Acta Appl Math (2013) 126:297–317
DOI 10.1007/s10440-013-9819-2

Independence entropy of $\mathbb{Z}^d$-shift spaces

Erez Louidor · Brian Marcus · Ronnie Pavlov

Received: 25 February 2010 / Accepted: 13 August 2010 / Published online: 26 March 2013

Abstract We introduce a concept of independence entropy for symbolic dynamical systems. This notion of entropy measures the extent to which one can freely insert symbols in positions without violating the constraint defined by the shift space. We show that for a certain class of one-dimensional shift spaces X, the independence entropy coincides with the limiting, as d tends to infinity, topological entropy of the dimensional shift defined by imposing the constraints of X in each of the d cardinal directions. This is of interest because for these shift spaces independence entropy is easy to compute. Thus, while in these cases, the topological entropy of the d-dimensional shift ($d \geq 2$) is difficult to compute, the limiting topological entropy is easy to compute. In some cases, we also compute the rate of convergence of the sequence of d-dimensional entropies. This work generalizes earlier work on constrained systems with unconstrained positions.

Mathematics Subject Classification 37B10 · 94A15

Keywords Entropy · Shift of finite type · Sofic shift · Unconstrained positions

1 Introduction/Overview

Topological entropy is the most fundamental numerical invariant associated to a d-dimensional shift space. When $d = 1$ and the shift space is a shift of finite type (SFT) or sofic

Dedicated to the memory of Hang Kim.

E. Louidor
Google, Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043, USA
e-mail: erez@google.com

B. Marcus (✉)
Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
e-mail: marcus@math.ubc.ca

R. Pavlov
Department of Mathematics, University of Denver, Denver, CO 80208, USA
e-mail: rpavlov@du.edu

shift, the topological entropy is easy to compute as the log of the largest eigenvalue of a nonnegative integer matrix. However, when $d = 2$ there is no known explicit expression for the topological entropy of SFTs, even for the simplest nontrivial examples. For instance, the one-dimensional golden mean shift is the set of all bi-infinite sequences of ‘0’s and ‘1’s such that 11 never appears. The topological entropy of this shift is well known to be the log of the golden mean, which is approximately $\ln(1.618)$. For the two-dimensional golden mean shift, defined to be the set of all configurations of ‘0’s and ‘1’s on $\mathbb{Z}^2$ such that 11 and $\substack{1\\1}$ never appear, there is no known expression for the topological entropy.

One analogously defines the d-dimensional golden mean shift for any $d \geq 1$. It is well known that the topological entropy is a non-increasing function in d and thus has a limit. It is not so well known that the value of this limit is known in this case; namely it is $\frac{\ln(2)}{2} \approx \ln(1.414)$ (see [12, 18]).

One way of interpreting this result is as follows. For any d, call a site in $\mathbb{Z}^d$ even (resp., odd) if the sum of its coordinates is even (resp., odd). Now, if we assign value 0 to each even site, then we are free to assign arbitrary binary values to each of the odd sites; the ‘0’s in the even sites prevent two adjacent sites from both taking value 1. Equivalently, for any configuration of ‘0’s and ‘1’s on all of $\mathbb{Z}^d$ such that all even sites have value 0, then if we freely change the values at all odd sites in any way, the resulting configuration belongs to the golden mean shift. The number of possible restrictions of configurations of this form to a d-dimensional rectangular cube of even side length L is $2^{L^d/2}$. The asymptotic growth rate of the number of such configurations, as $L \to \infty$, is $\frac{\ln(2)}{2}$. Thus, in any dimension, we get a contribution of $\frac{\ln(2)}{2}$ to the topological entropy of the d-dimensional golden mean shift. We call this quantity the “independence entropy” of the golden mean shift since it arises from independently choosing values at a maximal collection of specified sites, while fixing values at other sites. Note that the independence entropy is exactly the limiting entropy of the d-dimensional golden mean shift.

It is not hard to show that the topological entropy of the d-dimensional golden mean shift dominates its independence entropy and we show that for all d, there is a strict gap; evidently, this gap vanishes as $d \to \infty$.

In this paper, we define a general notion of independence entropy for an arbitrary d-dimensional shift space X over an arbitrary finite alphabet $\mathcal{A}$ (Sect. 4), using an associated “multi-choice” shift space (introduced in Sect. 3). The alphabet of this shift space is the collection of non-empty subsets of $\mathcal{A}$, and a configuration is allowed if and only if whenever one replaces each symbol $S \subseteq \mathcal{A}$ with any element of S, one obtains an allowed configuration in X.

In Sect. 4, we observe that the topological entropy of a shift space always dominates its independence entropy and show that the independence entropy can be computed explicitly for any 1-dimensional sofic shift.

In Sect. 5, we introduce the notion of an “axial product” of d 1-dimensional shift spaces X_i over the same alphabet: namely, the d-dimensional shift space consisting of all configurations whose restriction to every “row” in the i-th direction belongs to X_i, for each $i = 1, \dots, d$. For example, the d-dimensional golden mean shift is the axial product of d copies of the 1-dimensional golden mean shift. We show that the independence entropy of the axial product is dominated by the minimum of the independence entropies of the X_i. In the isotropic case, i.e., when all the X_i are the same shift space (call it X), we show that the independence entropy of the axial product is the same as that of X. We also note that the entropies of the d-dimensional versions of a $\mathbb{Z}$-shift space are nonincreasing in d, and therefore approach a “limiting d-dimensional entropy.”

In Sect. 6, we investigate the relationship between independence entropy and limiting d-dimensional entropy. First, we use the results of Sects. 4 and 5 to show that independence entropy is always less than or equal to limiting d-dimensional entropy. As mentioned above, for the golden mean shifts, as d tends to infinity, the gap between topological entropy and independence entropy disappears, and so the independence entropy and limiting d-dimensional entropy are equal. We show that this holds for some general classes of shifts of finite type as well as for a specific example, the 3-checkerboard SFT, thereby answering a question posed in [7]. For this SFT, as well as the golden mean shift, we also show that the d-dimensional topological entropy converges exponentially, and we compute the exact exponent of convergence (this is an improvement on an upper bound on the exponent given in [12]). We do not know of any example of a 1-dimensional shift space where the limiting topological entropy is not equal to the independence entropy.[1]

A d-dimensional shift space may naturally be regarded as a topological $\mathbb{Z}^d$-action with d generators: namely, the unit shifts in each of the d dimensions. As described above, a 1-dimensional shift space X defines a d-dimensional shift space, namely the isotropic axial product. In the "limit," X also defines an action of $\mathbb{Z}^{\mathbb{N}}$ on a sort of infinite dimensional isotropic axial product, and we note that the limiting topological entropy coincides with the topological entropy of this action (see [17] for an introduction to topological entropy of general amenable group actions).

Finally, we remark that our work was motivated by a scheme, known as "constrained systems with unconstrained positions," for combining error-correcting codes and modulation codes in magnetic recording [4, 16, 19, 20, 22]. The rough idea there is that a 1-dimensional SFT or sofic shift over $\{0, 1\}$ represents a set of allowable sequences that can be recorded along data tracks in a storage medium. Various physical considerations dictate constraints that define the shift space. While these constraints substantially reduce the likelihood of error introduced in the reading process, some errors are unavoidable due to imperfections and noise in the components. Thus, it is desirable to have some error-correcting (ECC) capability. This can be achieved by reserving some specific positions for ECC parity bits, which are computed based on the values in other positions. Since the parity bits are not known in advance, it is necessary that any selection of '0's and '1's be allowed in those positions, without violating the constraints of the shift space, so that the desired constraints on recorded sequences still hold. This results in a very special type of multi-choice shift space.

2 Definitions

We begin with some definitions from symbolic dynamics and graph theory.

Let $\mathcal{A}$ be a finite alphabet, and fix a positive integer d. For a subset $P \subseteq \mathbb{Z}^d$, a *configuration on P over $\mathcal{A}$* is a mapping $x : P \to \mathcal{A}$. For any such x and $\mathbf{i} \in P$ we denote the element of $\mathcal{A}$ corresponding to $\mathbf{i}$ by $x_{\mathbf{i}}$ or $x(\mathbf{i})$. A *shift* of a configuration x on $B \subseteq \mathbb{Z}^d$ over $\mathcal{A}$ is the configuration y on $\mathbf{t} + B = \{\mathbf{j} + \mathbf{t} : \mathbf{j} \in B\}$ over $\mathcal{A}$ given by $y_{\mathbf{j}+\mathbf{t}} = x_{\mathbf{j}}$, $\mathbf{j} \in B$, for some $\mathbf{t} \in \mathbb{Z}^d$. For two configurations $x \in \mathcal{A}^P$ and $y \in \mathcal{A}^Q$, where $P, Q \subset \mathbb{Z}^d$ are disjoint, we denote by $x \cup y$ the configuration on $P \cup Q$ satisfying $(x \cup y)|_P = x$ and $(x \cup y)|_Q = y$. A configuration is finite if it is on a finite set. We say that a configuration x on a set P *appears* in a configuration y if there exists $\mathbf{t} \in \mathbb{Z}^d$ such that for all $\mathbf{i} \in P$, $y_{\mathbf{i}+\mathbf{t}} = x_{\mathbf{i}}$. For

[1] T. Meyerovitch and R. Pavlov have recently shown that the limiting topological entropy and independence entropy always coincide.

a set $\mathcal{F}$ of finite configurations, the $\mathbb{Z}^d$ *shift space over* $\mathcal{A}$ *defined by* $\mathcal{F}$, denoted $X_{\mathcal{F}}$, is the set of all configurations on $\mathbb{Z}^d$ over $\mathcal{A}$ that do not have a configuration of $\mathcal{F}$ appearing in them. The set $\mathcal{F}$ is usually called a *forbidden list*. A $\mathbb{Z}^d$ shift space X is called a *shift of finite type* (abbreviated SFT) if it can be defined by a finite forbidden list. Let X be a $\mathbb{Z}^d$ shift space. A configuration x on some set $P \subset \mathbb{Z}^d$ is called *globally admissible* if it appears in an element of X. If X is defined by a forbidden list $\mathcal{F}$, such a configuration x is called *locally admissible with respect to* $\mathcal{F}$ if it does not contain an element of $\mathcal{F}$. For a set $P \subseteq \mathbb{Z}^d$, we denote by $B_P(X)$ the set of all globally admissible configurations on P. Local admissibility of finite configurations of a $\mathbb{Z}^d$ SFT is decidable: there is an algorithm that given a finite forbidden list $\mathcal{F}$ and a finite configuration determines whether the configuration is locally admissible with respect to $\mathcal{F}$. In contrast, such an algorithm for determining global admissibility of a finite configuration for a general SFT only exists when $d = 1$ [2].

For an alphabet $\mathcal{A}$, a word of length ℓ is a configuration on $\{0, \dots, \ell - 1\}$ over $\mathcal{A}$. We use ε to denote the empty word and $\mathcal{A}^*$ to denote the set of all words over $\mathcal{A}$. We use the following conventional notation: for words $x, y \in \mathcal{A}^*$ and a nonnegative integer n, we use $|x|$ to denote the length of x, xy to denote the word formed by "concatenating y to the end of x" in the usual sense, and x^n to denote the word formed by concatenating x to itself n times ($x^0 = \varepsilon$). A $\mathbb{Z}$-SFT is called an *m-step SFT* if it can be defined by a forbidden list of words $\mathcal{F}$, such that the maximum length of a word in $\mathcal{F}$ is $m + 1$. For a $\mathbb{Z}$ shift space over X, the *language* of X, denoted $B(X)$, is the set of all globally admissible words of X, that is $B(X) = \bigcup_{\ell=0}^{\infty} B_{\{0,\dots,\ell-1\}}(X)$.

Let $G = (V, E)$ be a directed graph with a finite set of vertices V and finite set of edges E. For an edge $e \in E$ we denote by $\sigma(e)$ and $\tau(e)$ the starting and terminating vertices of e in G. A bi-infinite path in G is a sequence $(e_i)_{i=-\infty}^{\infty} \subseteq E$ of edges such that for all integers i, $\tau(e_i) = \sigma(e_{i+1})$. We also deal with finite paths, which are finite sequences $(e_i)_{i=1}^{\ell} \subseteq E$ of some length ℓ such that $\tau(e_i) = \sigma(e_{i+1})$ for $i = 1, \dots, \ell - 1$. Such a path is said to *start* at the vertex $\sigma(e_1)$ and *end* at the vertex $\tau(e_\ell)$. A *cycle* is a finite path that starts and ends at the same vertex. A cycle $(e_i)_{i=1}^{\ell}$ is *simple* if the vertices $\tau(e_1), \dots, \tau(e_\ell)$ are distinct. The set of all bi-infinite paths in G is a $\mathbb{Z}$ SFT over E called the *edge-shift defined by G*.

A $\mathbb{Z}^d$ shift space is *sofic* if there exists an alphabet $\mathcal{A}'$, an SFT X' over $\mathcal{A}'$, and a mapping $\psi : \mathcal{A}' \to \mathcal{A}$ such that

$$X = \left\{x \in \mathcal{A}^{\mathbb{Z}^d} : \text{ there exists } x' \in X' \text{ s.t. for all } \mathbf{i} \in \mathbb{Z}^d, x_{\mathbf{i}} = \psi\left(x'_{\mathbf{i}}\right)\right\}.$$

(We note that our definition differs slightly from the usual one, in which ψ could be any continuous shift-commuting map. However, the definitions are equivalent.)

For $d = 1$, we will make use of two other equivalent definitions of a sofic shift space, each of which requires a preliminary definition.

For a word $w \in B(X)$, the *follower set* of w in X, denoted by $F_X(w)$, is defined by

$$F_X(w) = \left\{z \in B(X) : wz \in B(X)\right\}.$$

A $\mathbb{Z}$ shift space is sofic if and only if it has only finitely many follower sets.

A *labeled graph $\mathcal{G}$ with edge labels in $\mathcal{A}$* is a pair $(G, \mathcal{L})$ where $G = (V, E)$ is a finite directed graph and $\mathcal{L} : E \to \mathcal{A}$ is an edge-labeling function. The concepts of finite and bi-infinite paths of $\mathcal{G}$ are inherited from G. A bi-infinite path $(e_i)_{i=-\infty}^{\infty}$ in $\mathcal{G}$ is said to *generate* the bi-infinite sequence $(w_i)_{i=-\infty}^{\infty} \in \mathcal{A}^{\mathbb{Z}}$ if $w_i = \mathcal{L}(e_i)$. Similarly, a finite path of length ℓ *generates* a word of length ℓ over $\mathcal{A}$.

A $\mathbb{Z}$ shift space X is sofic if and only if there exists a labeled graph $\mathcal{G}$ such that X is equal to the set of all bi-infinite sequences generated by bi-infinite paths of $\mathcal{G}$. Any such labeled graph $\mathcal{G}$ is called a *presentation* of X.

See [14] for proofs and more details on $\mathbb{Z}$ shift spaces.

Finally, we use $\mathbb{N}$ to denote the set of positive integers.

3 Multi-choice Shift Spaces

In this section, we introduce a concept which will be necessary in Sect. 4 for defining independence entropy. For any finite alphabet $\mathcal{A}$, let $\hat{\mathcal{A}}$ denote the set of all nonempty subsets of $\mathcal{A}$. Let X be a $\mathbb{Z}^d$ shift space over $\mathcal{A}$ and for a configuration $\hat{x}$ on $P \subseteq \mathbb{Z}^d$ over $\hat{\mathcal{A}}$, define the set of "fillings" of $\hat{x}$, denoted $\Phi(\hat{x})$, by

$$\begin{aligned}\Phi(\hat{x}) &= \left\{x \in \mathcal{A}^P : \text{for all } \mathbf{i} \in P, x_{\mathbf{i}} \in \hat{x}_{\mathbf{i}}\right\}\\ &= \prod_{\mathbf{i} \in P} \hat{x}_{\mathbf{i}}\end{aligned}$$

We define the *multi-choice shift space corresponding to* X, denoted $\hat{X}$, by

$$\hat{X} = \left\{\hat{x} \in \hat{\mathcal{A}}^{\mathbb{Z}^d} : \Phi(\hat{x}) \subseteq X\right\}.$$

It's easy to verify that if $\mathcal{F}$ is a forbidden list defining X then the set

$$\hat{\mathcal{F}} = \left\{\hat{x} : \hat{x} \in \hat{\mathcal{A}}^P \text{ for some } P \subseteq \mathbb{Z}^d \text{ and } \Phi(\hat{x}) \cap \mathcal{F} \neq \emptyset\right\}$$

is a forbidden list defining $\hat{X}$; so $\hat{X}$ is indeed a shift space over $\hat{\mathcal{A}}$. Note that if $\mathcal{F}$ is finite then so is $\hat{\mathcal{F}}$, and therefore the multi-choice shift space of an SFT is an SFT.

Example 1 Define G to be the *golden mean SFT* on $\{0, 1\}$ with forbidden list $\mathcal{F} = \{11\}$. Then G consists of all biinfinite 0-1 sequences which do not contain consecutive 1's. The multi-choice SFT $\hat{\mathsf{G}}$ corresponding to G has alphabet $\{\{0\}, \{1\}, \{0, 1\}\}$, and has forbidden list $\{\{1\}\{1\}, \{1\}\{0, 1\}, \{0, 1\}\{1\}, \{0, 1\}\{0, 1\}\}$.

In contrast, for $d > 1$, we don't know if the multi-choice shift space corresponding to a sofic shift is sofic. For $d = 1$, [20] shows that this true when $\mathcal{A} = \{0, 1\}$. The proof can be generalized to larger alphabets, and for completeness we provide it in the next theorem. (We cannot extend our proof to the case $d > 1$ because it uses the follower set definition of a $\mathbb{Z}$-sofic shift, for which there is no known extension to $d > 1$.)

Theorem 1 *Let X be a $\mathbb{Z}$ sofic shift over an alphabet $\mathcal{A}$. Then $\hat{X}$ is a $\mathbb{Z}$ sofic shift over $\hat{\mathcal{A}}$.*

Proof We recall that one of the definitions of a $\mathbb{Z}$ sofic shift allows us to prove that $\hat{X}$ is sofic by constructing a labeled graph presentation $\hat{\mathcal{G}} = ((V, E), \mathcal{L})$ for $\hat{X}$. The set of vertices V consists of all (finite) intersections of follower sets of words in X:

$$V = \left\{\bigcap_{i=1}^{k} F_X(w_i) : w_1, \ldots, w_k \in B(X), k \in \mathbb{N}\right\}.$$

We recall that any $\mathbb{Z}$ sofic shift has only finitely many follower sets, and so since X is sofic, V is finite. We proceed to define E. For a vertex $v = \bigcap_{i=1}^{k} F_X(w_i) \in V$, and symbol $\hat{a} \in \hat{\mathcal{A}}$ with $\hat{a} \subseteq v$, define

$$\delta(v, \hat{a}) = \bigcap_{a \in \hat{a}} \bigcap_{i=1}^{k} F_X(w_i a)$$

(note that $w_i a \in B(X)$ for every $a \in \hat{a}$ and $i = 1, 2, \dots, k$). It's easy to verify that $\delta(v, \hat{a}) = \{w \in \mathcal{A}^* : aw \in v \text{ for all } a \in \hat{a}\}$, and therefore $\delta(v, \hat{a})$ does not depend on the choice of k and $w_1, \dots, w_k$. The set E is now defined by

$$E = \big\{ \big(v, \hat{a}, \delta(v, \hat{a})\big) : v \in V, \hat{a} \subseteq v \big\},$$

and for an edge $e = (v, \hat{a}, \delta(v, \hat{a})) \in E$ we define $\sigma(e) = v$, $\tau(e) = \delta(v, \hat{a})$ and $\mathcal{L}(e) = \hat{a}$. We claim that $\hat{\mathcal{G}}$ is a presentation of $\hat{X}$. The following fact is handy for proving this and is easily verified by induction on the length of the word w.

Fact *For any vertex $v = \bigcap_{i=1}^{k} F_X(w_i) \in V$ and any word $w \in \hat{\mathcal{A}}^*$, there is a path starting at v and generating w in $\hat{\mathcal{G}}$ if and only if $\Phi(w) \subseteq v$, in which case the path ends at the vertex $\bigcap_{i=1}^{k} \bigcap_{z \in \Phi(w)} F_X(w_i z)$.*

Now, let $\hat{x} \in \hat{\mathcal{A}}^{\mathbb{Z}}$ be a bi-infinite sequence generated by a bi-infinite path $(e_i)_{i=-\infty}^{\infty}$ of $\hat{\mathcal{G}}$. We claim that $\hat{x} \in \hat{X}$. Otherwise, there must be an $x \in \Phi(\hat{x})$ such that $x \notin X$. This implies that there exists a forbidden word z of X, $j \in \mathbb{Z}$ and a nonnegative integer n such that $z = x_j \cdots x_{j+n}$. Consider the finite path $(e_i)_{i=j}^{j+n}$, and let $\hat{z} = \hat{x}_j \cdots \hat{x}_{j+n}$ be the word that it generates. Let $v = \sigma(e_j)$ be the starting vertex of this path. By the Fact above, $\Phi(\hat{z}) \subseteq v$. Since clearly $z \in \Phi(\hat{z})$, it follows that $z \in v$, which implies $z \in B(X)$, a contradiction.

Conversely, let $\hat{x} \in \hat{X}$. Let $i \in \mathbb{Z}$ and pick a nonnegative integer n. Set $w = \hat{x}_i \cdots \hat{x}_{i+n}$. By our assumption, clearly $\Phi(w) \subseteq B(X)$. Since $B(X) = F(\varepsilon)$, where ε denotes the empty word, it follows by the Fact above that there is a path generating $w = \hat{x}_i \cdots \hat{x}_{i+n}$ in $\hat{\mathcal{G}}$ starting at $F(\varepsilon)$. Since this is true for all such i and n, it follows by an application of König's Infinity Lemma [5] (or a standard compactness argument) that there is a bi-infinite path in $\hat{\mathcal{G}}$ generating $\hat{x}$. □

4 Independence Entropy

In this section, we introduce the notion of "independence entropy" of a shift space. Roughly, this is the part of the entropy resulting from inter-symbol independence in elements of the shift space. We need the following generalization of Fekete's Subadditivity Lemma to multivariate functions.

For an n-tuple $\mathbf{m} = (m_1, \dots, m_n) \in \mathbb{N}^n$, let $[\mathbf{m}]$ denote the Cartesian product $\prod_i \{0, \dots, m_i - 1\}$ and let $\mathrm{d}(\mathbf{m}_i) = \min\{m_1, \dots, m_n\}$. We say that a sequence of such n-tuples $(\mathbf{m}_i)_{i=1}^{\infty}$ diverges to ∞, denoted $\mathbf{m}_i \to \infty$, if $(\mathrm{d}(\mathbf{m}_i))_{i=1}^{\infty}$ diverges to ∞. Let $\bar{\mathbb{R}}$ denote the extended real numbers $\mathbb{R} \cup \{+\infty, -\infty\}$. We say that a function $f : \mathbb{N}^d \to \bar{\mathbb{R}}$ has a limit $L \in \bar{\mathbb{R}}$, denoted $\lim_{\mathbf{m} \to \infty} f(\mathbf{m}) = L$, if for every sequence $(\mathbf{m}_i)_{i=1}^{\infty} \subseteq \mathbb{N}^d$ with $\mathbf{m}_i \to \infty$, we have $\lim_{i \to \infty} f(\mathbf{m}_i) = L$. We call a function $f : \mathbb{N}^d \to [-\infty, \infty)$ *entry-wise subadditive*, if for any $(m_1, \dots, m_d) \in \mathbb{N}^d$, $i \in \{1, 2, \dots, d\}$ and $n \in \mathbb{N}$, it holds that

$$\begin{aligned} &f(m_1, \dots, m_{i-1}, m_i + n, m_{i+1}, \dots, m_d) \\ &\quad \le f(m_1, \dots, m_d) + f(m_1, \dots, m_{i-1}, n, m_{i+1}, \dots, m_d). \end{aligned}$$

If f is an entry-wise subadditive function, then

$$\lim_{\mathbf{m}\to\infty} \frac{f(\mathbf{m})}{|[\mathbf{m}]|} = \inf_{\mathbf{m}} \frac{f(\mathbf{m})}{|[\mathbf{m}]|},$$

where for a set S we use $|S|$ to denote the cardinality of S. See [3] for a proof.

Fix $d \in \mathbb{N}$ and a finite alphabet $\mathcal{A}$. Let X be a nonempty $\mathbb{Z}^d$ shift space over $\mathcal{A}$. It's easy to verify that the mapping $\mathbf{m} \mapsto \ln |B_{[\mathbf{m}]}(X)|$ for $\mathbf{m} \in \mathbb{N}^d$ (we define $\ln 0 = -\infty$) is entry-wise subadditive. The *topological entropy* of X, denoted $h(X)$, is given by

$$\begin{aligned} h(X) &= \lim_{\mathbf{m}\to\infty} \frac{\ln |B_{[\mathbf{m}]}(X)|}{|[\mathbf{m}]|} \\ &= \inf_{\mathbf{m}} \frac{\ln |B_{[\mathbf{m}]}(X)|}{|[\mathbf{m}]|}. \end{aligned} \tag{1}$$

If we replace $|B_{[\mathbf{m}]}(X)|$ in the definition by the number of locally admissible patterns on $[\mathbf{m}]$ with respect to some fixed forbidden list, the topological entropy remains the same. (See [6] or [9] for a proof in the case where X is a $\mathbb{Z}^d$ SFT; the proof in [9] easily generalizes to the case where X is any $\mathbb{Z}^d$ shift space.) The topological entropy is an important invariant of shift spaces (under a suitably defined notion of isomorphism). For a $\mathbb{Z}$ sofic shift there is a closed formula (up to computing the largest root of a polynomial) for determining the topological entropy. In general, it is very difficult to "compute" the topological entropy of a $\mathbb{Z}^d$ SFT. There are only a few non-degenerate $\mathbb{Z}^d$-SFTs with $d > 1$ for which the topological entropy is known [1, 11, 13, 15, 21].

We now define "independence entropy." For any d-tuple $\mathbf{m}$ and configuration $\hat{w} \in B_{[\mathbf{m}]}(\hat{X})$, the real number $\ln(|\Phi(\hat{w})|)/|[\mathbf{m}]|$ can be thought of as the contribution to the topological entropy resulting from independence between entries in elements of X as "captured" by $\hat{w}$. We define the independence entropy as the limit (as $\mathrm{d}(\mathbf{m}) \to \infty$) of the maximum possible such contribution. Precisely, observe that the mapping $\mathbf{m} \mapsto \max\{\ln |\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}]}(\hat{X})\}$ for $\mathbf{m} \in \mathbb{N}^d$ is entry-wise subadditive. The following quantity, which we call "independence entropy," was defined in [20] for $\mathbb{Z}$-sofic shifts with alphabet $\{0, 1\}$ under the name "maximum insertion rate," and involved a logarithm with base 2 rather than base e.

The *independence entropy* of X, denoted $h_{\text{ind}}(X)$, is defined by

$$\begin{aligned} h_{\text{ind}}(X) &= \lim_{\mathbf{m}\to\infty} \frac{\max\{\ln |\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}]}(\hat{X})\}}{|[\mathbf{m}]|} \\ &= \inf_{\mathbf{m}} \frac{\max\{\ln |\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}]}(\hat{X})\}}{|[\mathbf{m}]|} \end{aligned} \tag{2}$$

Example 2 We return to G, the golden mean SFT, and compute its independence entropy. Note that for any pattern $\hat{w} \in \hat{\mathsf{G}}$, $|\Phi(\hat{w})| = 2^i$, where i is the number of occurrences of the symbol $\{0, 1\}$ in $\hat{w}$. Therefore, the maximum value of $|\Phi(\hat{w})|$ for $\hat{w}$ of a fixed length m will be achieved by maximizing the number of occurrence of $\{0, 1\}$ in $\hat{w}$. For any m, it is clear that there exists $\hat{w} \in B_{[m]}(\hat{\mathsf{G}})$ with $\lceil \frac{m}{2} \rceil$ occurrences of $\{0, 1\}$, namely $\hat{w} = \{0, 1\}\{0\}\{0, 1\} \ldots$. It is also clear that there is no $\hat{w} \in B_{[m]}(\hat{\mathsf{G}})$ with more than this many occurrences of $\{0, 1\}$, since such $\hat{w}$ would contain two consecutive $\{0, 1\}$'s. Therefore,

$$h_{\mathrm{ind}}(\mathsf{G}) = \lim_{m\to\infty} \frac{\ln 2^{\lceil m/2 \rceil}}{m} = \frac{\ln 2}{2}.$$

We next prove some properties of the independence entropy. In [20, Theorem 18] it is shown how the independence entropy of a sofic $\mathbb{Z}$ shift space can be determined from a presentation of $\hat{X}$ when $\mathcal{A} = \{0, 1\}$. This result is easily generalized to larger alphabets and we state and prove it in the next theorem.

Theorem 2 *Let X be a $\mathbb{Z}$ sofic shift over an alphabet $\mathcal{A}$, and pick any presentation $\hat{\mathcal{G}} = ((V, E), \mathcal{L})$ of $\hat{X}$. Then*

$$h_{\mathrm{ind}}(X) = \max\left\{ \frac{\ln |\Phi(\hat{w})|}{|\hat{w}|} \ :\ \hat{w} \in \hat{X} \text{ is generated by a simple cycle of } \hat{\mathcal{G}} \right\}. \tag{3}$$

Proof We first note that the max on the right-hand side clearly exists, since the finite graph (V, E) can only have finitely many simple cycles. Denote the RHS of (3) by ν^*. Let $\hat{w}_*$ be a word generated by a simple cycle of $\hat{\mathcal{G}}$ such that $\nu^* = (\ln |\Phi(\hat{w}_*)|)/|\hat{w}_*|$. Set $\ell = |\hat{w}_*|$. For any $n \in \mathbb{N}$, clearly $\hat{w}_*^n \in B_{[n\ell]}(\hat{X})$ and $(\ln |\Phi(\hat{w}_*^n)|)/(n\ell) = \nu^*$. It follows that $\nu^* \le (\ln \max\{|\Phi(\hat{w})| : \hat{w} \in B_{[n\ell]}(\hat{X})\})/(n\ell)$. Taking the limit as $n \to \infty$, we obtain $\nu^* \le h_{\mathrm{ind}}(X)$. To complete the proof, we will show that $h_{\mathrm{ind}}(X) \le \nu^*$. We first claim that if $\hat{w} \in \hat{\mathcal{A}}^*$ is a word generated by a (possibly non-simple) cycle of $\hat{\mathcal{G}}$, then

$$\nu^* \ge \frac{\ln |\Phi(\hat{w})|}{|\hat{w}|}. \tag{4}$$

This is easily proved by induction on $|\hat{w}|$. If $|\hat{w}| = 1$, then the cycle generating $\hat{w}$ is simple and obviously (4) holds. For $|\hat{w}|>1$, let $\pi = (e_i)_{i=1}^{\ell}$ be a cycle of $\hat{\mathcal{G}}$ generating $\hat{w}$. Obviously (4) holds if π is simple. Otherwise, there exist integers $1 \le j<k \le \ell$ such that $\tau(e_j) = \tau(e_k)$. So both $\alpha = e_{j+1}, \ldots, e_k$ and $\beta = e_1, \ldots, e_j, e_{k+1}, \ldots, e_\ell$ are cycles in $\hat{\mathcal{G}}$. Let $\hat{x}$, $\hat{y}$ denote the words generated by α and β respectively. Then using the induction hypothesis on $|\hat{x}|$ and $|\hat{y}|$, we get

$$\begin{aligned} \frac{\ln |\Phi(\hat{w})|}{|\hat{w}|} &= \frac{|\hat{x}|}{|\hat{w}|} \frac{\ln |\Phi(\hat{x})|}{|\hat{x}|} + \frac{|\hat{y}|}{|\hat{w}|} \frac{\ln |\Phi(\hat{y})|}{|\hat{y}|} \\ &\le \frac{|\hat{x}|}{|\hat{w}|}\nu^* + \frac{|\hat{y}|}{|\hat{w}|}\nu^* \\ &= \nu^*. \end{aligned}$$

Now, for $n \in \mathbb{N}$, let $\hat{z}(n) \in B_{[n]}(\hat{X})$ be a word such that $|\Phi(\hat{z}(n))| = \max\{|\Phi(\hat{w})|\ :\ \hat{w} \in B_{[n]}(\hat{X})\}$, and let $(e_i^{(n)})_{i=0}^{n-1}$ be a path in $\hat{\mathcal{G}}$ generating $\hat{z}(n)$. Then by [16], p. 1432, it may be decomposed as follows. There exist an integer $0 \le m \le |V|$ and $2m$ integers $0 \le s_1 \le t_1 < s_2 \le t_2 < \cdots < s_m \le t_m < n$ such that for each $k = 1, \ldots, m$, $(e_i^{(n)})_{i=s_k}^{t_k}$ is a cycle, and

$n - \sum_k (t_k - s_k + 1) \le |V|$. Set $S = \bigcup_{k=1}^m \{s_k, \ldots, t_k\}$. Using (4), we have

$$\begin{aligned}\frac{\ln|\Phi(\hat{z}(n))|}{n} &= \sum_{k=1}^{m} \frac{\ln|\Phi(\mathcal{L}(e_{s_k}^{(n)}) \cdots \mathcal{L}(e_{t_k}^{(n)}))|}{n} + \sum_{i \in [n] \setminus S} \frac{\ln(|\Phi(\mathcal{L}(e_i^{(n)}))|)}{n} \\ &\le \sum_{k=1}^{m} \left(\frac{\ln|\Phi(\mathcal{L}(e_{s_k}^{(n)}) \cdots \mathcal{L}(e_{t_k}^{(n)}))|}{t_k - s_k + 1} \cdot \frac{t_k - s_k + 1}{n} \right) + \frac{|V|}{n} \ln|\mathcal{A}| \\ &\le \nu^* + \frac{|V|}{n} \ln|\mathcal{A}|.\end{aligned}$$

The result follows by taking the limit as $n \to \infty$ of both sides of the last inequality. □

We next show that the independence entropy of a shift space cannot exceed its topological entropy.

Theorem 3 *For any $\mathbb{Z}^d$ shift space X over an alphabet $\mathcal{A}$, $h_{\text{ind}}(X) \le h(X)$.*

Proof For every $\mathbf{m} \in \mathbb{N}^d$, let $\hat{z}(\mathbf{m}) \in B_{[\mathbf{m}]}(\hat{X})$ be a configuration such that $|\Phi(\hat{z}(\mathbf{m}))| = \max\{|\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}]}(\hat{X})\}$. Obviously, $\Phi(\hat{z}(\mathbf{m})) \subseteq B_{[\mathbf{m}]}(X)$, which implies $(1/|[\mathbf{m}]|) \times \ln|\Phi(\hat{z}(\mathbf{m}))| \le (1/|[\mathbf{m}]|) \ln|B_{[\mathbf{m}]}(X)|$. Taking the limit of both sides as $\mathbf{m} \to \infty$, we obtain the result. □

5 Limiting *d*-Dimensional Entropy and Axial Products

We begin by defining a construction of a $\mathbb{Z}^d$ shift space from lower dimensional shift spaces (first defined in [15]), which will be used later in the section to define a notion of "limiting d-dimensional entropy" of a $\mathbb{Z}$ shift space. We then relate the topological and independence entropies of the resulting shift space to the respective entropies of the lower dimensional shift spaces used to construct it.

For a configuration $x \in \mathcal{A}^{\mathbb{Z}^d}$ on $\mathbb{Z}^d$ over $\mathcal{A}$ and $1 \le i \le d$, a *row in direction i* of x is a configuration y on $\mathbb{Z}$ over $\mathcal{A}$ defined by specifying some $m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_d \in \mathbb{Z}$ and then taking $y_k = x_{(m_1,\ldots,m_{i-1},k,m_{i+1},\ldots,m_d)}$ for all $k \in \mathbb{Z}$. Let $X_1, \ldots, X_d$ be $\mathbb{Z}$ shift spaces over $\mathcal{A}$, we define the *axial product* of $X_1, \ldots, X_d$ to be the $\mathbb{Z}^d$ shift space

$$X_1 \otimes \cdots \otimes X_d := \left\{ z \in \mathcal{A}^{\mathbb{Z}^d} : \forall i \in \{1, \ldots, d\} \text{ and row } y \text{ in direction } i \text{ of } z,\ y \in X_i \right\}.$$

More generally, one can define the axial product of a $\mathbb{Z}^{d_1}$ shift space with a $\mathbb{Z}^{d_2}$ shift space (both over the same alphabet), and show that this product is associative. Most of our results below can be generalized to this more general setting; however to simplify the notation, we only treat axial products of $\mathbb{Z}$ shift spaces. For $i \in \{1, 2, \ldots, d\}$, any configuration x on $B \subseteq \mathbb{Z}$ corresponds to the (d-dimensional) configuration $x^{(i)}$ on the Cartesian product $\{0\}^{i-1} \times B \times \{0\}^{d-i} \subseteq \mathbb{Z}^d$ defined by $x^{(i)}_{(0,\ldots,0,k,0,\ldots,0)} = x_k$ for all $k \in B$. It's easy to verify that if $\mathcal{F}_i$ is a forbidden list defining the shift space X_i, then $\mathcal{F} = \bigcup_i \{x^{(i)} : x \in \mathcal{F}_i\}$ is a forbidden list defining $X_1 \otimes \cdots \otimes X_d$; thus the axial product of shift spaces is a shift space. If $X_1 = \cdots = X_d = X$, we call $X_1 \otimes \cdots \otimes X_d$ the *d-fold axial power of* X and abbreviate it by $X^{\otimes d}$. If $\hat{X}_i$ is the multi-choice shift space corresponding to X_i for $i \in \{1, \ldots, d\}$, then it is easy to verify that the multi-choice shift space corresponding to $X_1 \otimes \cdots \otimes X_d$ is $\hat{X}_1 \otimes \cdots \otimes \hat{X}_d$.

The next theorem relates topological and independence entropies of $\mathbb{Z}$ shift spaces and their axial product.

Theorem 4 *Let $X_1, \dots, X_d, X$ be $\mathbb{Z}$ shift spaces over $\mathcal{A}$. Then the following statements hold*:

(1) $h_{\text{ind}}(X_1 \otimes \cdots \otimes X_d) \le \min_i h_{\text{ind}}(X_i)$
(2) $h(X_1 \otimes \cdots \otimes X_d) \le \min_i h(X_i)$
(3) $h_{\text{ind}}(X^{\otimes d}) = h_{\text{ind}}(X)$

Proof of (1) For each X_i let $\hat{X}_i$ be the multi-choice shift space corresponding to it and denote by S the multi-choice shift space $\hat{X}_1 \otimes \cdots \otimes \hat{X}_d$ corresponding to $X_1 \otimes \cdots \otimes X_d$. Fix $i \in \{1, \dots, d\}$. For $m \in \mathbb{N}$, define $\mathbf{m} = (m_1, \dots, m_d) \subseteq \mathbb{N}^d$ by $m_j = 1$ for $j \in \{1, \dots, d\} \backslash \{i\}$ and $m_i = m$. Since $B_{[\mathbf{m}]}(S) = B_{[\mathbf{m}]}(\hat{X}_1 \otimes \cdots \otimes \hat{X}_d) \subseteq B_{[m]}(\hat{X}_i)$,

$$\frac{\ln \max\{|\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}]}(S)\}}{|[\mathbf{m}]|} \le \frac{\ln \max\{|\Phi(\hat{w})| : \hat{w} \in B_{[m]}(\hat{X}_i)\}}{m}$$

(strictly speaking, a configuration on $[\mathbf{m}]$ is not a word of length m; rather, there is an obvious $1-1$ correspondence from $B_{[\mathbf{m}]}(S)$ to $B_{[m]}(\hat{X}_i)$ that preserves $|\Phi|$). Therefore by (2), $h_{\text{ind}}(X_1 \otimes \cdots \otimes X_d) \le h_{\text{ind}}(X_i)$. Since i is arbitrary, the result follows.

Proof of (2) This is similar to the proof of (1), so we omit it here.

Proof of (3) Let $Y = X^{\otimes d}$. By the proof of (1), it's enough to show $h_{\text{ind}}(Y) \ge h_{\text{ind}}(X)$. Let $f : \mathbb{N} \to \mathbb{N}$ be any function satisfying $\lim_{i\to\infty}(f(i)/i) = \infty$. For $i \in \mathbb{N}$, let $\mathbf{m}_i = (i, i, \dots, i, f(i)) \in \mathbb{N}^d$ be the d-tuple with every entry but the last equal to i, and the last entry equal to $f(i)$. Set $\ell(i) = (d-1)(i-1) + f(i)$, and let $\hat{z}(i) \in B_{[\ell(i)]}(\hat{X})$ be a word such that $|\Phi(\hat{z}(i))| = \max\{|\Phi(\hat{w})| : \hat{w} \in B_{[\ell(i)]}(\hat{X})\}$. Define the configuration $\hat{y}(i) \in \hat{\mathcal{A}}^{[\mathbf{m}_i]}$ by

$$\hat{y}(i)_{\mathbf{j}} = \hat{z}(i)_{\psi(\mathbf{j})}, \quad \mathbf{j} \in [\mathbf{m}_i],$$

where $\psi : \mathbb{Z}^d \to \mathbb{Z}$ is the function $\psi(j_1, \dots, j_d) = \sum_k j_k$. We claim that $\hat{y}(i) \in B_{[\mathbf{m}_i]}(\hat{Y})$. Indeed, let $\hat{z}'(i) \in \hat{X}$ be a configuration such that $\hat{z}'(i)_j = \hat{z}(i)_j$ for $j \in [\ell(i)]$ and consider the configuration $\hat{y}'(i) \in \hat{\mathcal{A}}^{\mathbb{Z}^d}$ defined by

$$\hat{y}'(i)_{\mathbf{j}} = \hat{z}'(i)_{\psi(\mathbf{j})}, \quad \mathbf{j} \in \mathbb{Z}^d.$$

Note that every row of $\hat{y}'(i)$ is a shift of $\hat{z}'(i)$; thus $\hat{y}'(i) \in \hat{Y}$. Since, clearly, $\hat{y}(i)$ appears in $\hat{y}'(i)$, it follows that $\hat{y}(i) \in B_{[\mathbf{m}_i]}(\hat{Y})$. Consequently,

$$\left|\Phi\left(\hat{y}(i)\right)\right| \le \max\left\{\left|\Phi(\hat{w})\right| : \hat{w} \in B_{[\mathbf{m}_i]}(\hat{Y})\right\}. \tag{5}$$

We next lower bound $\ln|\Phi(\hat{y}(i))|$. Set $S = \{j \in \mathbb{Z} : (i-1)(d-1) \le j < f(i)\}$. Then

$$\begin{aligned}
\ln\left|\Phi\left(\hat{y}(i)\right)\right| &= \sum_{\mathbf{j}\in[\mathbf{m}_i]} \ln\left|\Phi\left(\hat{y}(i)_{\mathbf{j}}\right)\right| \\
&= \sum_{k\in[\ell(i)]} \left(\left|\psi^{-1}\left(\{k\}\right) \cap [\mathbf{m}_i]\right| \ln\left|\Phi\left(\hat{z}(i)_k\right)\right|\right) \\
&\ge i^{d-1} \sum_{k\in S} \ln\left|\Phi\left(\hat{z}(i)_k\right)\right|
\end{aligned}$$

$$= i^{d-1}\Bigg(\sum_{k\in[\ell(i)]} \ln\big|\Phi\big(\hat{z}(i)_k\big)\big| - \sum_{k\in[\ell(i)]\setminus S} \ln\big|\Phi\big(\hat{z}(i)_k\big)\big|\Bigg)$$
$$\geq i^{d-1}\ln\big|\Phi\big(\hat{z}(i)\big)\big| - 2(d-1)(i-1)i^{d-1}\ln|\mathcal{A}|,$$

where we used the fact that for $k \in S$, it holds that $\psi^{-1}(\{k\}) \cap [\mathbf{m}_i] = \{(j_1, \ldots, j_{d-1}, k - \sum_t j_t) : (j_1, \ldots, j_{d-1}) \in [i]^{d-1}\}$. Combining the last inequality with (5), we have

$$\frac{\ln\max\{|\Phi(\hat{w})| : \hat{w} \in B_{[\mathbf{m}_i]}(\hat{Y})\}}{|[\mathbf{m}_i]|} \geq \frac{i^{d-1}(\ln|\Phi(\hat{z}(i))| - 2(d-1)(i-1)\ln|\mathcal{A}|)}{|[\mathbf{m}_i]|}$$
$$\geq \frac{\ln|\Phi(\hat{z}(i))|}{\ell(i)} - \frac{2(d-1)(i-1)\ln|\mathcal{A}|}{f(i)}.$$

Taking the limit of both sides as $i \to \infty$, we obtain $h_{\text{ind}}(Y) \geq h_{\text{ind}}(X)$. □

The main application of axial products will be the following definition. Let X be a $\mathbb{Z}$ shift space over $\mathcal{A}$. For any $\mathbf{m} \in \mathbb{N}^d$, any configuration of $B_{[\mathbf{m}]\times\{0\}}(X^{\otimes(d+1)})$ can clearly be thought of as a configuration in $B_{[\mathbf{m}]}(X^{\otimes d})$ by simply ignoring the last 0 coordinate of each entry. It follows by (1) that $h(X^{\otimes d})$ is non-increasing in d, and so the limit $\lim_{d\to\infty} h(X^{\otimes d})$ exists, and we denote it by $h_\infty(X)$. We think of $h_\infty(X)$ as a sort of limiting d-dimensional entropy of X.

6 The Relation Between h_{ind} and h_∞

The remainder of the paper is devoted to the relationship between the (seemingly unrelated) quantities $h_{\text{ind}}(X)$ and $h_\infty(X)$ for a $\mathbb{Z}$ shift space X. By Theorem 3 and Part 3 of Theorem 4, $h(X^{\otimes d}) \geq h_{\text{ind}}(X)$ for every d, and so

$$h_\infty(X) \geq h_{\text{ind}}(X). \tag{6}$$

For the few examples for which we can compute $h_\infty(X)$, it turns out to be equal to $h_{\text{ind}}(X)$. We suspect that this is true for all shift spaces X. In this section we exhibit a few families of $\mathbb{Z}$ shift spaces where this equality holds. In all but one of these, we prove that the rate of convergence of $h(X^{\otimes d})$ to $h_{\text{ind}}(X)$ is exponential.

6.1 The RLL$(\mathcal{D}, \infty)$ SFT

Let $\mathcal{D} \leq \mathcal{K}$ be nonnegative integers. The shift space RLL$(\mathcal{D}, \mathcal{K})$ is the $\mathbb{Z}$-SFT over the alphabet $\{0, 1\}$ consisting of all bi-infinite binary sequences in which every "run" of '0's has length at most $\mathcal{K}$ and every two consecutive '1's are separated by at least $\mathcal{D}$ '0's. It is defined by the forbidden list $\{0^{\mathcal{K}+1}\} \cup \{10^i1 : i \in [\mathcal{D}]\}$. The parameter $\mathcal{K}$ is allowed to be ∞, in which case there is no upper bound on the length of a run of '0's and RLL$(\mathcal{D}, \infty)$ is defined by the forbidden list $\{10^i1 : i \in [\mathcal{D}]\}$. These shifts are widely used in digital storage systems based on optical and magnetic recording. The independence entropy of RLL$(\mathcal{D}, \mathcal{K})$ is computed in [20, Theorem 25] and is given by

$$h_{\text{ind}}\big(\text{RLL}(\mathcal{D}, \mathcal{K})\big) = \frac{\lfloor(\mathcal{K}-\mathcal{D})/(\mathcal{D}+1)\rfloor \ln 2}{\lfloor(\mathcal{K}+1)/(\mathcal{D}+1)\rfloor(\mathcal{D}+1)} \tag{7}$$

$$h_{\mathrm{ind}}\big(\mathrm{RLL}(\mathcal{D},\infty)\big)=\frac{\ln 2}{\mathcal{D}+1} \tag{8}$$

(We remind the reader that in [20], the independence entropy is defined using $\log_2$ instead of ln. Accordingly, their formulas, presented here, are scaled by a factor of $\ln 2$.) In [18], Ordentlich and Roth show that for all nonnegative integers $\mathcal{D}$, $h_\infty(\mathrm{RLL}(\mathcal{D},\infty))=h_{\mathrm{ind}}(\mathrm{RLL}(\mathcal{D},\infty))$. In fact, they show that

$$h\big(\mathrm{RLL}(\mathcal{D},\infty)^{\otimes d}\big)=\frac{\ln 2}{\mathcal{D}+1}+O\bigg(\frac{\ln^2(d(\mathcal{D}+1))}{d(\mathcal{D}+1)}\bigg).$$

6.2 $\mathbb{Z}$-SFT's with Zero Independence Entropy

We now give another family of $\mathbb{Z}$-SFT's X for which $h_{\mathrm{ind}}(X)=h_\infty(X)$. Note that from (7), for $\mathcal{K}\le 2\mathcal{D}$, $h_{\mathrm{ind}}(\mathrm{RLL}(\mathcal{D},\mathcal{K}))=0$. In [10] it is shown that indeed $h_\infty(\mathrm{RLL}(\mathcal{D},\mathcal{K}))=0$ for such $\mathcal{D}$ and $\mathcal{K}$. In this section we show that the latter equality holds for all $\mathbb{Z}$-SFT's with zero independence entropy.

We need the concept of an irreducible component of a finite directed graph and we summarize it here. For a labeled graph $\mathcal{G}=((V,E),\mathcal{L})$ and a subset $U\subseteq V$, we call the graph $(U,E_U,\mathcal{L}|_{E_U})$ where $E_U=\{e\in E\ :\sigma(e)\in U,\tau(e)\in U\}$ the *subgraph of* $\mathcal{G}$ *induced by* U. For two vertices $u,v\in V$ we say that u *is reachable from* v if there is a (finite) path in $\mathcal{G}$ that starts in u and ends in v. We write $u\overset{\mathcal{G}}{\leftrightarrow}v$ if u is reachable from v and v is reachable from u. If $u\overset{\mathcal{G}}{\leftrightarrow}v$ does not hold we write $u\overset{\mathcal{G}}{\not\leftrightarrow}v$. The relation $\overset{\mathcal{G}}{\leftrightarrow}$ is an equivalence relation on the vertices of $\mathcal{G}$ and the equivalence classes are called the *irreducible components* of $\mathcal{G}$. For any irreducible component of $\mathcal{G}$, we shall sometimes also refer to the subgraph of $\mathcal{G}$ that it induces as an irreducible component. A graph $\mathcal{G}$ is called *irreducible* if it has only one irreducible component, or *reducible* otherwise.

Theorem 5 *Let X be an m-step $\mathbb{Z}$-SFT over $\mathcal{A}$ with $h_{\mathrm{ind}}(X)=0$. Then for all d,*

$$h\big(X^{\otimes(d+1)}\big)\le\frac{m}{m+1}h\big(X^{\otimes d}\big).$$

In particular, $h_\infty(X)=0$.

Remark The proof is a generalization of the proof of [10, Theorem 2].

Proof For any such X, we construct a presentation $\mathcal{G}=((V,E),L)$ of X. Its vertices are all m-letter words over $\mathcal{A}$, $V=\mathcal{A}^m$, and the edges and their labels are given by

$$E=\big\{(w_0\cdots w_{m-1},w_1\cdots w_m):w_0\cdots w_m\in B(X)\big\},$$

where for any $e=(w_0\cdots w_{m-1},w_1\cdots w_m)\in E$, $\sigma(e)=w_0\cdots w_{m-1}$, $\tau(e)=w_1\cdots w_m$, and $\mathcal{L}(e)=w_m$. It's easy to verify that $\mathcal{G}$ is indeed a presentation of X. We will need the next two lemmas. The first shows that if $h_{\mathrm{ind}}(X)=0$, knowledge of long enough prefixes and suffixes of a word in $B(X)$ is often sufficient to determine the middle of the word. $\square$

Lemma 1 *Let X be an m-step $\mathbb{Z}$-SFT over $\mathcal{A}$ with $h_{\mathrm{ind}}(X)=0$, and let $\mathcal{G}$ be as above. Let $x,y\in\mathcal{A}^*$ be words of length m such that $x\overset{\mathcal{G}}{\leftrightarrow}y$. Then there is at most one word of the form xay, where $a\in\mathcal{A}$, in $B(X)$.*

Proof Assume to the contrary that there are two such words $xay, xby \in B(X)$ where $a, b \in \mathcal{A}$ and $a \neq b$. Therefore there are two paths—one generating xay, and the other generating xby—in $\mathcal{G}$. By the construction of $\mathcal{G}$, any path generating x must terminate at x and the same holds for y. It follows that there are two paths, both starting at x and terminating at y such that one generates ay and the other generates by. Since, by our assumption, x and y are in the same irreducible component, there is a path generating some word $z \in \mathcal{A}^*$ starting at y and ending at x. Consequently, for all $k \in \mathbb{N}$ and sequence $(c_i)_{i=1}^k \subseteq \{a, b\}$, there is a cycle in $\mathcal{G}$ generating the word $c_1 yzc_2 yz \cdots c_k yz$. Let $y = y_0 \cdots y_{m-1}$, and $z = z_0 \cdots z_{\ell-1}$, where $y_i, z_j \in \mathcal{A}$ for $i \in [m], j \in [\ell]$, and denote by $\hat{y}, \hat{z}$ the words over $\hat{\mathcal{A}}$ given by $\hat{y} = \{y_0\} \cdots \{y_{m-1}\}$ and $\hat{z} = \{z_0\} \cdots \{z_{\ell-1}\}$. It follows that for all $k \in \mathbb{N}$, $(\{a, b\}\hat{y}\hat{z})^k \in B(\hat{X})$. But then, $h_{\mathrm{ind}}(X) \geq (\ln 2)/(m + \ell + 1)$, contradicting the fact that $h_{\mathrm{ind}}(X) = 0$. $\square$

The next lemma bounds the number of appearances of words of the form xay in a certain word of $B(X)$, where $x, y \in \mathcal{A}^m$ belong to different irreducible components and $a \in \mathcal{A}$.

Lemma 2 *Let X be an m-step SFT, $\mathcal{G}$ be as above and $C_{\mathcal{G}}$ denote the number of irreducible components of $\mathcal{G}$. For $\ell \in \mathbb{N}$ and $z \in B_{[(\ell+1)m+\ell]}(X)$, let $y^{(0)}, \ldots, y^{(\ell)} \in \mathcal{A}^m$ and $a^{(0)}, \ldots, a^{(\ell-1)} \in \mathcal{A}$ be defined by*

$$z = y^{(0)} a^{(0)} y^{(1)} a^{(1)} \cdots y^{(\ell-1)} a^{(\ell-1)} y^{(\ell)}.$$

Then

$$\left|\left\{i \in [\ell] \,:\, y^{(i)} \overset{\mathcal{G}}{\not\leftrightarrow} y^{(i+1)}\right\}\right| < C_{\mathcal{G}}.$$

Proof Assume to the contrary that there are $0 \leq i_1 < i_2 < \cdots < i_{C_{\mathcal{G}}} < \ell$ such that $y^{(i_j)} \overset{\mathcal{G}}{\not\leftrightarrow} y^{(i_j+1)}$, for $j = 1, \ldots, C_{\mathcal{G}}$. Then the sequence $y^{(i_1)}, y^{(i_2)}, \ldots, y^{(i_{C_{\mathcal{G}}})}, y^{(i_{C_{\mathcal{G}}}+1)}$ has $C_{\mathcal{G}} + 1$ vertices and therefore contains two belonging to the same irreducible component, say $y^{(s)}$ and $y^{(t)}$, for some integers $0 \leq s < t \leq \ell$, with $s \in \{i_1, \ldots, i_{C_{\mathcal{G}}}\}$. Since the word $y^{(s)} a^{(s)} y^{(s+1)} \in B(X)$, it is easy to verify that $y^{(s+1)}$ is reachable from $y^{(s)}$. Similarly, since $y^{(s+1)} a^{(s+1)} y^{(s+2)} a^{(s+2)} \cdots a^{(t-1)} y^{(t)} \in B(X)$, it holds that $y^{(t)}$ is reachable from $y^{(s+1)}$. But since $y^{(t)}$ is in the same irreducible component as $y^{(s)}$, it follows that $y^{(s)}$ is reachable from $y^{(s+1)}$ as well. Thus $y^{(s)} \overset{\mathcal{G}}{\leftrightarrow} y^{(s+1)}$ which contradicts $s \in \{i_1, \ldots, i_{C_{\mathcal{G}}}\}$. $\square$

We can now prove Theorem 5. Let ℓ be a positive integer. Denote by $\mathbf{l} \in \mathbb{N}^d$ the d-tuple with every entry equal to ℓ, and let $\mathbf{m}_\ell \in \mathbb{N}^{d+1}$ be given by $\mathbf{m}_\ell = ((\ell + 1)m + \ell, \ell, \ell, \ldots, \ell)$. We will give an upper bound on $|B_{[\mathbf{m}_\ell]}(X^{\otimes(d+1)})|$. Let $S \subseteq [(\ell + 1)m + \ell]$ be the set $[(\ell + 1)m + \ell] \setminus \{i(m + 1) + m : i \in [\ell]\}$ and set $\Gamma = S \times [\mathbf{l}]$. For the purposes of this proof we abbreviate $B_\Gamma(X^{\otimes(d+1)})$ by B_Γ. For a configuration $x \in B_\Gamma$, define the set $A(x) \subseteq B_{[\mathbf{m}_\ell]}(X^{\otimes(d+1)})$ by $A(x) = \{z \in B_{[\mathbf{m}_\ell]}(X^{\otimes(d+1)}) : z|_\Gamma = x\}$. Clearly,

$$\bigcup_{x \in B_\Gamma} A(x) = B_{[\mathbf{m}_\ell]}\big(X^{\otimes(d+1)}\big). \tag{9}$$

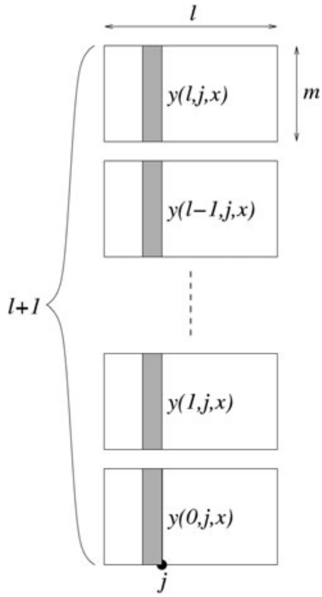
Let $x \in B_\Gamma$, and fix $\mathbf{j} \in [\mathbf{l}]$. For $i \in [\ell + 1]$, let $y(i, \mathbf{j}, x) \in \mathcal{A}^m$ be the word given by

$$y(i, \mathbf{j}, x) = x_{(i(m+1),\mathbf{j})} x_{(i(m+1)+1,\mathbf{j})} \cdots x_{(i(m+1)+m-1,\mathbf{j})}$$

and for $z \in A(x)$ let $w(z, \mathbf{j}) \in \mathcal{A}^{(\ell+1)m+\ell}$ be the word given by

$$w(z, \mathbf{j}) = z_{(0,\mathbf{j})} z_{(1,\mathbf{j})} \cdots z_{((\ell+1)m+\ell-1,\mathbf{j})}.$$

Fig. 1 $x \in B_\Gamma$



Note that for such z, since $z \in B_{[\mathbf{m}_\ell]}(X^{\otimes(d+1)})$, $w(z, \mathbf{j}) \in B(X)$. Since $z|_\Gamma = x$, we may write

$$w(z,\mathbf{j}) = y(0,\mathbf{j},x)a(0,\mathbf{j},z)y(1,\mathbf{j},x)a(1,\mathbf{j},z)\cdots y(\ell-1,\mathbf{j},x)a(\ell-1,\mathbf{j},z)y(\ell,\mathbf{j},x),$$

where $a(i,\mathbf{j},z) = z_{(i(m+1)+m,\mathbf{j})}$ for every $i \in [\ell]$ (see Fig. 1). Now, if $i \in [\ell]$ has the property that $y(i,\mathbf{j},x)\overset{\mathcal{G}}{\leftrightarrow}y(i+1,\mathbf{j},x)$ then, by Lemma 1, all $z \in A(x)$ have the same $a(i,\mathbf{j},z)$. On the other hand, by Lemma 2, since $A(x) \neq \emptyset$, we have $|\{i \in [\ell] : y(i,\mathbf{j},x)\overset{\mathcal{G}}{\not\leftrightarrow}y(i+1,\mathbf{j},x)\}| < C_\mathcal{G}$. It follows that $|\{w(z,\mathbf{j}) : z \in A(x)\}| \le |\mathcal{A}|^{C_\mathcal{G}}$, and consequently

$$\left|A(x)\right| \le \prod_{\mathbf{j}\in[\mathbf{l}]} \left|\left\{w(z,\mathbf{j}) : z \in A(x)\right\}\right| \le |\mathcal{A}|^{C_\mathcal{G}\ell^d}. \tag{10}$$

Since for any $x \in B_\Gamma$, and any $i \in S$, it holds that $x|_{\{i\}\times[\mathbf{l}]} \in B_{[\mathbf{l}]}(X^{\otimes d})$ (where we identify a configuration on $\{i\} \times [\mathbf{l}]$ with a configuration on $[\mathbf{l}]$ in the obvious manner), we get that $|B_\Gamma| \le |B_{[\mathbf{l}]}(X^{\otimes d})|^{|S|}$. Combining this with (9) and (10), we have

$$\left|B_{[\mathbf{m}_\ell]}\left(X^{\otimes(d+1)}\right)\right| = \sum_{x\in B_\Gamma} \left|A(x)\right| \le |\mathcal{A}|^{C_\mathcal{G}\ell^d} \left|B_{[\mathbf{l}]}\left(X^{\otimes d}\right)\right|^{(\ell+1)m}.$$

Taking the natural logarithm of both sides and dividing by $|[\mathbf{m}_\ell]|$, we obtain

$$\frac{\ln|B_{[\mathbf{m}_\ell]}(X^{\otimes(d+1)})|}{|[\mathbf{m}_\ell]|} \le \frac{(\ell+1)m}{(\ell+1)m+\ell}\,\frac{\ln|B_{[\mathbf{l}]}(X^{\otimes d})|}{\ell^d} + \frac{C_\mathcal{G}\ln|\mathcal{A}|}{(\ell+1)m+\ell}$$

The theorem follows by taking the limit as $\ell \to \infty$. □

6.3 The Golden Mean SFT

We recall that the *golden mean SFT* G is the $\mathbb{Z}$-SFT on $\{0, 1\}$ with forbidden list $\mathcal{F} = \{11\}$, and note that G is the RLL$(1, \infty)$ SFT as defined in Sect. 6.1. As was already mentioned, it is shown in [18] that $h_\infty(\mathsf{G}) = \frac{\ln 2}{2}$ so indeed $h_\infty = h_{\text{ind}}$. Korshunov and Sapozhenko [12] show that

$$\lim_{d\to\infty} \frac{|B_{[2]^d}(\mathsf{G}^{\otimes d})|}{2\sqrt{e}2^{2^{d-1}}} = 1,$$

which, since $\ln(|B_{[2]^d}(\mathsf{G}^{\otimes d}))/2^d \geq h(\mathsf{G}^{\otimes d})$, implies that for large enough d,

$$h\big(\mathsf{G}^{\otimes d}\big) \leq \frac{\ln 2}{2} + 2^{o(d)}2^{-d}$$

So $h(\mathsf{G}^{\otimes d})$ approaches $(\ln 2)/2$ at least exponentially fast. Our next theorem improves this bound, and also provides a lower bound. Together these show that the rate is indeed exponential. In what follows, we call two sites $x, y \in \mathbb{Z}^d$ *adjacent*, if $x - y \in \{\pm \mathbf{e}_i : i = 1, \dots, d\}$, where $\mathbf{e}_1, \dots, \mathbf{e}_d$ are the standard basis vectors for $\mathbb{R}^d$. In this section, we say that a configuration is locally admissible in G^d if it is locally admissible with respect to the forbidden list consisting of all configurations on pairs of adjacent sites in $\mathbb{Z}^d$ having '1's in both sites.

Theorem 6 *For sufficiently large d,*

$$\frac{\ln 2}{2} + \frac{\ln 2}{2}2^{-2d} \leq h\big(\mathsf{G}^{\otimes d}\big) \leq \frac{\ln 2}{2} + 2^{o(d)}2^{-2d}.$$

Proof We begin with the upper bound. The main result which we will use is from [8], and concerns the phase transition exhibited by $\mathsf{G}^{\otimes d}$ for large values of d. To state the result, we need a few notations. Call $v \in \mathbb{Z}^d$ *even* if the sum of the entries of v is even, and *odd* if the sum of the entries of v is odd. For $n \in \mathbb{N}$, let $\Lambda_n = ([-n+1, n] \cap \mathbb{Z})^d$, and $\partial^* \Lambda_n = \Lambda_{n+1} \setminus \Lambda_n$. Define the configuration $\delta_{\Lambda_n,e}$ on $\partial^* \Lambda_n$ which has '0's at even sites and '1's at odd sites, and the probability measure $\mu_{\Lambda_n,e}$ on $\{0, 1\}^{\Lambda_n}$ which is uniform over all configurations x on Λ_n for which the configuration $x \cup \delta_{\Lambda_n,e}$ on $\Lambda_n \cup \partial^* \Lambda_n$ is locally admissible in $\mathsf{G}^{\otimes d}$.

It is shown in [8, Theorem 1.2] that there exists $D \in \mathbb{N}$ so that $\mu_{\Lambda_n,e}(x(v) = 1) < \alpha(d)$, for any $n \in \mathbb{N}$, $d > D$, and even vertex $v \in \Lambda_n$, where $\alpha(d) = 2^{-(2-o(1))d}$. (In fact, their result involves a parameter λ; the result we have quoted corresponds to $\lambda = 1$.)

Roughly speaking, this result states that when d is very large, the vast majority of configurations on a large d-dimensional cube which have a "checkerboard" pattern on the boundary continue the checkerboard pattern on the interior with relatively few exceptions. We assume that D is so large that $\alpha(d) < \frac{1}{4}$ when $d > D$.

Now, we will use a counting argument to give an upper bound on the size of the set $S_{n,e}$ of configurations x on Λ_n for which $x \cup \delta_{R,e}$ is locally admissible in $\mathsf{G}^{\otimes d}$. Since $\mu_{\Lambda_n,e}(x(v) = 1) < \alpha(d)$ for all even $v \in \Lambda_n$,

$$\sum_{x\in S_{n,e}} \big(\#\text{ of '1's at even sites in } x\big) < \alpha(d)|S_{n,e}|\frac{(2n)^d}{2}.$$

It then follows that the number of configurations in $S_{n,e}$ with total number of '1's at even sites fewer than $\alpha(d)(2n)^d$ is at least $\frac{|S_{n,e}|}{2}$. Denote this set of configurations by $S'_{n,e}$. We note that for any way of filling the even sites of Λ_n with '0's and '1's, there are obviously at most

$2^{\frac{(2n)^d}{2}}$ ways to fill the leftover odd sites. By this observation and by indexing by the number k of '1's at even sites within configurations in $S'_{n,e}$, we see that

$$\left|S'_{n,e}\right| \le 2^{\frac{(2n)^d}{2}} \sum_{k=0}^{\lfloor \alpha(d)(2n)^d \rfloor} \binom{\frac{(2n)^d}{2}}{k} \le 2^{\frac{(2n)^d}{2}} \big(\alpha(d)(2n)^d\big) \binom{\frac{(2n)^d}{2}}{\lfloor \alpha(d)(2n)^d \rfloor}.$$

Therefore, by Stirling's formula, for $d > D$,

$$\limsup_{n\to\infty} \frac{\ln |S'_{n,e}|}{(2n-2)^d} \le \frac{\ln 2}{2} + \frac{-2\alpha(d)\ln(2\alpha(d)) - (1-2\alpha(d))\ln(1-2\alpha(d))}{2}, \tag{11}$$

which is less than $\frac{\ln 2}{2} + 2^{-(2-o(1))d}$ for sufficiently large d. We recall that $|S'_{n,e}| \ge \frac{|S_{n,e}|}{2}$, and also note that any configuration in $B_{\Lambda_{n-1}}(\mathsf{G}^{\otimes d})$ can be extended to a configuration in $S_{n,e}$ by surrounding it with '0's. Therefore, $|S_{n,e}| \ge B_{\Lambda_{n-1}}(\mathsf{G}^{\otimes d})$. It is then clear that if $S'_{n,e}$ were replaced by $B_{\Lambda_{n-1}}(\mathsf{G}^{\otimes d})$ in (11), it would still hold, and so $h(\mathsf{G}^{\otimes d}) \le \frac{\ln 2}{2} + 2^{-(2-o(1))d}$ for large enough d.

We now establish the lower bound on $h(\mathsf{G}^{\otimes d})$. For any n and d, define E_n to be the set of even sites in the rectangular prism Λ_n, and define the *neighbor set* of any $x \in \mathbb{Z}^d$ as $N_x := \{y \in \mathbb{Z}^d : y \text{ is adjacent to } x\}$. Define the pattern $p = 0^{N_0}$ (where we abbreviate and write 0^{N_0} for $\{0\}^{N_0}$) containing '0's at the $2d$ neighbors of the origin. (Obviously translations of p contain '0's at the neighbors of some other site in $\mathbb{Z}^d$.) Define the shape

$$C = \Lambda_{n-1} \cup \big(E_n \cap \partial^*(\Lambda_{n-1})\big).$$

Note that every configuration $w \in \{0,1\}^{E_n}$ can be extended to some set of configurations in $B_C(\mathsf{G}^{\otimes d})$ by filling in the odd sites in Λ_{n-1} with '0's and '1's. For each w, one can do this in 2^t ways, where t is the number of occurrences of p in w. This is true since the only odd sites which are not forced to be 0 given w are those which are surrounded by '0's (i.e. those which are at the center of an occurrence of p in w), and since no two odd sites are adjacent, each of these t odd sites may be independently filled with 0 or 1 given w.

Denote by $f(w)$ the number of occurrences of p in any $w \in \{0,1\}^{E_n}$. Then,

$$\big|B_C\big(\mathsf{G}^{\otimes d}\big)\big| = \sum_{w\in\{0,1\}^{E_n}} 2^{f(w)} \ge 2^{|E_n|} 2^{(2^{-|E_n|}\sum_{w\in\{0,1\}^{E_n}} f(w))}$$

by Jensen's inequality and convexity of 2^x. It is clear, however, that $2^{-|E_n|}\sum_{w\in\{0,1\}^{E_n}} f(w)$ is just the expected value of the function $f(w)$ with respect to the $(\frac{1}{2},\frac{1}{2})$ Bernoulli measure on $\{0,1\}^{E_n}$, which is $2^{-2d}|C \setminus E_n| = 2^{-2d}2^{\frac{(2n-2)^d}{2}}$. We have then shown that

$$\big|B_C\big(\mathsf{G}^{\otimes d}\big)\big| \ge 2^{\frac{(2n)^d}{2}} 2^{2^{-2d}\frac{(2n-2)^d}{2}}. \tag{12}$$

Since $\Lambda_{n-1} \subseteq C \subseteq \Lambda_n$, $\lim_{n\to\infty} \frac{\ln|B_C(\mathsf{G}^{\otimes d})|}{(2n)^d} = h(\mathsf{G}^{\otimes d})$. So, by taking logs of both sides of (12) and dividing by $(2n)^d$, we see that $h(\mathsf{G}^{\otimes d}) \ge \frac{\ln 2}{2} + \frac{\ln 2}{2}2^{-2d}$. □

Remark In fact, our reading of the proof in [8] suggests that the expression $2^{o(d)}$ in the statement of the theorem could be replaced by cd^k for some constants c and k.

6.4 The 3-Checkerboard

The 3-*checkerboard SFT* C is the $\mathbb{Z}$-SFT on $\{0, 1, 2\}$ with forbidden list $\mathcal{F} = \{00, 11, 22\}$. In this section, we will use the results of Sect. 6.3 to show that $h_\infty(\mathsf{C}) = h_{\text{ind}}(\mathsf{C}) = \frac{\ln 2}{2}$. This answers a question of [7]. In this section, we say that a configuration is locally admissible in C^d if it is locally admissible with respect to the forbidden list consisting of all configurations over $\{0, 1, 2\}$ on pairs of adjacent sites in $\mathbb{Z}^d$ having the same symbol.

Lemma 3 *Let $S \subseteq \mathbb{Z}^d$ be finite, let $x \in S$, and let $U_x = N_x \cap S$. Then for any partition $\{A_1, A_2\}$ of U_x,*

$$\left|B_{S\setminus\{x\}}\left(\mathsf{G}^{\otimes d}\right)\right| + \left|B_{S\setminus(\{x\}\cup(U_x))}\left(\mathsf{G}^{\otimes d}\right)\right| \geq \left|B_{S\setminus(\{x\}\cup A_1)}\left(\mathsf{G}^{\otimes d}\right)\right| + \left|B_{S\setminus(\{x\}\cup A_2)}\left(\mathsf{G}^{\otimes d}\right)\right|.$$

Proof Choose such S and x, and fix any partition $\{A_1, A_2\}$ of U_x. By the definition of $\mathsf{G}^{\otimes d}$, adding '0's to a globally admissible configuration always yields a globally admissible configuration. Therefore, we can define an injection $\phi_1 : B_{S\setminus(\{x\}\cup A_1)}(\mathsf{G}^{\otimes d}) \to B_{S\setminus\{x\}}(\mathsf{G}^{\otimes d})$, where $u \mapsto u \cup 0^{A_1}$. Similarly, define $\phi_2 : B_{S\setminus(\{x\}\cup A_2)}(\mathsf{G}^{\otimes d}) \to B_{S\setminus\{x\}}(\mathsf{G}^{\otimes d})$, where $u \mapsto u \cup 0^{A_2}$. Then it is clear that

$$\begin{aligned}\left|B_{S\setminus\{x\}}\left(\mathsf{G}^{\otimes d}\right)\right| \geq{}& \left|B_{S\setminus(\{x\}\cup A_1)}\left(\mathsf{G}^{\otimes d}\right)\right| + \left|B_{S\setminus(\{x\}\cup A_2)}\left(\mathsf{G}^{\otimes d}\right)\right| \\ &- \left|\phi_1\left(B_{S\setminus(\{x\}\cup A_1)}\left(\mathsf{G}^{\otimes d}\right)\right) \cap \phi_2\left(B_{S\setminus(\{x\}\cup A_2)}\left(\mathsf{G}^{\otimes d}\right)\right)\right|.\end{aligned}$$

We now note that for any $v \in \phi_1(B_{S\setminus(\{x\}\cup A_1)}(\mathsf{G}^{\otimes d})) \cap \phi_2(B_{S\setminus(\{x\}\cup A_2)}(\mathsf{G}^{\otimes d}))$, $v|_{A_1} = 0^{A_1}$ and $v|_{A_2} = 0^{A_2}$, and so $v|_{U_x} = 0^{U_x}$. But then clearly any such v is determined by its letters on $S \setminus (\{x\} \cup U_x)$, and so

$$\left|\phi_1\left(B_{S\setminus(\{x\}\cup A_1)}\left(\mathsf{G}^{\otimes d}\right)\right) \cap \phi_2\left(B_{S\setminus(\{x\}\cup A_2)}\left(\mathsf{G}^{\otimes d}\right)\right)\right| \leq \left|B_{S\setminus(\{x\}\cup U_x)}\left(\mathsf{G}^{\otimes d}\right)\right|,$$

completing the proof. □

Our next result gives an upper bound on the topological entropy of a $\mathbb{Z}^d$ SFT X in terms of a $(d-1)$-dimensional expression.

Lemma 4 *Let $d > 1$, X be a $\mathbb{Z}^d$ SFT over $\mathcal{A}$, defined by a forbidden list $\mathcal{F}$, and $\mathbf{m} \in \mathbb{N}^{d-1}$. For $w \in B_{[\mathbf{m}]\times\{0\}}(X)$ we define $S(w) := |\{v \in \mathcal{A}^{[\mathbf{m}]\times\{1\}} : v \cup w$ is locally admissible with respect to $\mathcal{F}\}|$. Then*

$$h(X) \leq \frac{1}{|[\mathbf{m}]|} \ln \max_{w \in B_{[\mathbf{m}]\times\{0\}}(X)} S(w).$$

Proof Let $\mathbf{m} = (m_1, \ldots, m_{d-1})$ and for $n \in \mathbb{N}$ set $\mathbf{n} = (m_1, \ldots, m_{d-1}, n)$. One can construct every configuration in $B_{[\mathbf{n}]}(X)$ by filling the values of the sites of each "layer", $[\mathbf{m}] \times \{i\}$, in turn, for $i = 0, 1, \ldots, n-1$, so that at each stage our configuration is globally admissible. There are $|B_{[\mathbf{m}]\times\{0\}}(X)|$ choices for the first such layer, and at most $\max_{w\in B_{[\mathbf{m}]\times\{0\}}(X)} S(w)$ for each successive layer given the previous one. It follows that

$$\begin{aligned}\left|B_{[\mathbf{n}]}(X)\right| &\leq \left|B_{[\mathbf{m}]\times\{0\}}(X)\right|\Big(\max_{w\in B_{[\mathbf{m}]\times\{0\}}(X)} S(w)\Big)^{n-1} \\ &\leq |\mathcal{A}|^{|[\mathbf{m}]|}\Big(\max_{w\in B_{[\mathbf{m}]\times\{0\}}(X)} S(w)\Big)^{n}.\end{aligned}$$

Fig. 2 Defining $\phi_{(5,5)}$

0/1	0/2	0/1	0/2	0/1
0/2	0/1	0/2	0/1	0/2
0/1	0/2	0/1	0/2	0/1
0/2	0/1	0/2	0/1	0/2
0/1	0/2	0/1	0/2	0/1

By taking the logarithm of both sides and dividing by $|[\mathbf{n}]|$, we get

$$h(X) \le \frac{\ln |B_{[\mathbf{n}]}(X)|}{|[\mathbf{n}]|} \le \frac{\ln |\mathcal{A}|}{n} + \frac{1}{|[\mathbf{m}]|} \ln \max_{w \in B_{[\mathbf{m}]\times\{0\}}(X)} S(w),$$

which upon letting $n \to \infty$, completes the proof. $\square$

We now state the main result of this subsection, which will allow us to use the fact that $h_\infty(\mathsf{G}) = h_{\text{ind}}(\mathsf{G})$ to show that $h_\infty(\mathsf{C}) = h_{\text{ind}}(\mathsf{C})$.

Theorem 7 *For any $d > 1$, $h(\mathsf{G}^{\otimes d}) \le h(\mathsf{C}^{\otimes d}) \le h(\mathsf{G}^{\otimes (d-1)})$.*

Proof We begin with proof of the lower bound on $h(\mathsf{C}^{\otimes d})$. For any d and for any d-tuple $\mathbf{m}$, we define an injection from $B_{[\mathbf{m}]}(\mathsf{G}^{\otimes d})$ to $B_{[\mathbf{m}]}(\mathsf{C}^{\otimes d})$. We again define $\mathbf{j} \in \mathbb{Z}^d$ to be even if the sum of its entries is even, and odd if the sum of its entries is odd. Then, for any $\mathbf{m}$ and any $w \in B_{[\mathbf{m}]}(\mathsf{G}^{\otimes d})$, define $\phi_{\mathbf{m}}(w)$ by

$$\big(\phi_{\mathbf{m}}(w)\big)(\mathbf{j}) = \begin{cases} 0 & \text{if } w(\mathbf{j}) = 1 \\ 1 & \text{if } w(\mathbf{j}) = 0 \text{ and } \mathbf{j} \text{ is even} \\ 2 & \text{if } w(\mathbf{j}) = 0 \text{ and } \mathbf{j} \text{ is odd} \end{cases}$$

It is clear that $\phi_{\mathbf{m}}(w)$ is locally admissible; since w did not contain two adjacent '1's, $\phi_{\mathbf{m}}(w)$ does not contain two adjacent '0's, and by definition $\phi_{\mathbf{m}}(w)$ does not contain two adjacent '1's or '2's, since adjacent sites have opposite parity. It is easy to see that in fact $\phi_{\mathbf{m}}(w)$ is globally admissible by applying a similar mapping to the configuration $w' \in \mathsf{G}^{\otimes d}$ in which w appears. Figure 2 illustrates this concept; any configuration w in $B_{[(5,5)]}(\mathsf{G}^{\otimes 2})$ yields a locally admissible configuration in $B_{[(5,5)]}(\mathsf{C}^{\otimes 2})$ by letting a 0 at a site in w correspond to the second choice at that site, and a 1 at a site in w correspond to the first choice at that site.

Therefore, $|B_{[\mathbf{m}]}(\mathsf{G}^{\otimes d})| \le |B_{[\mathbf{m}]}(\mathsf{C}^{\otimes d})|$, and by taking logarithms and normalizing we get the first inequality of Theorem 7.

The proof of the upper bound is more involved. The main idea is to use Lemma 4. We claim that for any $d > 1$, any $(d-1)$-tuple $\mathbf{m} \in \mathbb{N}^{d-1}$, and any $w \in B_{[\mathbf{m}]}(\mathsf{C}^{\otimes (d-1)})$, $S(w) \le |B_{[\mathbf{m}]}(\mathsf{G}^{\otimes (d-1)})|$. If we can prove this claim, then Lemma 4 implies $h(\mathsf{C}^{\otimes d}) \le h(\mathsf{G}^{\otimes (d-1)})$. We first think of $S(w)$ in a slightly different way.

For any w as defined above, we can define an associated diagram, which we call the *up-followers diagram* of w and denote by $U(w)$. The up-followers diagram $U(w)$ of any $w \in B_{[\mathbf{m}]}(\mathsf{C}^{\otimes d})$ is a member of $\{\{0,1\},\{0,2\},\{1,2\}\}^{[\mathbf{m}]}$, where for any $s \in [\mathbf{m}]$, $(U(w))(s)$

0	1	2	1	0
1	2	0	2	1
2	0	2	1	0
1	2	1	0	1
0	1	0	1	2

w

1/2	0/2	0/1	0/2	1/2
0/2	0/1	1/2	0/1	0/2
0/1	1/2	0/1	0/2	1/2
0/2	0/1	0/2	1/2	0/2
1/2	0/2	1/2	0/2	0/1

Fig. 3 A configuration w and up-followers diagram $U(w)$ induced by it

contains $\{0, 1, 2\} \setminus \{w(s)\}$. Then clearly, $S(w)$ is the number of ways to fill $U(w)$ (by choosing one of the two options at each site) with a locally admissible configuration in $\mathsf{C}^{\otimes(d-1)}$. As an example, in Fig. 3, $S(w)$ is equal to the number of ways to choose a letter at each site in the up-followers diagram $U(w)$ on the right to make a locally admissible pattern in $\mathsf{C}^{\otimes 2}$.

We are then attempting, for any w, to bound the maximum number of ways to fill $U(w)$ with locally admissible configurations in $\mathsf{C}^{\otimes(d-1)}$. It is easy to see that any way of assigning the pairs $\{0, 1\}$, $\{0, 2\}$, and $\{1, 2\}$ to sites such that adjacent sites receive distinct pairs is an up-followers diagram for some locally admissible w in $\mathsf{C}^{\otimes(d-1)}$. (This restriction comes from the fact that w itself must be a locally admissible pattern in $\mathsf{C}^{\otimes(d-1)}$, and so it does not contain the same letter in adjacent sites.)

We will now again recast this problem. Since the intersection of any two distinct 2-element subsets of $\{0, 1, 2\}$ is a singleton, for each adjacent pair of sites s, t in an up-followers diagram $U(w)$ of some configuration $w \in B_{[\mathbf{m}]}(\mathsf{C}^{\otimes(d-1)})$, exactly one of the four possible ways to choose letters at s and t is disallowed in $\mathsf{C}^{\otimes(d-1)}$. We use this fact to associate to any up-followers diagram $U = U(w)$ (on the set of sites $\mathbf{m} \subseteq \mathbb{Z}^d$) an object called a *forbidden adjacency chart* or FAC. The FAC associated to U consists of, for every set of 2 adjacent positions $\{s, t\} \subseteq [\mathbf{m}]$, a forbidden adjacency $w_{\{s,t\}} \in \{a, b\}^{\{s,t\}}$ defined by

$$w_{\{s,t\}}(s) = \begin{cases} a & \text{if } \min U(s) = U(s) \cap U(t) \\ b & \text{if } \max U(s) = U(s) \cap U(t) \end{cases} \quad \text{and}$$

$$w_{\{s,t\}}(t) = \begin{cases} a & \text{if } \min U(t) = U(s) \cap U(t) \\ b & \text{if } \max U(t) = U(s) \cap U(t). \end{cases}$$

As an example, Fig. 4 contains an up-followers diagram and its associated FAC. (The pair of symbols flanking each edge denotes the forbidden adjacency for the pair of sites sharing that edge.) It is straightforward to check that the number of ways to fill U with a locally admissible configuration in $\mathsf{C}^{\otimes(d-1)}$ is the same as the number of ways to fill the FAC associated to U by choosing either a or b at each site and avoiding all of the forbidden adjacencies.

It is now sufficient to show that for any FAC, the number of ways to fill it with letters a, b without using any of the forbidden adjacencies is at most the number of globally admissible configurations with the same shape in $\mathsf{G}^{\otimes(d-1)}$. We will prove this claim for FACs with any shape in $\mathbb{Z}^{d-1}$ (not just rectangular prisms), and will prove it by induction on the number of

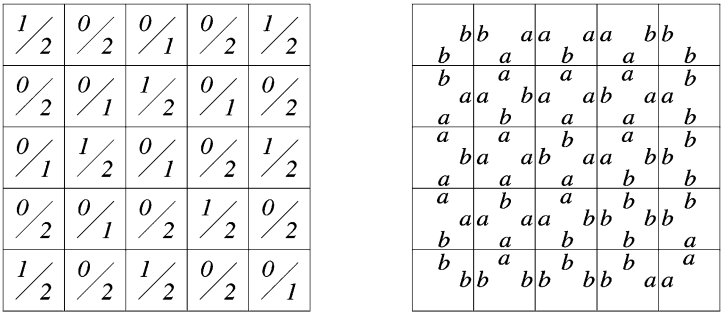


Fig. 4 An up-followers diagram U and the FAC induced by it

sites in the FAC. To be rigorous, we now say that a general FAC F consists of any finite set $S \subseteq \mathbb{Z}^{d-1}$ of sites and, for every pair $\{s,t\}$ of adjacent sites in S, some forbidden adjacency $w_{\{s,t\}} \in \{a,b\}^{\{s,t\}}$.

The case where F has 0 sites is trivial. Assume that the claim is true whenever $|S| < k$, and consider an arbitrary FAC F with $|S| = k$. Choose any $x \in S$. Then we may partition the legal ways of filling F by the letter at x. Due to the fact that each pair of adjacent vertices in S is associated with exactly one forbidden adjacency in F, we may define a partition $\{A, B\}$ of $N_x \cap S$, where A is the set of sites in $N_x \cap S$ whose letters are forced when an a is placed at x, and B is the set of sites in $N_x \cap S$ whose letters are forced when a b is placed at x. This means that the number of ways of filling F is less than or equal to the number of ways of filling $F|_{S\setminus(\{x\}\cup A)}$ plus the number of ways of filling $F|_{S\setminus(\{x\}\cup B)}$, where for any $T \subseteq S$, $F|_T$ is the FAC with set of sites T and adjacencies inherited from F.

However, by the inductive hypothesis, the number of ways of filling $F|_{S\setminus(\{x\}\cup A)}$ is less than or equal to $|B_{S\setminus(\{x\}\cup A)}(\mathsf{G}^{\otimes(d-1)})|$, and the number of ways of filling $F|_{S\setminus(\{x\}\cup B)}$ is less than or equal to $|B_{S\setminus(\{x\}\cup B)}(\mathsf{G}^{\otimes(d-1)})|$. Lemma 3 then implies that the number of ways of filling F is less than or equal to $|B_{S\setminus\{x\}}(\mathsf{G}^{\otimes(d-1)})| + |B_{S\setminus(\{x\}\cup(N_x\cap S))}(\mathsf{G}^{\otimes(d-1)})|$. However, by partitioning $B_S(\mathsf{G}^{\otimes(d-1)})$ by the letter which appears at x, it is fairly clear that $|B_S(\mathsf{G}^{\otimes(d-1)})| = |B_{S\setminus\{x\}}(\mathsf{G}^{\otimes(d-1)})| + |B_{S\setminus(\{x\}\cup(N_x\cap S))}(\mathsf{G}^{\otimes(d-1)})|$, completing the induction and the proof. □

The following is now obvious by Theorems 6 and 7.

Corollary 1 $h_\infty(\mathsf{C}) = h_{\mathrm{ind}}(\mathsf{C}) = \frac{\ln 2}{2}$, *and for sufficiently large* d, $\frac{\ln 2}{2} + \frac{\ln 2}{2}2^{-2d} \le h(\mathsf{C}^{\otimes d}) \le \frac{\ln 2}{2} + 2^{o(d)}2^{-2d}$.

References

1. Baxter, R.J.: Hard hexagons: exact solution. J. Phys. A, Math. Gen. **13**, L61–L70 (1980)
2. Berger, R.: The Undecidability of the Domino Problem, vol. 66. American Mathematical Society, Providence (1966)
3. Capobianco, S.: Multidimensional cellular automata and generalization of Fekete's lemma. Discrete Math. Theor. Comput. Sci. **10**, 95–104 (2008)
4. de Souza, J.C., Marcus, B.H., New, R., Wilson, B.A.: Constrained systems with unconstrained positions. IEEE Trans. Inf. Theory **48**, 866–879 (2002)
5. Even, S.: Graph Algorithms. Computer Science Press, New York (1979)
6. Friedland, S.: On the entropy of $\mathbb{Z}^d$ subshifts of finite type. Linear Algebra Appl. **252**, 199–220 (1997)

7. Friedland, S.: Multi-dimensional capacity, pressure and Hausdorff dimension. In: Mathematical Systems in Biology, Communications, Computations, and Finance, vol. 134, pp. 183–222 (2003)
8. Galvin, D., Kahn, J.: On phase transition in the hard-core model on $\mathbb{Z}^d$. Comb. Probab. Comput. **13**(2), 137–164 (2004)
9. Hochman, M., Meyerovitch, T.: A characterization of the entropies of multidimensional shifts of finite type. Ann. Math. **171**, 2011–2038 (2010)
10. Ito, H., Kato, A., Nagy, Zs., Zeger, K.: Zero capacity region of multidimensional run length constraints. Electron. J. Comb. **6**, R33 (1999)
11. Kastelyn, P.W.: The statistics of dimers on a lattice. Physica A **27**, 1209–1225 (1961)
12. Korshunov, A.D., Sapozhenko, A.A.: The number of binary codes with distance 2. Probl. Kibern. **40**, 111–130 (1983)
13. Lieb, E.H.: Residual entropy of square ice. Phys. Rev. **162**(1), 162–172 (1967)
14. Lind, D., Marcus, B.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge (1995). Reprinted 1999
15. Louidor, E., Marcus, B.H.: Improved lower bounds on capacities of symmetric 2-dimensional constraints using Rayleigh quotients. IEEE Trans. Inf. Theory, 1624–1639 (2010)
16. Louidor, E., Poo, T.L., Chaichanavong, P., Marcus, B.H.: Maximum insertion rate and capacity of multidimensional constraints. In: IEEE International Symposium on Information Theory, ISIT 2008. IEEE Press, New York (2008)
17. Moulin Ollagnier, J.: Ergodic Theory and Statistical Mechanics. Springer Lecture Notes in Mathematics (1985)
18. Ordentlich, E., Roth, R.: Independent sets in regular hypergraphs and multidimensional runlength-limited constraints. SIAM J. Discrete Math. **17**(4), 615–623 (2004)
19. Poo, T.L.: Optimal code rates for constrained systems with unconstrained positions: An approach to combining error correction codes with modulation codes for digital storage systems. Ph.D. thesis, Department of Electrical Engineering, Stanford (2005)
20. Poo, T.L., Chaichanavong, P., Marcus, B.H.: Trade-off functions for constrained systems with unconstrained positions. IEEE Trans. Inf. Theory **52**, 1425–1449 (2006)
21. Shwartz, M.: Constrained codes as networks of relations. IEEE Trans. Inf. Theory **54**, 2179–2195 (2008)
22. van Wijngaarden, A.J., Immink, K.A.S.: Maximum runlength-limited codes with error control capabilities. IEEE J. Sel. Areas Commun. **19**, 602–611 (2001)