epsilon < \frac{1}{8d+8}$ and $\epsilon < \frac{1}{384d^2}$.

Proof of (A). Recall that for any finite $S \subseteq \mathbb{Z}^d$ and $\delta \in L_{\partial S}(X)$, Λ^δ is the conditional distribution on S given δ associated to any measure of maximal entropy, which is uniformly distributed over all configurations $w \in A^S$ which form a locally admissible configuration when combined with δ. We will show that there is only a single measure μ with these conditional distributions, implying that there is only a single measure of maximal entropy. Our method is similar to that of [2] in that we construct a coupling of Λ^δ and $\Lambda^{\delta'}$ for pairs of boundaries $\delta \neq \delta'$ of large connected shapes, and show that this coupling gives a high probability of agreement far from δ and δ', implying that Λ^δ and $\Lambda^{\delta'}$ behave similarly far from δ and δ'. (Informally, the influence of a boundary decays with distance.) However, we must begin with the special case where δ and δ' consist entirely of G-letters.

Choose any finite connected sets $C, C' \subset \mathbb{Z}^d$ with nonempty intersection, any site $s \in C \cap C'$, and any $\delta \in L_{\partial C}(X)$ and $\delta' \in L_{\partial C'}(X)$ consisting only of G-letters. Define $D := \min(d(s, \partial C), d(s, \partial C'))$. We will construct a coupling λ of Λ^δ and $\Lambda^{\delta'}$ which gives very small probability to a disagreement at s (when D is large).

Define $\overline{C} = C \cup \partial C$ and $\overline{C'} = C' \cup \partial C'$. Fix any ordering on the set $\overline{C} \cup \overline{C'}$; from now on when we talk about any notion of size for sites in $\overline{C} \cup \overline{C'}$, it is assumed we are speaking of this ordering. For convenience, we will extend C and C' to configurations on $\overline{C}$ and $\overline{C'}$ by appending δ and δ' respectively. Therefore, λ will be defined on pairs of configurations (w_1, w_2) where w_1 has shape $\overline{C}$ and w_2 has shape $\overline{C'}$; the marginalization of λ which leads to a true coupling of Λ^δ and $\Lambda^{\delta'}$ should be clear. We will define λ on one site at a time, assigning values to both $w_1(s)$ and $w_2(s)$ when s is in $\overline{C} \cap \overline{C'}$, and just assigning one of these two values if s is only one of the sets. We use α_1 and α_2 to denote the (incomplete) configurations on $\overline{C}$ and $\overline{C'}$ respectively at any step. We therefore begin with $\alpha_1 = \delta$ and $\alpha_2 = \delta'$. At any step of the construction, we use W to denote the set of vertices in $\overline{C} \cup \overline{C'}$ on which either α_1 or α_2 have already received values. (In particular, at the beginning, $W = \partial C \cup \partial C'$.) This means that α_1 is always defined on $W \cap \overline{C}$, and α_2 is always

defined on $W \cap \overline{C'}$. At an arbitrary step of the construction, we choose the next site s on which to assign values in α_1 and/or α_2 as follows:

(i) If there exists any site in $(\overline{C} \cup \overline{C'}) \setminus W$ which is adjacent to a site in W at which either α_1 or α_2 has been assigned a B-letter, then take s to be the smallest such site.

(ii) If (i) does not apply, but there exists a site in $(\overline{C} \cup \overline{C'}) \setminus W$ which is adjacent to a site in $\overline{C} \cap \overline{C'} \cap W$ (i.e. a site at which both α_1 and α_2 have been defined), and their values disagree, then take s to be the smallest such site.

(iii) If (i) and (ii) do not apply, but there exists a site in $(\overline{C} \cup \overline{C'}) \setminus W$ which is not in $\overline{C} \cap \overline{C'}$, then take s to be the smallest such site.

(iv) If none of (i)-(iii) apply, then take s to be the smallest site in $(\overline{C} \cup \overline{C'}) \setminus W$.

Now we are ready to define λ on s. If s is in $\overline{C}$ but not $\overline{C'}$ (i.e. chosen according to case (iii)), then assign $\alpha_1(s)$ randomly according to the marginalization of the distribution Λ^{α_1} to s, and if s is in $\overline{C'}$ but not $\overline{C}$, then assign $\alpha_2(s)$ randomly according to the marginalization of the distribution Λ^{α_2} to s.

If $s \in \overline{C} \cap \overline{C'}$ (i.e. chosen according to case (i) or case (ii)), then assign $\alpha_1(s)$ and $\alpha_2(s)$ according to an optimal coupling of the marginalizations of the distributions Λ^{α_1} and Λ^{α_2} to s. Since λ is defined sitewise, and at each step is assigned according to Λ^{α_1} in the first coordinate and Λ^{α_2} in the second, it is indeed a coupling of Λ^{δ} and $\Lambda^{\delta'}$. The key property of λ is the following:

Fact 3.6. *For any site $s \in \overline{C} \cap \overline{C'}$, λ-a.s., $w_1(s) \neq w_2(s)$ if and only if there exists a path γ from s to $\partial C \cup \partial C'$ contained within $\overline{C} \cap \overline{C'}$ such that for each site $t \in \gamma$, either one of $w_1(t)$ or $w_2(t)$ is a B-letter, or w_1 and w_2 disagree at t, i.e. $w_1(t) \neq w_2(t)$.*

Proof. Assume for a contradiction that $w_1(s) \neq w_2(s)$ and that no such path γ exists. Then there is a closed contour Γ containing s and contained within $\overline{C} \cap \overline{C'}$ so that $w_1|_\Gamma = w_2|_\Gamma \in G^\Gamma$. Denote by F the set of sites inside Γ. Then regardless of the order of the sites on which λ is defined, the first site in F which is assigned is done so by case (iv); since it is the first site in F to be assigned, its neighbors are either unassigned or in Γ, and so cases (i)-(iii) cannot apply. Call this site t.

Consider the state of λ when t is assigned under case (iv). The sets of undefined sites for α_1 and α_2 must be the same (since case (iii) was not applied), and every site adjacent to a site in $(\overline{C} \cup \overline{C'}) \setminus W$ must be a location at which α_1 and α_2 agree. (since case (ii) was not applied) Then the distributions Λ^{α_1} and Λ^{α_2} are identical. This means that their optimal coupling has support contained in the diagonal, and $\alpha_1(t) = \alpha_2(t)$ λ-a.s. It is then easy to see that λ-a.s., we remain in case (iv) for the remainder of the construction. Therefore, λ-a.s., α_1 and α_2 agree on all of F. Since $s \in F$, this clearly contradicts $w_1(s) \neq w_2(s)$.

$\square$

We will now show that the λ-probability of such a path (consisting entirely of disagreements and sites where w_1 or w_2 contains a B-letter) is very low. Consider any γ a path from s to $\partial C \cup \partial C'$ contained within $\overline{C} \cap \overline{C'}$, with length $L > D$. By passing to a subpath if necessary, we can assume that γ is minimal, i.e. it does not contain a proper subset which is also a path with the same initial and terminal vertices, and that γ is contained entirely within $C \cap C'$. Denote the length of γ by

L; clearly $L \geq D$. We will give an upper bound on the λ-probability that γ consists entirely of disagreements and locations where either w_1 or w_2 has a B-letter. Define $\gamma' = \gamma \cup \partial\gamma$, the "thickened" version of the path γ. Since $\gamma \subset C \cap C'$, $\gamma' \subset \overline{C} \cap \overline{C'}$, and so both w_1 and w_2 are defined on γ' for (w_1, w_2) in the support of λ. It should be obvious that $|\gamma'| \leq 2dL$. We need a definition:

Definition 3.7. A site $s \in \gamma$ is B-**proximate** if there is $t \in N_s \cup \{s\} \subset \gamma'$ for which $w_1(t)$ or $w_2(t)$ is a B-letter.

We now separate into two cases depending on whether the number of sites in γ which are B-proximate is greater than or equal to $\frac{L}{2}$ or not.

Case 1: γ contains at least $\frac{L}{2}$ B-proximate sites

In this case, γ' contains at least $\frac{L}{4d}$ sites at which either w_1 or w_2 has a B-letter, since a single such site can "induce" at most $2d$ B-proximate sites in γ. (It is not possible for such a site and all $2d$ of its neighbors to be in γ, or else γ would not be minimal.) Therefore, either $w_1|_{\gamma'}$ or $w_2|_{\gamma'}$ (w.l.o.g. we say $w_1|_{\gamma'}$) contains $\frac{L}{8d}$ B-letters.

Fix any set $T \subseteq \gamma'$ with size $\frac{L}{8d}$. By Lemma 3.5, $\Lambda^\delta(x|_T \in B^T) < N^{-\frac{L}{8d}}$, where $N = \lfloor \frac{1}{2}(\epsilon^{-1} - 4d - 4) \rfloor$. Since $\epsilon < \frac{1}{8d+8}$, $N > \frac{1}{5\epsilon}$. The number of possible such T for given γ is bounded from above by $\binom{2dL}{\frac{L}{8d}} \leq (16ed^2)^{\frac{L}{8d}}$. Therefore, the Λ^δ-probability that there exists any such T for our fixed γ is bounded from above by $(80ed^2\epsilon)^{\frac{L}{8d}}$. Since there are fewer than $(2d)^L$ possible γ of length L, the Λ^δ-probability that any such γ has at least $\frac{L}{8d}$ B-letters is less than

$$\sum_{L=D}^{\infty} \left(80ed^2(2d)^{8d}\epsilon\right)^{\frac{L}{8d}} \leq \sum_{L=D}^{\infty} 2^{-\frac{L}{8d}} = \frac{1}{1 - \sqrt[8d]{0.5}} 2^{-\frac{D}{8d}}$$

since $\epsilon \leq \frac{1}{160e}d^{-2}(2d)^{-8d}$. The same is true of $\Lambda^{\delta'}$, and since λ is a coupling of Λ^δ and $\Lambda^{\delta'}$, the λ-probability that there exists any path γ for which at least half of its sites are B-proximate is less than $\frac{2}{1 - \sqrt[8d]{0.5}} 2^{-\frac{D}{8d}}$.

Case 2: γ contains fewer than $\frac{L}{2}$ B-proximate sites

In this case, there exists $R \subset \gamma$, $|R| > \frac{L}{2}$, such that no site in R is B-proximate. Since γ consists entirely of sites where either one of w_1 and w_2 is a B-letter or w_1 and w_2 disagree, this implies that for each $r \in R$, $w_1(r) \neq w_2(r)$. Also, by the definition of B-proximate, for each $r \in R$, both $w_1|_{N_r}$ and $w_2|_{N_r}$ contain only G-letters. Order the elements of R as $r_1, \ldots, r_K$, $K > \frac{L}{2}$. Our fundamental claim is that for any $i \in [1, K]$,

$$(1) \quad \lambda\big(w_1(r_i) \neq w_2(r_i) \mid w_1(r_j) \neq w_2(r_j), 1 \leq j < i$$
$$\text{and } w_1|_{N_{r_{j'}}}, w_2|_{N_{r_{j'}}} \in G^{N_{r_{j'}}}, 1 \leq j' \leq i\big) < 12d\epsilon.$$

To prove (1), we fix some $i \in [1, K]$ and condition on the facts that $w_1(r_j) \neq w_2(r_j)$ for $1 \leq j < i$, and that $w_1|_{N_{r_{j'}}}, w_2|_{N_{r_{j'}}} \in G^{N_{r_{j'}}}$ for $1 \leq j' \leq i$. Then the conditional λ-distribution on r_i is a weighted average of the λ-distribution assigned at site r_i, taken over all possible evolutions of w_1 and w_2 in the definition of λ. For any such evolution of w_1 and w_2, at the step where $w_1(r_i)$ and $w_2(r_i)$ were (simultaneously) assigned, no unassigned site in either C or C' was adjacent to an assigned B-letter.

(Otherwise, the smallest such site lexicographically would be used instead of r_i, under case (i) in the definition of λ.)

Again, independently of which evolution of w_1 and w_2 we consider, Lemma 3.5 implies that for any possible α_1 when r_i was assigned, $\Lambda^{\alpha_1}(x|_{N_{r_i}} \in G^{N_{r_i}}) > 1 - \frac{2d}{N} > 1 - 10d\epsilon$. This means that for any possible α_1 when r_i was assigned, $\Lambda^{\alpha_1}(x(r_i))$ was a weighted average of the conditional distributions $\Lambda^{\alpha_1}(x(r_i) \mid x|_{N_{r_i}})$, where all but at most $10d\epsilon$ of the weights are associated to $x|_{N_{r_i}}$ consisting entirely of G-letters. For any such $x|_{N_{r_i}}$, $\Lambda^{\alpha_1}(x(r_i) \mid x|_{N_{r_i}})$ is a uniform distribution over a subset of A of size at least $|A|(1 - 2d\epsilon)$ by ϵ-fullness of X. Therefore, for any such $x|_{N_{r_i}}$,

$$d\left(\Lambda^{\alpha_1}(x(r_i) \mid x|_{N_{r_i}}), U\right) < 2d\epsilon,$$

where we use U to denote the uniform distribution over all of A. The analogous estimate also holds for Λ^{α_2} by exactly the same argument. Since at least $1 - 10d\epsilon$ of the measures $\Lambda^{\alpha_1}(x(r_i))$ and $\Lambda^{\alpha_2}(x(r_i))$ have been decomposed as weighted averages of distributions within $2d\epsilon$ of U,

$$d\left(\Lambda^{\alpha_1}(x(r_i)), \Lambda^{\alpha_2}(x(r_i))\right) < 12d\epsilon.$$

Since the marginalization of λ to r_i is an optimal coupling of these two measures, this marginalization gives a probability of less than $12d\epsilon$ to the event $w_1(r_i) \neq w_2(r_i)$. Since the same is true for every evolution of w_1 and w_2, we have shown that conditioned on $w_1(r_j) \neq w_2(r_j)$ for $1 \leq j < i$ and $w_1|_{N_{r_{j'}}}, w_2|_{N_{r_{j'}}} \in G^{N_{r_{j'}}}$ for $1 \leq j' \leq i$, $\lambda(w_1(r_i) \neq w_2(r_i)) < 12d\epsilon$, verifying (1).

From this, it is clear that $\lambda(\text{no site in } R \text{ is } B-\text{proximate}) < (12d\epsilon)^{\frac{L}{2}}$ by decomposing it as a product of conditional probabilities. There are at most $(2d)^L$ choices for γ and at most 2^L choices for the subset R, so the λ-probability that there is any path γ with at least $\frac{1}{2}$ non-B-proximate sites is less than

$$\sum_{L=D}^{\infty} (192d^2\epsilon)^{-\frac{L}{4}} < 2^{-\frac{D}{4}}$$

since $\epsilon < \frac{1}{384d^2}$.

Clearly any path γ from s to $\partial C \cup \partial C'$ contained within $\overline{C} \cap \overline{C'}$ which consists entirely of disagreements and locations where either w_1 or w_2 has a B-letter must be in either Case 1 or Case 2, so we have shown that the λ-probability that there exists any such path is less than $\frac{2}{1 - \sqrt[8d]{0.5}} 2^{-\frac{D}{8d}} + 2^{-\frac{D}{4}} < Z 2^{-\frac{D}{8d}}$ for a constant Z independent of D. By Fact 3.6, λ-a.s., $w_1(s) \neq w_2(s)$ if and only if there exists such a γ, and so $\lambda(w_1(s) \neq w_2(s)) < Z 2^{-\frac{D}{8d}}$. Clearly this implies that for any shape S consisting of sites at a distance at least D from ∂C and $\partial C'$, $\lambda(w_1|_S \neq w_2|_S) < Z|S|2^{-\frac{D}{8d}}$.

Since λ is a coupling of Λ^δ and $\Lambda^{\delta'}$, we have shown the following:

Fact 3.8. *For any $\delta \in L_{\partial C}(X)$ and $\delta' \in L_{\partial C'}(X)$ consisting only of G-letters, and for any shape $S \in C \cap C'$ such that $D = \min(d(S, \partial C), d(S, \partial C')) > 20$,*

$$(2) \qquad\qquad d\left(\Lambda^\delta(x|_S), \Lambda^{\delta'}(x|_S)\right) < Z|S|2^{-\frac{D}{8d}}.$$

We note that (2) is very close to the classical condition of (weak) spatial mixing with exponential rate (see [5] for a survey of various results and discussions involving spatial mixing), but with the important difference that it only applies here to

boundaries consisting entirely of G-letters. To finish the proof, we must now consider general boundaries η. For this portion of the proof, we will use only the fact that $d\left(\Lambda^\delta(x|_S), \Lambda^{\delta'}(x|_S)\right)$ decays to 0 as $D \to \infty$, ignoring the exponential rate.

Roughly speaking, the strategy is to show that for any connected set $C \subseteq \mathbb{Z}^d$, any measure of maximal entropy μ on X and for any finite shape $S \subset C$ far from ∂C, "most" possible boundary conditions $\eta \in L_{\partial C}(X)$ (in the sense of μ-measure) have the following property: with very high Λ^η-probability there exists a closed contour δ of G-letters inside C such that S is contained in the interior of δ, and $d(S, \delta)$ is large. Then, for any such η, most of $\Lambda^\eta(x|_S)$ can be written as a weighted average over $\Lambda^\delta(x|_S)$ for such δ. We have already shown that $\Lambda^\delta(x|_S)$ has very little dependence on δ consisting only of G-letters when $d(S, \delta)$ is large, and so this will show that for "most" η, $\Lambda^\eta(x|_S)$ has little dependence on η. Therefore, by taking C larger and larger, we can write $\mu|_S$ as a limit of weighted averages of $\Lambda^\eta(x|_S)$, and since the above shows that dependence on η fades as C becomes large, $\mu|_S$ has only one possible value, showing uniqueness of μ.

In the sequel, we use the notation $S \leftrightarrow T$ to denote the event that there is a path of B-letters connecting some site in S to some site in T. We first need to prove the following:

Fact 3.9. *For any measure of maximal entropy μ on X,*

$$(3) \qquad \lim_{n \to \infty} \mu(0 \leftrightarrow \partial[-n, n]^d) = 0.$$

Proof. For a contradiction, suppose (3) is false. Then since the events $0 \leftrightarrow \partial[-n, n]^d$ are decreasing, there exists $\alpha > 0$ so that for all n, $\mu(0 \leftrightarrow \partial[-2n, 2n]^d) > \alpha$. Then by stationarity, for each $s \in [-n, n]^d$, $\mu(s \leftrightarrow (s + \partial[-2n, 2n]^d)) > \alpha$. Since $s + [-2n, 2n]^d \supset [-n, n]^d$, $s \leftrightarrow (s + \partial[-2n, 2n]^d)$ implies $s \leftrightarrow \partial[-n, n]^d$, and so $\mu(s \leftrightarrow \partial[-n, n]^d) > \alpha$ for all $s \in [-n, n]^d$. Then

$$(4) \qquad \mu\left(|\{s \in [-n, n]^d \ : \ s \leftrightarrow \partial[-n, n]^d\}| > 0.5\alpha|[-n, n]^d|\right) > 0.5\alpha.$$

Since μ is an MRF with conditional probabilities $\{\Lambda^\delta\}$, we may write $\mu|_{[-n,n]^d}$ as a weighted average:

$$\mu|_{[-n,n]^d} = \sum \alpha_i \Lambda^{\delta_i},$$

where δ_i ranges over configurations in $L_{\partial[-n,n]^d}(X)$ (and $\alpha_i = \mu([\delta_i])$.) By (4), at least 0.25α of the weights α_i are associated to δ_i for which

$$(5) \qquad \Lambda^{\delta_i}(|\{s \in [-n, n]^d \ : \ s \leftrightarrow \partial[-n, n]^d\}| > 0.5\alpha|[-n, n]^d|) > 0.25\alpha.$$

In other words, if we denote the set of δ_i which satisfy (5) by P, then $\mu([P]) > 0.25\alpha$. Make the notation $K = |\{w \in L_{[-n-1,n+1]^d}(X) \ : \ w|_{\partial[-n,n]^d} \in P\}|$. Recall that $\sum_{w \in L_{[-n-1,n+1]^d}(X)} -\mu([w]) \log \mu([w]) \geq |[-n-1, n+1]^d| h(\mu)$, which equals

$|[-n-1, n+1]^d|h(X)$ since μ is a measure of maximal entropy. Then

$$
(6) \quad |[-n-1, n+1]^d|h(X) \leq \sum_{w \in L_{[-n-1,n+1]^d}(X)} -\mu([w]) \log \mu([w]) =
$$

$$
\sum_{\delta \in P^c} \sum_{w \in L_{[-n-1,n+1]^d}(X), w|_{\partial[-n,n]^d}=\delta} -\mu([w]) \log \mu([w])
$$

$$
+ \sum_{\delta \in P} \sum_{w \in L_{[-n-1,n+1]^d}(X), w|_{\partial[-n,n]^d}=\delta} -\mu([w]) \log \mu([w])
$$

$$
\leq (1-0.25\alpha) \log |L_{[-n-1,n+1]^d}(X)| + 0.25\alpha \log K - \log(0.25\alpha),
$$

where the last inequality uses the easily checked fact that for any finite set of reals $S = \{s_1, \ldots, s_k\}$ with sum s, $\sum(-g_i \log s_i) \leq s(\log k - \log s)$.

By definition of topological entropy, for any $\eta > 0$, there exists N_η such that $(1 - 0.25\eta\alpha) \log |L_{[-n,n]^d}(X)| < |[-n,n]^d|h(X) + \log(0.25\alpha)$ for $n > N_\eta$. This means, in particular, that for $n \geq N_\eta$,

$$
(1-0.25\eta\alpha)\log|L_{[-n-1,n+1]^d}(X)| < |[-n-1,n+1]^d|h(X) + \log(0.25\alpha)
$$

$$
\leq (1-0.25\alpha)\log|L_{[-n-1,n+1]^d}(X)| + 0.25\alpha \log K \quad \text{(by (6))}
$$

$$
\implies \log K > (1-\eta)\log|L_{[-n-1,n+1]^d}(X)|.
$$

Therefore, for $n > N_\eta$, $K > |L_{[-n-1,n+1]^d}(X)|^{1-\eta}$. Since there are fewer than $|A|^{|\partial[-n,n]^d|}$ elements of P, there exists $\delta \in P$ and a set of at least $\frac{|L_{[-n-1,n+1]^d}(X)|^{1-\eta}}{|A|^{|\partial[-n,n]^d|}}$ configurations w for which $w\delta \in L(X)$. Then, since δ satisfies (5), there is a set of configurations $S \subseteq L_{[-n,n]^d}(X)$ of size at least $\frac{0.25\alpha|L_{[-n-1,n+1]^d}(X)|^{1-\eta}}{|A|^{|\partial[-n,n]^d|}}$, each of which contains at least $0.5\alpha|[-n,n]^d|$ sites connected to $\partial[-n,n]^d$ by paths of B-letters.

We now perform a very similar replacement procedure to the one used in the proof of Lemma 3.5. We will not give an exact description, and just summarize the changes from the previous procedure. Consider any $u \in S$, and take $C_i(u)$, $1 \leq i \leq k$, to be the maximal connected components of locations of B-letters in u which have nonempty intersection with $\partial[-n,n]^d$. Since $u \in S$, $\sum |C_i(u)| > 0.5\alpha|[-n,n]^d|$. For each i, define $B_i(u)$ to be the subconfiguration $u|_{C_i(u)}$ of u occupying $C_i(u)$. Then, remove all $B_i(u)$ from u, and fill the holes in various ways using the same procedure as in Lemma 3.5, where each site s is filled with a G-letter which encodes the information about the B-letter $u(s)$ and whether s was on the boundary of its component $C_i(u)$ or in the interior. As in Lemma 3.5, this yields a set $f(u) \subseteq L_{[-n,n]^d}(X)$ of configurations of size at least $N^{0.5\alpha|[-n,n]^d|}$, where $N > \frac{1}{5\epsilon}$. Here, we will actually only need the fact that $N > 2$. Therefore,

$$
\sum_u |f(u)| > |S| 2^{0.5\alpha|[-n,n]^d|}.
$$

In Lemma 3.5, we showed that all of the sets $f(u)$ were disjoint, which is not necessarily the case here. However, it is still true that if there exist $C_i(u)$ and $C_j(u')$ which are unequal but have nonempty intersection, then $f(u) \cap f(u') = \varnothing$. It is also still true that if there exist $C_i(u)$ and $C_j(u')$ which are equal, but $B_i(u) \neq B_j(u')$, then $f(u) \cap f(u') = \varnothing$. The only new case under which $f(u)$ and $f(u')$ might

not be disjoint is if all pairs $C_i(u)$ and $C_j(u')$ are either disjoint or equal, and if $B_i(u) = B_i(u')$ whenever $C_i(u) = C_j(u')$; suppose we are in this case. Fix any $v \in L_{[-n,n]^d}(X)$, and let us bound the size of $F_v := \{u \ : \ v \in f(u)\}|$ from above.

For each $s \in \partial[-n,n]^d$, either s is in some $C_i(u_s)$ for $u_s \in F_v$ or not. If it is, then by the above analysis, then denote the configuration $B_i(u_s)$ by $B(s)$. By the above analysis, for every $u \in F_v$, either $B(s) = B_i(u)$ for some i or all $B_i(u)$ are disjoint from $B(s)$ (in particular, this would imply that no $B_i(u)$ contains s). This in turn implies that for every $u \in F_v$, the set $\{B_i(u)\}$ is just a subset of $\{B(s)\}_{s \in \partial[-n,n]^d}$. Since knowing $\{B_i(u)\}$ along with v uniquely determines u, and since there are at most $|\partial[-n,n]^d|$ sets $B(s)$, $|F_v| \le 2^{|\partial[-n,n]^d|}$. In other words, each v is in at most $2^{|\partial[-n,n]^d|}$ of the sets $f(u)$. Since this is true for any v, we have shown that

$$|L_{[-n-1,n+1]^d}(X)| \ge \frac{\sum_u |f(u)|}{2^{|\partial[-n,n]^d|}} \ge \frac{|S|2^{0.5\alpha|[-n-1,n+1]^d|}}{2^{|\partial[-n,n]^d|}}$$

$$\ge |L_{[-n-1,n+1]^d}(X)|^{1-\eta} \frac{0.25\alpha 2^{0.5\alpha|[-n-1,n+1]^d|}}{(2|A|)^{|\partial[-n,n]^d|}}$$

$$\implies |L_{[-n-1,n+1]^d}(X)| \ge \left(\frac{0.25\alpha 2^{0.5\alpha|[-n-1,n+1]^d|}}{(2|A|)^{|\partial[-n,n]^d|}} \right)^{\eta^{-1}}.$$

However, since $|L_{[-n-1,n+1]^d}(X)| < |A|^{|[-n-1,n+1]^d|}$, this clearly gives a contradiction for small enough η and sufficiently large n (both larger than N_η and large enough so that $\frac{|\partial[-n,n]^d|}{|[-n,n]^d|}$ is much smaller than 0.5α). Therefore, our original assumption was wrong and (3) is true, i.e. $\lim_{n \to \infty} \mu(0 \leftrightarrow \partial[-n,n]^d) = 0$.

□

We are now ready to complete the proof of (A). Choose μ to be any measure of maximal entropy on X, and fix any k, l, and $\epsilon > 0$. By Fact 3.9, we can choose $n > k + l$ large enough that $\mu(0 \leftrightarrow \partial[-n-k-l, n+k+l]^d) < \frac{\epsilon}{|[-k-l,k+l]^d|}$. Then by stationarity of μ, $\mu(v \leftrightarrow v + \partial[-n-k-l, n+k+l]^d) < \frac{\epsilon}{|[-k-l,k+l]^d|}$ for all $v \in [-k-l, k+l]^d$. Since $[-n,n]^d \subseteq v + [-n-k-l, n+k+l]^d$ for all such v, the event $v \leftrightarrow \partial[-n,n]^d$ is contained in the event $v \leftrightarrow v + \partial[-n-k-l, n+k+l]^d$ for all such v, and so $\mu(v \leftrightarrow \partial[-n,n]^d) < \frac{\epsilon}{|[-k-l,k+l]^d|}$. Summing over all $v \in [-k-l, k+l]^d$ yields $\mu([-k-l, k+l]^d \leftrightarrow \partial[-n,n]^d) < \epsilon$.

This implies that for any n, there is a set $U_n \subseteq L_{[-n,n]^d}(X)$ with $\mu([U_n]) > 1 - \epsilon$ for which any $w \in U_n$ contains a closed contour consisting entirely of G-letters containing $[-k-l, k+l]^d$ in its interior. For any $w \in U_n$, if $\{\gamma_i\} = \{\partial S_i\}$ is the collection of all such contours, then clearly $\gamma(w) := \partial(\bigcup S_i)$ is the unique maximal such contour, i.e. any other such closed contour γ' for w is contained in the interior of $\gamma(w)$. Define $B(w)$ to consist of the set of all sites of $[-n,n]^d$ on or outside $\gamma(w)$, and $D(w) = w|_{B(w)}$. We note that $[U_n]$ can be written as a disjoint union of the sets $[L_{[-n,n]^d}(X) \cap \langle D(w)\rangle]$ over all possible choices for $D(w)$. (For clarity, we note that $L_{[-n,n]^d}(X) \cap \langle D(w)\rangle$ consists of all words x in $L_{[-n,n]^d}(X)$ for which $x|_{B(w)} = D(w)$.) This means that μ, restricted to U_n and then marginalized to $[-k,k]^d$, can be written as a weighted average of the conditional measures $\mu\left(x|_{[-k,k]^d} \mid [D(w)]\right)$ over possible values of $D(w)$, and since μ is an MRF, this is actually a weighted average of $\Lambda^{\gamma(w)}|_{[-k,k]^d}$.

However, each $\gamma(w)$ is a closed contour of G-letters with distance greater than l from $[-k,k]^d$, and so by Fact 3.8, for any $\gamma(w)$ and any $\eta \in L_{\partial[-n,n]^d}(X)$ consisting only of G-letters,

$$(7) \qquad d\left(\Lambda^{\gamma(w)}|_{[-k,k]^d}, \Lambda^{\eta}|_{[-k,k]^d}\right) < Z|[-k,k]^d|2^{-\frac{l}{8d}}.$$

Since the set U_n has μ-measure at least $1 - \epsilon$, and since $\mu|_{[-k,k]^d}$ restricted to U_n can be decomposed as a weighted average of measures $\Lambda^{\gamma(w)}|_{[-k,k]^d}$, (7) implies that

$$(8) \qquad d\left(\mu|_{[-k,k]^d}, \Lambda^{\eta}|_{[-k,k]^d}\right) < Z|[-k,k]^d|2^{-\frac{l}{8d}} + \epsilon.$$

By taking $l \to \infty$ and $\epsilon \to 0$ (thus forcing $n \to \infty$, since n was chosen larger than $k + l$), we see that $\mu|_{[-k,k]^d}$ is in fact uniquely determined by the conditional probabilities Λ^{δ}. Since k was arbitrary, μ is the unique MRF with conditional probabilities Λ^{δ}, implying by Proposition 2.16 that μ is the unique measure of maximal entropy on X, proving (A).

$\square$

We now state two easy corollaries of the proof of (A), which will be useful later for the proofs of (B) - (E).

Corollary 3.10. *If X is an ϵ_d-full nearest neighbor $\mathbb{Z}^d$ SFT with unique m.m.e. μ, then μ is the (unique) weak limit (as $n \to \infty$) of Λ^{η_n} for any sequence $\eta_n \in L_{\partial[-n,n]^d}(X)$ of boundary configurations consisting only of G-letters.*

Proof. Choose any such sequence η_n. For any k, (8) implies that as $n \to \infty$, $\Lambda^{\eta_n}|_{[-k,k]^d}$ approaches $\mu|_{[-k,k]^d}$ weakly. Therefore, $\Lambda^{\eta_n} \to \mu$. $\square$

Corollary 3.11. *If X is an ϵ_d-full nearest neighbor $\mathbb{Z}^d$ SFT with unique m.m.e. μ, then any configuration $u \in L(X)$ containing only G-letters on its inner boundary has positive μ-measure.*

Proof. Consider any such configuration $u \in L_S(X)$. It was shown in the proof of (A) that there exists $T \supset S$ and a closed contour $\delta \in L_{\partial T}$ of G-letters containing S for which $\mu([\delta]) > 0$. (Specifically, take k, l, n large enough that $S \subseteq [-k-l, k+l]^d$ and $\mu([U_n]) > 0$, choose $w \in U_n$ with $\mu([w]) > 0$, and then take $\delta = \gamma(w)$.) Since the concatenation $u\delta$ has inner boundary consisting only of G-letters, by Corollary 3.4 it is globally admissible. Therefore, there exists a configuration $v \in L(X)$ with $v|_S = u$ and $v|_{\partial T} = \delta$, implying that $\Lambda^{\delta}([u]) > 0$. Then $\mu([u]) \geq \mu([u\delta]) = \mu([\delta])\mu([u] \mid [\delta]) = \mu([\delta])\Lambda^{\delta}([\mu]) > 0$. $\square$

Remark 3.12. At first glance, our argument may appear to be an extension of the main reuslt from [2], which guaranteed uniqueness of MRFs corresponding to certain classes of conditional probabilities. However, this is not the case; even for ϵ arbitrarily close to 0, ϵ-full SFTs may still support multiple MRFs with the same conditional probabilities Λ^{δ}, some corresponding to limits of boundary conditions involving B-letters. (For instance, in Example 2.33, both the point mass at $0^{\mathbb{Z}^d}$ and the Bernoulli measure of maximal entropy on $\{1, \ldots, n\}^{\mathbb{Z}^d}$ are MRFs with conditional probabilities Λ^{δ}.) We very much need the extra condition of maximal entropy to rule out all but one of these MRFs as "degenerate."

Proof of (B). Our strategy is to first show that we can compute $h(X)$ by taking the exponential growth rate of globally admissible configurations whose boundaries contain only G-letters, and that we can bound the rate at which these approximations approach $h(X)$. Then, we can easily write an algorithm which counts such configurations, since by Corollary 3.4, a configuration with boundary containing only G-letters is globally admissible iff it is locally admissible.

Fix any n, and denote by Γ the set $\partial[1,n]^d$. For any $\delta \in L_\Gamma(X)$, we will show that $|L_{[1,n]^d}(X) \cap \langle\delta\rangle| \le |L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|$, i.e. the number of globally admissible configurations with shape $[1,n]^d$ whose restriction to Γ equals δ is less than or equal to the number of globally admissible configurations with shape $[1,n]^d$ whose restriction to Γ consists entirely of G-letters. The proof involves a similar replacement procedure as in the proof of Lemma 3.5; the difference is that for any $u \in L_{[1,n]^d}(X) \cap \langle\delta\rangle$, we will define only a single configuration $f(u)$ in $L_{[1,n]^d}(X)$, rather than a set. Take $C_i(u)$, $1 \le i \le k(u)$, to be the set of maximal connected components of locations of B-letters in u which have nonempty intersection with Γ, and for each i define $B_i(u) = u|_{C_i(u)}$ to be the subconfiguration of u occupying $C_i(u)$. Just as in the proof of Lemma 3.5, remove all $B_i(u)$ from u, and fill the holes with words of G-letters, where the letter chosen to fill a site s encodes the letter $u(s)$ and the information of whether s was on the boundary of its component $C_i(u)$ or in the interior. For exactly the same reasons as in Lemma 3.5, $u \ne u' \Rightarrow f(u) \ne f(u')$. We also note that all $f(u)$ are in $L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle$, meaning that their restrictions to Γ consist entirely of G-letters.

We have then shown that $|L_{[1,n]^d}(X) \cap \langle\delta\rangle| \le |L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|$. By summing over all choices for δ, we see that $|L_{[1,n]^d}(X)| \le |A|^{|\Gamma|}|L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|$. This means that

$$(9) \quad h(X) \le \frac{1}{n^d}\log|L_{[1,n]^d}(X)| \le \frac{1}{n^d}\left(\log|L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle| + |\Gamma|\log|A|\right)$$
$$\le \frac{1}{n^d}\left(\log|L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|\right) + \frac{2d\log|A|}{n}.$$

We now make a simple observation: for any k and any configurations $w_v \in L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle$, $v \in [1,k]^d$, define the concatenation u of all w_v, which has shape $\bigcup_v \prod[1+(v_i-1)(n+1), v_i(n+1)-1]$ and is defined by $u|_{\prod[1+(v_i-1)(n+1), v_i(n+1)-1]} = w_v$. Then u is made up of a union of locally admissible configurations where each pair is separated by a distance of at least 1, so it is locally admissible. Then by Corollary 3.4, since the outer boundary of u consists only of G-letters and $|A|(1-(2d+1)\epsilon) \ge 1$, u is also globally admissible, meaning in particular that it is a subconfiguration of a configuration in $L_{[1,k(n+1)]^d}(X)$. This implies that for any $k > 0$,

$$|L_{[1,k(n+1)]^d}(X)| \ge |L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|^{k^d}.$$

By taking logs of each side, dividing by $(k(n+1))^d$, and letting $k \to \infty$, we see that

$$(10) \qquad h(X) \ge \frac{1}{(n+1)^d}\left(\log|L_{[1,n]^d}(X) \cap \langle G^\Gamma\rangle|\right).$$

The upper and lower bounds on $h(X)$ given by (9) and (10) differ by

$$(11) \quad \frac{2d \log |A|}{n} + \left(\frac{1}{n^d} - \frac{1}{(n+1)^d}\right) \log \left(|L_{[1,n]^d}(X) \cap \langle G^\Gamma \rangle|\right) \leq \frac{2d \log |A|}{n}$$

$$+ \frac{(n+1)^d - n^d}{n^d(n+1)^d} n^d \log |A| \leq \frac{2d \log |A|}{n} + \frac{d(n+1)^{d-1}}{(n+1)^d} \log |A| \leq \frac{3d \log |A|}{n}.$$

Since $\frac{1}{n^d} \log \left(|L_{[1,n]^d}(X) \cap \langle G^\Gamma \rangle|\right)$ is between the bounds from (9) and (10), it is within $\frac{3d \log |A|}{n}$ of $h(X)$. And by Corollary 3.4, $\log \left(|L_{[1,n]^d}(X) \cap \langle G^\Gamma \rangle|\right) = \log \left(|LA_{[1,n]^d}(X) \cap \langle G^\Gamma \rangle|\right)$. Finally, we note that $\log \left(|LA_{[1,n]^d}(X) \cap \langle G^\Gamma \rangle|\right)$ can be computed algorithmically, in $|A|^{n^d(1+o(1))}$ steps, by simply writing down all possible configurations with alphabet A and shape $[1,n]^d$ and counting those which are locally admissible and have restriction to Γ consisting only of G-letters.

Since we may invest $|A|^{n^d(1+o(1))}$ steps to get an approximation to $h(X)$ with tolerance $\frac{3d \log |A|}{n}$, clearly $h(X)$ is computable in time $|A|^{(3d \log |A|n)^d(1+o(1))}$, verifying (B).

$\square$

Proof of (C). For any n, we define X_n to be the $\mathbb{Z}^d$ SFT consisting of all points of X in which all connected components of B-letters have size less than n. We will show that each X_n has the UFP and that $h(X_n) \to h(X)$ as $n \to \infty$.

We first verify that X_n has the UFP with distance $2n$. Consider any k, l with $l > k+3n$ and any configurations $w \in L_{[-k,k]^d}(X_n)$ and $w' \in L_{[-l,l]^d \setminus [-k-2n,k+2n]^d}(X_n)$. We will exhibit $x \in X_n$ with $x|_{[-k,k]^d} = w$ and $x|_{[-l,l]^d \setminus [-k-2n,k+2n]^d} = w'$, proving the UFP. (This is not exactly the definition of the UFP we gave earlier, which involved configurations on $[-k,k]^d$ and $\mathbb{Z}^d \setminus [-k-2n,k+2n]^d$, but clearly one can take limits as $l \to \infty$.)

We first use the fact that w, w' are globally admissible in X_n to extend them to configurations $v \in L_{[-k-n+1,k+n-1]^d}$ and $v' \in L_{[-l,l]^d \setminus [-k-n-1,k+n+1]^d}$ respectively. Then, in both v and v', remove any connected components of B-letters which have empty intersection with w or w'. Fill these with G-letters in some locally admissible way, in the same way as in several arguments thus far, creating new configurations u and u' respectively. Since connected components of B-letters in X_n must have size less than n, $u|_{\underline{\partial}[-k-n+1,k+n-1]^d}$ and $u'|_{\underline{\partial}[-k-n-1,k+n+1]^d}$ consist only of G-letters. (If this were not the case, then either u contained a connected component of B-letters intersecting both $[-k,k]^d$ and $\underline{\partial}[-k-n+1,k+n-1]^d$ or u' contained a connected component of B-letters intersecting both $[-l,l]^d \setminus [-k-2n,k+2n]^d$ and $\underline{\partial}[-k-n-1,k+n+1]^d$, and in either case such a component would have had size at least n, which is impossible since $v, v' \in L(X_n)$.) Then by Corollary 3.4, the empty region $\underline{\partial}[-k-n,k+n]^d$ between u and u' can be filled with G-letters in a locally admissible way, creating a new locally admissible configuration v'' with shape $[-l,l]^d$. Finally, we note that since w' was globally admissible, there exists $x' \in X$ with $x'|_{[-l,l]^d \setminus [-k-2n,k+2n]^d} = w'$. Finally, we note that w' has "thickness" at least n, and that no letters on w' were changed in the construction of v''. Therefore, since X_n is an SFT defined by forbidden configurations of size at most n, the point $x \in A^{\mathbb{Z}^d}$ defined by $x|_{[-l,l]^d} = v''$ and $x|_{\mathbb{Z}^d \setminus [-l,l]^d} = x'|_{\mathbb{Z}^d \setminus [-l,l]^d}$ is in X_n, and we are done. (See Figure 1 for an illustration of the creation of x.)
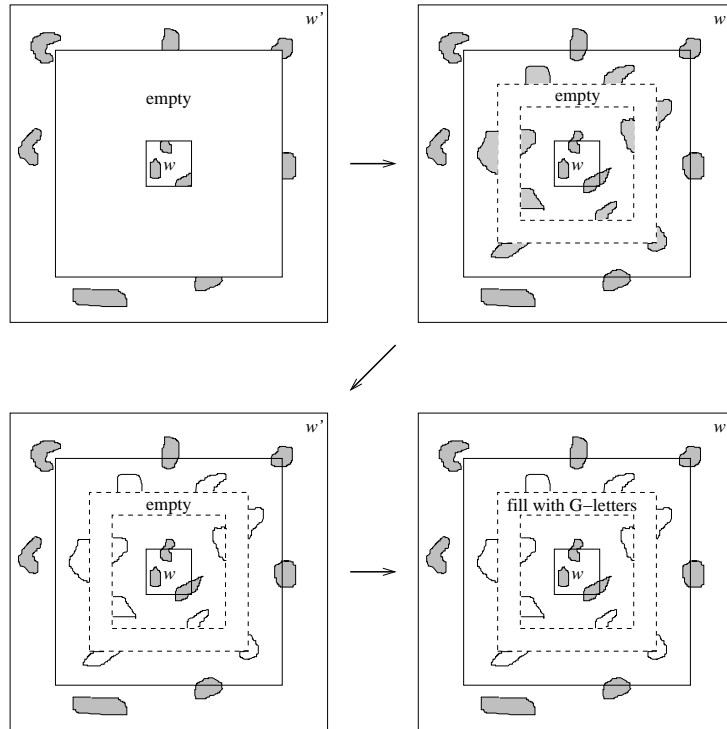
FIGURE 1. Filling between w and w' (shaded areas represent B-letters, white areas represent G-letters)

We finish by verifying that $h(X_n) \to h(X)$. We showed in the proof of (B) that for any collection $w_v \in L_{[1,n]^d}(X) \cap \langle G^{\partial[1,n]^d}\rangle$, $v \in [1,k]^d$, the concatenation u of all w_v, which has shape $\bigcup_v \prod[1 + (v_i - 1)(n + 1), v_i(n + 1) - 1]$ and is defined by $u|_{\prod[1+(v_i-1)(n+1),v_i(n+1)-1]} = w_v$, is in $L(X)$.

To prove this fact, we invoked Corollary 3.4, which in fact says a bit more; it implies that u can be extended to a point $x \in X$ by appending only G-letters to u. Note that since each w_v contains only n^d sites, and since all letters of x outside u are in G, x does not contain any connected components of B-letters with size more than n^d. Therefore, $x \in X_{n^d}$, which implies that $u \in L(X_{n^d})$. By counting the possible choices for the collection (w_v), we see that

$$|L_{[1,k(n+1)]^d}(X_{n^d})| \geq |L_{[1,n]^d}(X) \cap \langle G^{\partial[1,n]^d}\rangle|^{k^d}.$$

By taking logs of both sides, dividing by $(k(n + 1))^d$, and letting $k \to \infty$, we see that

$$h(X_{n^d}) \geq \frac{1}{(n+1)^d} \log |L_{[1,n]^d}(X) \cap \langle G^{\partial[1,n]^d}\rangle|.$$

We now recall that in the proof of (B), we showed that $\frac{1}{(n+1)^d} \log |L_{[1,n]^d}(X) \cap \langle G^{\partial[1,n]^d}\rangle|$ is within $\frac{3d \log |A|}{n}$ of $h(X)$, and so we have shown that $h(X_{n^d}) \geq h(X) - \frac{3d \log |A|}{n}$ for all n, implying that $h(X_n) \to h(X)$ as $n \to \infty$.

For any $\mathbb{Z}^d$ aperiodic ergodic measure-theoretic dynamical system (Y, μ, S_v) with $h(\mu) < h(X)$, there then exists n for which $h(\mu) < h(X_n)$. We recall from Section 2 that any $\mathbb{Z}^d$ SFT with the UFP is a measure-theoretic universal model, and so there exists a measure ν on X_n so that $(X_n, \nu, \sigma_v) \cong (Y, \mu, S_v)$. Since the support of ν is contained in X_n, clearly it is contained in X as well, and we have verified (C).

$\square$

Proof of (D). We prove that μ is Bernoulli by using the property of quite weak Bernoulli as defined in [10].

Definition 3.13. A measure μ on $A^{\mathbb{Z}^d}$ is called **quite weak Bernoulli** if for all $\epsilon > 0$,

$$\lim_{n \to \infty} d\left(\mu|_{[-n(1-\epsilon),n(1-\epsilon)]^d \cup \mathbb{Z}^d \setminus [-n,n]^d}, \mu|_{[-n(1-\epsilon),n(1-\epsilon)]^d} \times \mu|_{\mathbb{Z}^d \setminus [-n,n]^d}\right) = 0.$$

It is known that quite weak Bernoulli measures are Bernoulli (for instance, in [10], they note that it implies the property of very weak Bernoulli as defined in [17], and that Theorem 1.1 of [17] shows that very weak Bernoulli measures are Bernoulli), and so it suffices to show that μ is quite weak Bernoulli.

It is shown in [10] that μ is quite weak Bernoulli if and only if for all $\epsilon > 0$,

$$(12) \quad \lim_{n \to \infty} \min\left\{\alpha \; : \; \mu\big(\{\eta \in L_{\mathbb{Z}^d \setminus [-n,n]^d}(X) \; : \right.$$

$$\left. d\left(\mu|_{[-n(1-\epsilon),n(1-\epsilon)]^d}, \mu^\eta|_{[-n(1-\epsilon),n(1-\epsilon)]^d}\right) < \alpha\}\big) > 1 - \alpha\right\} = 0,$$

where μ^η is the conditional distribution on $[-n,n]^d$ of μ given η. Since μ is an MRF with conditional probabilities Λ^δ, we can replace μ^η by Λ^δ, where $\delta := \eta|_{\partial[-n,n]^d}$. Therefore, it suffices to show that for all $\epsilon > 0$,

$$(13) \quad \lim_{n \to \infty} \min\left\{\alpha \; : \; \mu\big(\{\delta \in L_{\partial[-n,n]^d}(X) \; : \right.$$

$$\left. d\left(\mu|_{[-n(1-\epsilon),n(1-\epsilon)]^d}, \Lambda^\delta|_{[-n(1-\epsilon),n(1-\epsilon)]^d}\right) < \alpha\}\big) > 1 - \alpha\right\} = 0.$$

We first note that combining Lemma 3.5 with Corollary 3.10 yields the fact that for any finite $T \subseteq \mathbb{Z}^d$, $\mu([B^T]) \leq N^{-|T|}$, where $N = \lfloor \frac{1}{2}(\epsilon^{-1} - 4d - 4)\rfloor > \frac{1}{5\epsilon}$. By summing over all possible paths of B-letters from $\partial[-n(1-\epsilon), n(1-\epsilon)]^d$ to $\underline{\partial}[-n,n]^d$, this implies that

$$\mu\left([-n(1-\epsilon), n(1-\epsilon)]^d \leftrightarrow \underline{\partial}[-n,n]^d\right) \leq \sum_{L=n\epsilon}^\infty (2d)^L N^{-L} = \frac{2d}{N-2d}\left(\frac{2d}{N}\right)^{n\epsilon}.$$

Since $N > 4d$, $\mu\left([-n(1-\epsilon), n(1-\epsilon)]^d \leftrightarrow \underline{\partial}[-n,n]^d\right) < 2^{-\epsilon n}$. Therefore, with μ-probability at least $1 - 2^{-\epsilon n}$ there exists a closed contour of G-letters containing $[-n(1-\epsilon), n(1-\epsilon)]^d$ in its interior and contained within $[-n,n]^d$. Since μ is an MRF with conditional probabilities Λ^δ, $\mu|_{[-n,n]^d} = \sum_\delta \mu([\delta])\Lambda^\delta$, where the sum is over $\delta \in L_{\partial[-n,n]^d}(X)$. Clearly, there then exists a set $S \subseteq L_{\partial[-n,n]^d}(X)$ with $\mu([S]) > 1 - 2^{-0.5\epsilon n}$ so that for any $\delta \in S$, the Λ^δ-probability that there exists a closed contour of G-letters containing $[-n(1-\epsilon), n(1-\epsilon)]^d$ and contained in $[-n,n]^d$ is at least $1 - 2^{-0.5\epsilon n}$.

As in the proof of (A), for any $\delta \in S$ and any $u \in L_{[-n,n]^d}(X)$ such that $u\delta \in L(X)$, we define $\gamma(u)$ to be the unique maximal closed contour of G-letters contained within $[-n,n]^d$, which contains $[-n(1-\epsilon), n(1-\epsilon)]^d$ with Λ^δ probability at

least $1-2^{-0.5\epsilon n}$ by definition of S. Also define $B(u)$ to be the set of sites of $[-n,n]^d$ outside $\gamma(u)$, and $D(u) = u|_{B(u)}$. Then $[\{u \in L_{[-n,n]^d}(X) : u\delta \in L(X)\}]$ can be written as a disjoint union of sets $[L_{[-n,n]^d}(X) \cap \langle D(u)\rangle)]$, meaning that except for a set of Λ^δ-measure at most $2^{-0.5\epsilon n}$, $\Lambda^\delta|_{[-n(1-\epsilon),n(1-\epsilon)]^d}$ can be written as a weighted average of $\mu\left(x|_{[-n(1-\epsilon),n(1-\epsilon)]^d} \mid [D(u)]\right)$ over possible values of $D(u)$. Since μ is an MRF, this is in fact a weighted average of $\Lambda^{\gamma(u)}|_{[-n(1-\epsilon),n(1-\epsilon)]^d}$. Finally, by Fact 3.8 from the proof of (A), for any $\gamma(u) \neq \gamma(u')$ containing $[-n(1-\epsilon),n(1-\epsilon)]^d$,

$$d\left(\Lambda^{\gamma(u)}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}, \Lambda^{\gamma(u')}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}\right) < Z(2n)^d 2^{-\frac{n\epsilon}{8d}}.$$

So, for each $\delta \in S$, except for a set of measure at most $2^{-0.5\epsilon n}$, $\Lambda^\delta|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$ can be decomposed as a weighted average of $\Lambda^{\gamma(u)}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$, and each pair $\Lambda^{\gamma(u)}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$, $\Lambda^{\gamma(u')}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$ has total variational distance less than $Z(2n)^d 2^{-\frac{n\epsilon}{8d}}$ for some universal constant Z. Therefore, for any $\delta, \delta' \in S$, the pair $\Lambda^\delta|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$, $\Lambda^{\delta'}|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$ has total variational distance less than $2^{-0.5\epsilon n} + Z(2n)^d 2^{-\frac{n\epsilon}{8d}}$.

Recall that except for a set of μ-measure at most $2^{-0.5\epsilon n}$, $\mu|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$ can be decomposed as a weighted average over $\Lambda^\delta|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}$ for $\delta \in S$. This means that for any $\delta \in S$,

$$d\left(\mu|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}, \Lambda^\delta|_{[-n(1-2\epsilon),n(1-2\epsilon)]^d}\right) < 2 \cdot 2^{-0.5\epsilon n} + Z(2n)^d 2^{-\frac{n\epsilon}{8d}}.$$

As $n \to \infty$, the right-hand side of this inequality approaches 0, and $\mu(S)$ approaches 1. We have then verified (13), so μ is quite weak Bernoulli, and therefore Bernoulli. $\qquad\square$

Proof of (E). Denote by X_μ the support of μ. Consider any topological factor map $\phi : X_\mu \to Y$ where Y does not consist of a single fixed point. Since ϕ is a sliding block code, there are configurations $w, w' \in L(X_\mu)$ with common shape S and $a \neq b$ in the alphabet of Y for which $\phi([w]) \subseteq [a]$ and $\phi([w']) \subseteq [b]$.

Since $w, w' \in L(X_\mu)$, $\mu([w])$ and $\mu([w'])$ are positive. We note that by Corollary 3.10, μ is the weak limit of Λ^{δ_N} for any sequence $\delta_N \in L_{\partial[-N,N]^d}(X)$ of boundaries consisting only of G-letters. Therefore, there exist N and $\delta = \delta_N \in L_{\partial[-N,N]^d}(X)$ consisting only of G-letters for which $\Lambda^\delta([w]), \Lambda^\delta([w']) > 0$. In particular, this means that there exist configurations $u, u' \in L_{[-N,N]^d}(X)$ so that $u|_{\partial[-N,N]^d} = u'|_{\partial[-N,N]^d} = \delta$, $u|_S = w$, and $u'|_S = w'$.

By Corollary 3.11, $\mu([\delta]) > 0$ since δ consists only of G-letters. Since μ is a MRF with conditional probabilities Λ^δ and u, u' both have δ on their inner boundaries and are in $L(X)$, $\mu([u]), \mu([u']) > 0$ as well. Since μ is the unique measure of maximal entropy, it is ergodic, and so there exists a point $x \in X_\mu$ such that the limiting frequency as $n \to \infty$ of occurrences of δ in $x|_{[-n,n]^d}$ is positive. It is easy to see that we can take a subset of these occurrences of δ which contains no pair of overlapping δ and which still has a positive limiting frequency β. Denote by $V \subset \mathbb{Z}^d$ the set of centers of this set of disjointly occurring δ in x; then $d(V) := \lim_{n\to\infty} \frac{|V \cap [-n,n]^d|}{|[-n,n]^d|} = \beta$.

Then, for any M, define $x_M = x|_{[-M,M]^d} \in L(X_\mu)$ and $\eta_M = x_M|_{\partial[-M,M]^d} \in L(X_\mu)$. Clearly, since $\eta_M \in L(X_\mu)$, $\mu([\eta_M]) > 0$. Create a family F_M of configurations by independently replacing each of the configurations $x_M|_{v+[-N,N]^d}$ $(v \in V)$ contained entirely within x_M by either u or u'. Since u and u' agree on their inner

boundary, all configurations in F_M are locally admissible for X. Then, since each configuration in F_M has η_M on the inner boundary and since μ is an MRF with conditional probabilities Λ^δ, $F_M \subseteq L(X_\mu)$. However, since the frequency $d(V)$ of V is equal to $\beta > 0$,

$$(14) \qquad \lim_{M\to\infty} \frac{\log |F_M|}{|[-M,M]^d|} = \beta \log 2.$$

Finally, we note that ϕ is injective on F_M; for any nonequal $y, y' \in F_M$, there is $v \in V$ so that $y|_{v+[-N,N]^d} = u$ and $y'|_{v+[-N,N]^d} = u'$ or vice versa, implying that $y|_{v+S} = w$ and $y|_{v+S} = w'$ or vice versa, finally implying that $(\phi(y))(v) = a$ and $(\phi(y'))(v) = b$ or vice versa. Either way, $\phi(y) \neq \phi(y')$. We have then shown that $\phi(F_M) \subseteq L(\phi(X_\mu))$ and $|F_M| = |\phi(F_M)|$, and combining this with (14) shows that $h(\phi(X_\mu)) > 0$. Since ϕ was arbitrary, we have shown that X_μ has topologically completely positive entropy, verifying (E).

$\square$

4. Unrelatedness of the ϵ-fullness condition to mixing conditions

Often in $\mathbb{Z}^d$ symbolic dynamics, useful properties of an SFT follow from some sort of mixing condition, such as topological mixing, block gluing or uniform filling. Examples of such properties following from some uniform mixing condition are being a measure-theoretically universal model ([25]), the existence of dense periodic points ([20]), and entropy minimality (nonexistence of proper subshift of full entropy) ([24]). In this section, we give some examples demonstrating that the ϵ-fullness condition has no relationship whatsoever with these mixing conditions by exhibiting ϵ-full nearest neighbor $\mathbb{Z}^d$ SFTs at various levels of the mixing hierarchy for arbitrarily small ϵ.

Example 4.1 (Non-topologically mixing)**.** The SFT from Example 2.33 can be made ϵ-full for arbitrarily small ϵ by increasing the parameter n, but is never topologically mixing, since there are no points which contain both a 0 and a 1.

Example 4.2 (Topologically mixing but not block gluing)**.** In [24], a $\mathbb{Z}^2$ SFT called the **checkerboard island shift** is defined. We briefly describe its properties here, but refer the reader to [26] for more details. The checkerboard island shift C is defined by a set of legal 2×2 configurations, namely those appearing in Figure 2, plus the 2×2 configuration of all blank symbols. Note that C is not a nearest neighbor SFT; we will deal with this momentarily. It is shown in [24] that C is topologically mixing, and in fact more is observed; any finite configuration $w \in L(C)$ is a subconfiguration of a configuration $w' \in L(C)$ with only blank symbols on the inner boundary. It is also shown in [24] that C does not have the uniform filling property. In fact, their proof also shows that C is not block gluing; they observe that any square checked configuration surrounded by arrows (e.g. the central 8×8 block of Figure 2) forces a square configuration containing it of almost twice the size (e.g. the 14×14 configuration in Figure 2), which clearly precludes block gluing.

First, we define a version of C which is nearest neighbor by passing to the second higher block presentation: define C' to have alphabet A' consisting of the 2×2 configurations from $L(C)$, and the only adjacency rule is that adjacent 2×2 blocks must agree on the pair of letters along their common edge. (For instance, for letters a, b, c, d in the alphabet of C, $\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}$ $\begin{smallmatrix} b & c \\ d & a \end{smallmatrix}$ would be a legal adjacent pair of letters in
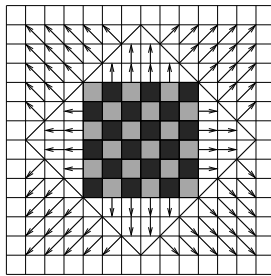
FIGURE 2. A sample configuration from C

C'.) C' is topologically conjugate to C, and so shares all properties of C described above. The reader can check that the alphabet A' of C' has size 79.

We can now make versions of C' which are ϵ-full for small ϵ; for any N, define C'_N to have alphabet $A'_N = G \cup A'$, where G is a set of N "free" symbols with the following adjacency rules: each G-letter can appear next to any other G-letter, and each G-letter can legally appear next to any letter from A' consisting of four blank symbols from the original alphabet of C. The reader may check that the addition of these new symbols does not affect the above arguments proving topological mixing and absence of block gluing, and so each C'_N is topologically mixing, but not block gluing. Also, since any G-letter can be legally followed in any direction by any other G-letter, C'_N is $\frac{79}{N+79}$-full. Clearly C'_N can then be made ϵ-full for arbitrarily small ϵ by increasing the parameter N.

This example can be trivially extended to a nearest neighbor $\mathbb{Z}^d$ SFT $C'^{(d)}_N$ by keeping the same alphabet and adjacency rules and adding no transition rules along the extra $d-2$ dimensions. Clearly $C'^{(d)}_N$ is still topologically mixing and not block gluing, and can be made ϵ-full for arbitrarily small ϵ by taking N large.

Example 4.3 (Block gluing but not uniform filling). In [26], a nearest neighbor $\mathbb{Z}^2$ SFT called the **wire shift** is defined. We briefly describe its properties here, but refer the reader to [26] for more details. For any integer N, the wire shift W_N has alphabet $A_N = G \cup B$, where B consists of six "grid symbols" illustrated in Figure 3 and G is a set of "blank tiles" labeled with integers from $[1, N]$. The adjacency rules are that neighboring letters must have edges which "match up" in the sense of Wang tiles; for instance, the leftmost symbol from Figure 3 could not appear immediately above the second symbol from the left.
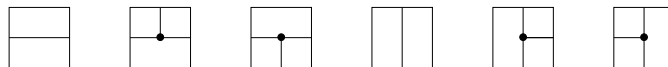

FIGURE 3. The symbols B from the alphabet of W_N

It is shown in Corollary 3.3 from [26] that W_N is block gluing for any N. In Lemma 3.4 from [26], it is shown that for any $N \geq 2$, W_N is not entropy minimal, in other words W_N contains a proper subshift with topological entropy $h(W_N)$. However, Lemma 2.7 from [26] shows that any $\mathbb{Z}^d$ SFT with the UFP is entropy minimal, and so W_N does not have the UFP for any $N \geq 2$.

Again any letter of G can be legally followed in any direction by any other letter of G, implying that W_N is $\frac{6}{N+6}$-full. Clearly then W_N can be made ϵ-full for arbitrarily small ϵ by increasing the parameter N.

This example can also be trivially extended to a nearest neighbor $\mathbb{Z}^d$ SFT $W_N^{(d)}$ by keeping the same alphabet and adjacency rules and adding no transition rules along the extra $d-2$ dimensions. Clearly $W_N^{(d)}$ is still block gluing and does not have the UFP, and can be made ϵ-full for arbitrarily small ϵ by taking N large.

Example 4.4 (Uniform filling)**.** Any full shift is ϵ-full for any $\epsilon > 0$, and obviously has the UFP.

Remark 4.5. We note that we can say a bit more about the extended checkerboard island shift C_N' of Example 4.2 for large N. As we already noted, it was shown in [24] that every $w \in L(C)$ can be extended to a configuration $w' \in L(C)$ with only blank symbols from the alphabet of C on the inner boundary. Then, for any N, we claim that any configuration $w \in L(C_N')$ can be extended to a configuration $w' \in L(C_N')$ with only G-letters on the inner boundary: any configuration $w \in L(C_N')$ looks like a recoded version of a configuration from C, possibly with some G-letters replacing letters of A' consisting of four blanks from the original alphabet of A, and so the same extension proved in [24] guarantees that w can be extended to w' with only G-letters and four-blank letters from A', which can itself be surrounded by a boundary of G-letters. If we take any N for which $\frac{79}{N+79} < \epsilon_d$ and denote the unique measure of maximal entropy on C_N' by μ, then by Corollary 3.11, $\mu([w']) > 0$, clearly implying that $\mu([w]) > 0$. We have then shown that μ has full support, and in particular have shown the following:

Theorem 4.6. *For any $\epsilon > 0$, there exists an ϵ-full nearest neighbor $\mathbb{Z}^2$ SFT with unique measure of maximal entropy μ whose support is not block gluing.*

In other words, there really is no uniform mixing condition implied by the ϵ-fullness property, even hidden within the support of the unique measure of maximal entropy μ.

5. COMPARISON WITH EXISTING SUFFICIENT CONDITIONS FOR UNIQUE M.M.E.

We for now focus on property (A) from Theorem 3.1, i.e. the fact that for any d and small enough ϵ, any ϵ-full nearest neighbor $\mathbb{Z}^d$ SFT has a unique measure of maximal entropy. In this section, we will attempt to give proper context by giving some examples of conditions in the literature related to our condition, and some examples suggesting refinements of our result. We first need a definition.

Definition 5.1. For a nearest neighbor $\mathbb{Z}^d$ SFT X, a letter a of the alphabet A is a **safe symbol** if a is a legal neighbor of every letter of A in every cardinal direction $\pm \vec{e}_i$.

Example 5.2. In [22], the classical Dobrushin uniqueness criterion for Markov random fields is used to prove the following result:

Theorem 5.3 ([22], Proposition 5.1)**.** *For any nearest neighbor $\mathbb{Z}^d$ SFT X with alphabet A such that at least $|A|\left(\frac{2d}{\sqrt{4d^2+1}+1}\right)$ of the letters of A are safe symbols, X has a unique measure of maximal entropy.*

This seems to be the first result to show that a large proportion of safe symbols is enough to guarantee uniqueness.

Example 5.4. In [9], methods from percolation theory, following techniques from [1], were used to prove the following result.

Theorem 5.5 ([9], Theorem 1.17). *For any nearest neighbor $\mathbb{Z}^d$ SFT X with alphabet A such that at least $|A|(\sqrt{1 - p_c(\mathbb{Z}^d)})$ of the letters of A are safe symbols, X has a unique measure of maximal entropy, where $p_c(\mathbb{Z}^d)$ is the critical probability for site percolation on the d-dimensional hypercubic lattice.*

(This is not the precise statement of their theorem, but is an equivalent reformulation which better contrasts with Theorem 5.3.) We will not define $p_c(\mathbb{Z}^d)$ or discuss percolation theory here; for a good introduction to the subject, see [14].

Theorem 5.5 is stronger than Theorem 5.3 for $d = 2$: it is known that $p_c(\mathbb{Z}^2) > 0.5$, so $\sqrt{1 - p_c(\mathbb{Z}^2)} < \sqrt{0.5} < \frac{4}{\sqrt{17}+1}$. However, for large d, $p_c(\mathbb{Z}^d) = \frac{1+o(1)}{2d}$ ([18]), therefore $\frac{2d}{\sqrt{4d^2+1}+1} \approx \frac{2d}{2d+1} < 1 - \frac{1}{4d} \approx \sqrt{1 - p_c(\mathbb{Z}^d)}$, implying that Theorem 5.3 is stronger for large d.

Example 5.6. A slightly more general result comes from [15], which requires another definition.

Definition 5.7. For any nearest neighbor $\mathbb{Z}^d$ SFT X with alphabet A, the **generosity of X** is

$$G(X) = \frac{1}{|A|} \min_{\delta \in L_{N_0}(X)} |\{a \in A \; : \; a\delta \in L(X)\}|.$$

Theorem 5.8 ([15], Theorem 1.12). *Any nearest neighbor $\mathbb{Z}^d$ SFT with generosity at least $\frac{1}{1+p_c(\mathbb{Z}^d)}$ has a unique measure of maximal entropy.*

The strength of Theorem 5.8 is that it allows one to consider SFTs without safe symbols. For instance, the n-**checkerboard** $\mathbb{Z}^d$ SFT, defined by alphabet $A = \{1, \dots, n\}$ and the adjacency rule that no letter may be adjacent to itself in any cardinal direction, has generosity $1 - \frac{2d}{n}$, which satisfies the hypotheses of Theorem 5.8 for large n. However, it has no safe symbols, and so cannot satisfy the hypotheses of Theorems 5.3 and 5.5.

Remark 5.9. We note that Haggstrom also showed in [15] that it is not possible for any $d \geq 2$ to give a lower bound on $G(X)$ in Theorem 5.8 which would imply uniqueness of measure of maximal entropy for all $\mathbb{Z}^d$ SFTs (not just nearest neighbor): Theorem 1.13 from [15] states that for any $d \geq 2$ and any $\epsilon > 0$, there exists a $\mathbb{Z}^d$ SFT X with more than one measure of maximal entropy such that $G(X) \geq 1 - \epsilon$. This implies that such a uniform lower bound would be impossible for our Theorem 3.1 as well, since any nearest neighbor $\mathbb{Z}^d$ SFT with generosity more than $1 - \epsilon$ is clearly ϵ-full.

One common property of each of these conditions is that they require all letters of A to satisfy some fairly stringent adjacency properties. For instance, if A contains even one letter which has only a single allowed neighbor in some direction, then it has at most one safe symbol and its generosity equals $\frac{1}{|A|}$ (the minimum possible amount). The strength of the ϵ-fullness condition is that it allows the existence of a small set of letters with bad adjacency properties, as long as the rest of the symbols are "close enough" to being safe symbols.

6. CRITICAL VALUE OF δ_d

Given Corollary 3.2, it is natural to define the optimal value of δ_d which guarantees uniqueness of the measure of maximal entropy.

Definition 6.1. For every d, we define

$$\alpha_d := \inf\{\alpha \ : \ \exists \text{a nearest neighbor } \mathbb{Z}^d \text{ SFT } X$$

with more than one measure of maximal entropy for which $h(X) \geq (\log|A|) - \alpha.\}$

We can determine α_1 exactly.

Theorem 6.2. $\alpha_1 = \log 2$.

Proof. Suppose that X is a nearest neighbor $\mathbb{Z}$ SFT with alphabet A and more than one measure of maximal entropy. Since any measure of maximal entropy of X is supported in an irreducible component of X, clearly X has at least two irreducible components of entropy $h(X)$. (See [21] for the definition of irreducible components and a simple proof of this fact.) Each of these components must then have alphabet size at least $e^{h(X)}$, and so $|A| \geq 2e^{h(X)}$. This implies that $(\log|A|) - h(X) \geq \log 2$. Therefore, if $(\log|A|) - h(X) < \log 2$, X has a unique measure of maximal entropy, implying that $\alpha_1 \geq \log 2$.

However, the same idea shows that there exists a nearest neighbor $\mathbb{Z}$ SFT X with multiple measures of maximal entropy and for which $(\log|A|) - h(X) = \log 2$; take $X = \{1,\ldots,n\}^{\mathbb{Z}} \cup \{n+1,\ldots,2n\}^{\mathbb{Z}}$, the union of two disjoint full shifts on n symbols. Therefore, $\alpha_1 \leq \log 2$, completing the proof.

$\square$

We will now show that α_2 is less than $\log 2$, meaning that in two dimensions, it is possible to have an SFT X in which two disjoint portions of the alphabet induce different measures of maximal entropy and can still coexist within the same point of X (unlike the one-dimensional case).

Theorem 6.3. $\alpha_2 < \log 2$.

Proof. In [8], an example is given of a strongly irreducible nearest neighbor $\mathbb{Z}^2$ SFT with two measures of maximal entropy, later called the **iceberg model** in the literature. We quickly recall the definition of this SFT.

The iceberg model, which we denote by $\mathcal{I}_M$, is defined by a positive integer parameter M. The alphabet is $A_M = \{-M,\ldots,-2,-1,1,2,\ldots,M\}$. The adjacency rules of $\mathcal{I}_M$ are that any letters with the same sign may neighbor each other, but a positive may only sit next to a negative if they are 1 and -1. It is shown in [8] that for any $M > 4e28^2$, $\mathcal{I}_M$ has exactly two measures of maximal entropy.

We will now give a lower bound on $h(\mathcal{I}_M)$ which is strictly greater than $\log M$, which will imply that $(\log|A_M|) - h(\mathcal{I}_M) < \log 2$. For any n, define the set P_n of all configurations with shape $[1,n]^2$ consisting of letters from $\{1,\ldots,M\}$. Clearly $P_n \subset L_{[1,n]^d}(\mathcal{I}_M)$. By ergodicity of the Bernoulli measure giving each positive letter equal probability (or simply the Strong Law of Large Numbers), if we define G_n to be the set of configurations in P_n with at least $\frac{n^2}{M^5}$ occurrences of the configuration $\begin{smallmatrix} & 1 & \\ 1 & 1 & 1 \\ & 1 & \end{smallmatrix}$, then $\lim_{n\to\infty} \frac{\log|G_n|}{n^2} = \log M$. In any $u \in G_n$, it is simple to choose at least $\frac{n^2}{5M^5}$ occurrences of $\begin{smallmatrix} & 1 & \\ 1 & 1 & 1 \\ & 1 & \end{smallmatrix}$ with disjoint centers. Then, one can construct a set $f(u)$

of at least $2^{\frac{n^2}{5M^5}}$ configurations in $L_{[1,n]^2}(\mathcal{I}_M)$ by independently either replacing the center of each of these 1-crosses by -1 or leaving it as a 1. It is easy to see that $f(u) \cap f(u') = \varnothing$ for any $u \neq u' \in G_n$. Therefore, $|L_{[1,n]^2}(\mathcal{I}_M)| \geq |\bigcup_{u \in G_n} f(u)| \geq 2^{\frac{n^2}{5M^5}}|G_n|$, implying that

$$h(\mathcal{I}_M) = \lim_{n \to \infty} \frac{\log |L_{[1,n]^2}(\mathcal{I}_M)|}{n^2} \geq \lim_{n \to \infty} \frac{\log |G_n|}{n^2} + \frac{\log 2}{5M^5} = \log M + \frac{\log 2}{5M^5}.$$

Therefore, $(\log |A_M|) - h(\mathcal{I}_M) \leq \log 2 - \frac{\log 2}{5M^4}$, and so since $\mathcal{I}_M$ has two measures of maximal entropy for $M = 10000 > 4e28^2$, $\alpha_2 \leq \log 2 - \frac{\log 2}{5 \cdot 10^{20}}$.

$\square$

Theorem 3.1, along with Theorem 2.32, implies that $\alpha_d \geq \frac{\epsilon_d^2}{4d} > 10^{-6}d^{-5}(2d)^{-16d}$. It is then natural to look for upper bounds on α_d as well, since this lower bound is almost certainly nonoptimal. Unsurprisingly, the sequence α_d is nonincreasing.

Theorem 6.4. $\alpha_{d+1} \leq \alpha_d$ for all d.

Proof. Fix any d and any $\epsilon > 0$. By definition, there exists a nearest neighbor $\mathbb{Z}^d$ SFT X (with alphabet A) with $\mu_1 \neq \mu_2$ measures of maximal entropy for which $h(X) > (\log |A|) - \alpha_d - \epsilon$. We may then define $X^{\mathbb{Z}}$ to be the nearest neighbor $\mathbb{Z}^{d+1}$ SFT containing all $x \in A^{\mathbb{Z}^{d+1}}$ for which each $x|_{\mathbb{Z}^d \times \{j\}} \in X$. In other words, $X^{\mathbb{Z}}$ has the adjacency rules for X in the $x_1, \ldots, x_d$-directions, and no restrictions at all in the x_{d+1}-direction.

Then clearly $h(X^{\mathbb{Z}}) = h(X) > (\log |A|) - \alpha_d - \epsilon$. Also, it is not hard to check that $\mu_1^{\mathbb{Z}}$ and $\mu_2^{\mathbb{Z}}$ are measures of maximal entropy for $X^{\mathbb{Z}}$, where $\mu_i^{\mathbb{Z}}$ is the independent product of countably many copies of μ_i in the x_{d+1}-direction. Therefore, $\alpha_{d+1} \leq \alpha_d + \epsilon$. Since ϵ was arbitrary, we are done.

$\square$

Our final result shows that $\lim_d \alpha_d = 0$, and that the decay rate is at least polynomial.

Theorem 6.5. *There exists a constant B so that $\alpha_d \leq \frac{1}{\lfloor Bd^{0.25}(\log d)^{-0.75} \rfloor}$ for all d.*

Proof. The main tool in our construction is a theorem of Galvin and Kahn ([13]) about phase transitions for the hard-core shift with activities. Specifically, define a Gibbs measure with activity $\lambda \in \mathbb{R}^+$ on the $\mathbb{Z}^d$ hard-core shift $\mathcal{H}_d$ to be a measure μ with the property that for any n and any configurations w with shape S and δ with shape ∂S such that $w\delta \in L(\mathcal{H}_d)$,

$$\mu([w] \mid [\delta]) = \frac{\lambda^{\# \text{ ones in } w}}{\sum_{v \in \{0,1\}^S \text{ s.t. } v\delta \in L(\mathcal{H}_d)} \lambda^{\# \text{ ones in } v}}.$$

In other words, the conditional probability of w given a fixed δ is proportional to $\lambda^{\# \text{ ones in } w}$. One way to create Gibbs measures is to fix any boundary conditions $\delta_n \in L_{\partial[-n,n]^d}(\mathcal{H}_d)$ on larger and larger cubes $[-n,n]^d$, and take a weak limit point of the sequence of conditional measures $\mu(\cdot \mid [\delta])$. For the hard-core model, two boundary conditions of interest are $\delta_{e,n}$, which contains 1 on all $v \in \partial[-n,n]^d$ with $\sum v_i$ odd and 0 on all $v \in \partial[-n,n]^d$ with $\sum v_i$ even, and $\delta_{o,n}$, which does the reverse.

The main result of [13] states that there exists a universal constant C with the following property: for any dimension d and activity level λ which is greater than

$Cd^{-0.25}(\log d)^{0.75}$, the sequences of conditional measures $\mu(\cdot \mid [\delta_{e,n}])$ and $\mu(\cdot \mid [\delta_{o,n}])$ approach respective weak limits μ_e and μ_o, which are distinct Gibbs measures with activity λ on $\mathcal{H}_d$. Though it is not explicitly stated in [13], it is well-known that each of μ_o and μ_e is a shift by one unit in any cardinal direction of the other, and in particular that both μ_o and μ_e are invariant under any shift in $(2\mathbb{Z})^d$. (This is mentioned in, among other places, [3].) The strategy used in [13] to show that $\mu_o \neq \mu_e$ is quite explicit: it is shown that for λ satisfying the hypothesis of the theorem, $\mu_e(x(0) = 1) < \mu_o(x(0) = 1)$.

We now define $\mathcal{H}_{N,d}$ to be the nearest-neighbor $\mathbb{Z}^d$ SFT with N safe symbols $\{0_1, \ldots, 0_N\}$ and a symbol 1 which cannot appear next to itself in any cardinal direction; $\mathcal{H}_{N,d}$ is a version of the usual hard-core shift where the symbol 0 has been "split" into N copies. For any Gibbs measure μ on the hard-core model $\mathcal{H}_d$ with activity $\lambda = \frac{1}{N}$, define a measure $\hat{\mu}$ on $\mathcal{H}_{N,d}$ by "splitting" the measure of any cylinder set $[w]$ uniformly over all $N^{\#\text{zeroes in } w}$ ways of assigning subscripts to 0 symbols in w. It is easily checked that any such measure $\hat{\mu}$ has the uniform conditional probabilities property from the conclusion of Proposition 2.16 (Proposition 1.20 from [9]), which stated that all measures of maximal entropy have this property. In fact, Proposition 1.21 from [9] gives a partial converse: for strongly irreducible SFTs, any shift-invariant measure with uniform conditional probabilities must be a measure of maximal entropy.

Since $\mathcal{H}_{N,d}$ is clearly strongly irreducible, this would show that $\widehat{\mu_e}$ and $\widehat{\mu_o}$ are measures of maximal entropy on $\mathcal{H}_{N,d}$, were it not for the fact that these measures are not shift-invariant. However, their average $\frac{1}{2}(\widehat{\mu_e} + \widehat{\mu_o})$ clearly shares the uniform conditional probability property, and is shift-invariant, and so is a measure of maximal entropy on $\mathcal{H}_{N,d}$. In fact, it is the unique measure of maximal entropy on $\mathcal{H}_{N,d}$. We will show, however, that the direct product of $\mathcal{H}_{N,d}$ with itself can have multiple measures of maximal entropy.

Define the nearest neighbor $\mathbb{Z}^d$ SFT $\mathcal{H}_{N,d}^2$ with alphabet $\{0_1, \ldots, 0_N, 1\}^2$, where the adjacency rules from $\mathcal{H}_{N,d}$ are separately enforced in each coordinate. In other words, $(0_1, 1)$ may appear next to $(1, 0_2)$ since $0_1 1$ and $1 0_2$ are each legal in $\mathcal{H}_{N,d}$, but $(0_1, 1)$ cannot appear next to $(1, 1)$: though $0_1 1$ is legal in $\mathcal{H}_{N,d}$, 11 is not. Define the measures $\nu_1 := \frac{1}{2}(\widehat{\mu_o} \times \widehat{\mu_o} + \widehat{\mu_e} \times \widehat{\mu_e})$ and $\nu_2 := \frac{1}{2}(\widehat{\mu_o} \times \widehat{\mu_e} + \widehat{\mu_e} \times \widehat{\mu_o})$ on $\mathcal{H}_{N,d}^2$. It should be obvious that both ν_1 and ν_2 have the uniform conditional probability property mentioned above and that both are shift-invariant. Since $\mathcal{H}_{N,d}^2$ is strongly irreducible (it still has a safe symbol $(0_1, 0_1)$), both ν_1 and ν_2 are therefore measures of maximal entropy on $\mathcal{H}_{N,d}^2$.

We claim that $\nu_1 \neq \nu_2$ as long as $N < C^{-1} d^{0.25} (\log d)^{-0.75}$. If this condition holds, then $\frac{1}{N} > Cd^{-0.25}(\log d)^{0.75}$, and so by [13], $\mu_e(x(0) = 1) < \mu_o(x(0) = 1)$ for the hard-core shift with activity $\lambda = \frac{1}{N}$. Clearly, $\widehat{\mu_e}(x(0) = 1) = \mu_e(x(0) = 1)$ and $\mu_o(x(0) = 1) = \widehat{\mu_o}(x(0) = 1)$, so $\widehat{\mu_e}(x(0) = 1) < \widehat{\mu_o}(x(0) = 1)$. For brevity, denote these probabilities by α and β respectively. By definition, $\nu_1(x(0) = (1,1)) = \frac{1}{2}(\alpha^2 + \beta^2)$ and $\nu_2(x(0) = (1,1)) = \frac{1}{2}(\alpha\beta + \alpha\beta)$. These are equal if and only if $\alpha = \beta$, which is not the case. Therefore, $\nu_1 \neq \nu_2$ and $\mathcal{H}_{N,d}^2$ has multiple measures of maximal entropy.

Our final observation is that any way of placing arbitrary letters on even sites and only pairs $(0_i, 0_j)$ on odd sites yields configurations in $L(\mathcal{H}_{N,d}^2)$, and so $h(\mathcal{H}_{N,d}^2) \geq \frac{1}{2}(\log(N+1)^2 + \log N^2) = \log N(N+1)$. Therefore, for this SFT, $\log|A| - h(\mathcal{H}_{N,d}^2) \leq$

$\log(N+1)^2 - \log N(N+1) = \log\left(1 + \frac{1}{N}\right) < \frac{1}{N}$. Choosing $B = C^{-1}$ and $N = \lfloor C^{-1} d^{0.25}(\log d)^{-0.75} \rfloor$ now completes our proof.

$\square$

We therefore know that $d^{d(-16-o(1))} < \alpha_d < d^{-0.25+o(1)}$. We imagine that the true decay rate is much closer to the upper bound than the lower, but have no conjectures as to the exact rate.

Acknowledgments

References

[1] J. van den Berg, *A uniqueness condition for Gibbs measures, with application to the 2-dimensional ising antiferromagnet*, Comm. Math. Phys. **152** (1993), no. 1, 161–166.

[2] J. van den Berg and C. Maes, *Disagreement percolation in the study of Markov random fields*, Ann. Probab. **22** (1994), no. 2, 749–763.

[3] J. van den Berg and J.E. Steif, *Percolation and the hard-core lattice gas model*, Stochastic Process. Appl. **49** (1994), no. 2, 179–197.

[4] R. Berger, *The undecidability of the domino problem*, Mem. Amer. Math. Soc. **66** (1966)

[5] J. Bertoin, F. Martinelli, and Y. Peres, *Lectures on probability theory and statistics: Ecole d'Ete de Probabilites de Saint-Flour XXVII*, Springer, Lecture Notes in Mathematics, vol. 1717, 1999.

[6] F. Blanchard, *Fully positive topological entropy and topological mixing*, Symbolic dynamics and applications (in honor of R.L. Adler), AMS Contemporary Mathematics, Providence, RI, **135** (1992), 95–105.

[7] M. Boyle, R. Pavlov and M. Schraudner, *Multidimensional sofic shifts without separation, and their factors*, Trans. Amer. Math. Soc., **362** (2010), 4617–4653.

[8] R. Burton and J. Steif, *Non-uniqueness of measures of maximal entropy for subshifts of finite type*, Ergodic Theory Dynam. Systems, **14** (1994), 213–335.

[9] R. Burton and J. Steif, *New results on measures of maximal entropy*, Israel J. Math., **89** (1995), 275–300.

[10] R. Burton and J. Steif, *Quite weak Bernoulli with exponential rate and percolation for random fields*, Stochastic Process. Appl. **58** (1995), 35–55.

[11] R. Burton and J. Steif, *Some 2-d symbolic dynamical systems: entropy and mixing*, Ergodic Theory and $\mathbb{Z}^d$-actions, London Math. Soc. Lecture Notes Series, Cambridge, **228** (1996), 297–306.

[12] M. Campanino and L. Russo, *An upper bound on the critical percolation probability for the three-dimensional cubic lattice*, Ann. Probab. **13** (1985), no. 2, 478–491.

[13] D. Galvin and J. Kahn, *On phase transition in the hard-core model on $\mathbb{Z}^d$*, Combin. Probab. Comput. **13** (2004), 137–164.

[14] Geoffrey Grimmett, *Percolation*, Springer-Verlag, Grundlehren der mathematischen Wissenschaften, vol. 321, 1999.

[15] O. Häggström, *On phase transitions for subshifts of finite type*, Israel J. Math. **94** (1996), 319–352.

[16] M. Hochman and T. Meyerovitch, *A Characterization of the Entropies of Multidimensional Shifts of Finite Type*, Annals of Math., to appear.

[17] J. W. Kammeyer, *A complete classification of the two-point extensions of a multidimensional Bernoulli shift*, J. Anal. Math. **54** (1990), no. 1, 113–163.

[18] H. Kesten, *Asymptotics in high dimensions for percolation*, Disorder in physical systems, Oxford Univ. Press, New York, (1990), 219–240.

[19] Ker-I Ko, *Complexity theory of real functions*, Progress in Theoretical Computer Science, Birkhauser, 1991.

[20] S. Lightwood, Morphisms from non-periodic $\mathbb{Z}^2$ subshifts I: constructing embeddings from homomorphisms, Ergodic Theory Dynam. Systems, **23** (2003), 587–609.

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

[22] N. Markley and M. Paul, *Maximal measures and entropy for $\mathbb{Z}^\nu$ subshifts of finite type*, *Classical mechanics and dynamical systems*, Lecture Notes in Pure and Appl. Math. **70** (1981), 135–157.

[23] M. Misiurewicz, *A short proof of the variational principle for a $\mathbb{Z}_+^n$-action on a compact space*, Asterisque, **40** (1975), 147–157.

[24] A. Quas and A. Sahin, *Entropy gaps and locally maximal entropy in $\mathbb{Z}^d$ subshifts*, Ergodic Theory Dynam. Systems **23** (2003), 1227–1245.

[25] E. A. Robinson Jr. and A. A. Sahin., *Modeling ergodic, measure preserving actions on $\mathbb{Z}^d$ shifts of finite type*, Monatsh. Math. **132** (2001), no. 3, 237–253.

[26] M. Schraudner, *Projectional entropy and the electrical wire shift*, Discrete Cont. Dyn. Syst. A **26** (2010), no. 1, 333–346.

[27] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, Graduate Texts in Mathematics, vol. 79, 1982.

[28] H. Wang, *Proving theorems by pattern recognition II*, AT&T Bell Labs. Tech. J. **40** (1961), 1–41.

[29] S.A. Zuev, *A lower bound for a percolation threshold for a square lattice*, (Russian) Vestnik Moskov. Univ. Ser. I Mat. Mekh. 1988, no. 5, 59–61; translation in Moscow Univ. Math. Bull. **43** (1988), no. 5, 66–69.

RONNIE PAVLOV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, 2360 S. GAYLORD ST., DENVER, CO 80208

E-mail address: rpavlov@du.edu

URL: www.math.du.edu/∼rpavlov/