ENTROPY ON REGULAR TREES

KARL PETERSEN \( ^\ast \)
Department of Mathematics
CB 3250 Phillips Hall
Chapel Hill, NC 27599, USA

IBRAHIM SALAMA
School of Business
North Carolina Central University
Durham, NC 27707, USA

(Communicated by Alejandro Maass)

Abstract. We show that the limit in our definition of tree shift topological entropy is actually the infimum, as is the case for both the topological and measure-theoretic entropies in the classical situation when the time parameter is \( \mathbb{Z} \). As a consequence, tree shift entropy becomes somewhat easier to work with. For example, the statement that the topological entropy of a tree shift defined by a one-dimensional subshift dominates the topological entropy of the latter can now be extended from shifts of finite type to arbitrary subshifts. Adapting to trees the strip method already used to approximate the hard square constant on \( \mathbb{Z}^2 \), we show that the entropy of the hard square tree shift on the regular \( k \)-tree increases with \( k \), in contrast to the case of \( \mathbb{Z}^k \). We prove that the strip entropy approximations increase strictly to the entropy of the golden mean tree shift for \( k = 2, \ldots, 8 \) and propose that this holds for all \( k \geq 2 \). We study the dynamics of the map of the simplex that advances the vector of symbol counts as the width of the approximating strip is increased, providing a fairly complete description for the golden mean subshift on the \( k \)-tree for all \( k \). This map provides an efficient numerical method for approximating the entropies of tree shifts defined by nearest neighbor restrictions. Finally, we show that counting configurations over certain other patterns besides the natural finite subtrees yields the same value of entropy for tree SFT's.

1. Introduction. Entropy is a single number attached to a topological or measure-theoretic dynamical system that in a limited but precise way describes the complexity or richness of the system. In recent years increasing attention has been paid to the calculation of entropy for systems in which the “time” is not \( \mathbb{Z} \) nor \( \mathbb{R} \), but perhaps \( \mathbb{Z}^d \) for some \( d \geq 2 \), or an arbitrary amenable group, or even a free or arbitrary countable group. We will not attempt to review the extensive and rapidly developing literature here (nor the connections with information theory, statistical mechanics, and other areas), referring only to [11, 7, 10] for background and references.

2010 Mathematics Subject Classification. 37B10, 37B40, 54H20.
Key words and phrases. Tree shift, complexity function, entropy.
\( ^\ast \) Corresponding author: Karl Petersen.
Aubrun and Béal [1, 2, 3, 4, 5] proposed studying subshifts on trees, since for such systems the “time” has both higher-dimensional and directional aspects, making them perhaps somehow between one- and higher-dimensional subshifts. Steve Piantadosi [19] studied the topological entropy of the hard square model on free groups \( \mathbb{F}_k \). He obtained an explicit formula in terms of a rapidly converging infinite series and used it to show numerically but rigorously that the entropy increases with \( k \) for a range of \( k \). Here we investigate some of these same questions for trees, with different methods but with some closely related results.

In a previous paper [18] we gave a definition of entropy for tree shifts and showed that the limit in the definition exists. We proved that for a 2-tree shift defined by nearest neighbor constraints, the tree-shift entropy dominates the entropy of the corresponding one-dimensional shift of finite type. We also provided estimates for the entropies of various 2-tree shifts, especially the ones determined by the “golden mean” (or “hard square” or “hard core”) condition that no two adjacent nodes have identical labels.

Here we prove (Theorem 2.1) that the limit in the definition is actually an infimum. As a corollary (Corollary 1) we show that the entropy comparison between a one-dimensional shift of finite type and the tree shift it defines holds for all subshifts. Then we adapt the “strip method” used for lattice shifts [16, 14] to study the entropy \( h^{(k)} \) of a subshift on the regular \( k \)-tree. We focus on the golden mean example, which is already sufficiently interesting and challenging; the general case is sketched in Section 8. Generalizing and improving the result in [18], we show in Theorem 3.7 that, for the golden mean, \( h^{(k)} \) is strictly increasing in \( k \). This contrasts with the apparent decrease of the entropy for the golden mean SFT’s on \( \mathbb{Z}^k \) for \( k = 1, 2, 3, 4 \) [9]. In Theorems 3.1 and 4.1 we show that for each fixed \( k = 2, \ldots, 8 \) the strip entropies \( h^{(k)}_n \) increase strictly to \( h^{(k)} \). We believe that the statement holds for all \( k \geq 2 \). As in [19], a related map of the interval (or, for the case of more general tree subshifts, simplex) appears as one considers ratios of symbol counts in the improving approximations. We produce a thorough analysis of this map for the case of golden mean restrictions (see Theorem 6.1) and show in Section 8 how to use it to obtain rapidly converging approximations to the entropies of more general tree shifts. Finally, we count configurations over extensions of the patterns that in this setting correspond to intervals in the one-dimensional case, showing in Corollary 3 that for tree SFT’s the resulting entropy is the same.

Apparently the definition of entropy considered here and in [19, 18] is not the same as sofic entropy (see [11, 7]), since the latter is a conjugacy invariant while the entropy considered here can increase under higher block codes.

1.1. Notation and setup. Let \( k \geq 2 \) and let \( \Sigma_k = \{0, 1, \ldots, k-1\} \). The set \( \Sigma^*_k \) of all finite words on the alphabet \( \Sigma_k \) is the \( k \)-tree, which is naturally visualized as the Cayley graph of the free semigroup on \( k \) generators. The empty word \( \epsilon \) corresponds to the root of the tree and the neutral element of the semigroup. Let \( A = \{0, 1, \ldots, d-1\} \) be an alphabet on \( d \) symbols. A labeled tree is a function \( \tau : \Sigma^*_k \to A \). For \( w \in \Sigma^*_k \), we think of \( \tau(w) \) as the label attached to the node determined by \( w \). For each \( n \geq 0 \) let \( \Delta_n = \cup_{0 \leq i \leq n} \Sigma^*_k \) denote the initial height-\( n \) subtree of the \( k \)-tree. The cardinality of \( \Delta_n \) is \( |\Delta_n| = 1 + k + \cdots + k^n \). An \( n \)-block is a function \( B : \Delta_n \to A \), which we think of as a labeling of \( \Delta_n \) or a configuration on \( \Delta_n \). We say that an \( n \)-block \( B \) appears in a labeled tree \( \tau \) if there is a node \( x \in \Sigma^*_k \) such that \( \tau(xw) = B(w) \) for all \( w \in \Delta_n \). A tree shift \( Z \) is the set of all labeled trees.
which omit all of a certain set (possibly infinite) of forbidden blocks. These are 
exactly the closed shift-invariant subsets of the full tree shift space \( T(A) = A^{\Sigma^n} \).
A tree shift \( Z \) is called \textit{transitive} if it contains a labeled tree \( \tau \) such that for every 
\( \xi \in Z \) every block that appears in \( \xi \) also appears in \( \tau \). Such a labeled tree \( \tau \) is called a \textit{transitive point}.

The complexity function \( p_\tau \) of the labeled tree \( \tau \) assigns to each \( n \geq 0 \) the number 
of \( n \)-blocks that appear in \( \tau \). The complexity function \( p_Z(n) \) of a tree shift \( Z \) gives for 
each \( n \geq 0 \) the number of \( n \)-blocks among all labeled trees in \( Z \). We are interested 
in studying the complexity functions of trees that are labeled according to certain 
restrictions, in particular nearest-neighbor constraints specified by \( d \)-dimensional 
0,1 transition matrices. In [18] it was proved that for any labeled tree \( \tau \) the limit
\[
\lim_{n \to \infty} \frac{\log p_\tau(n)}{1 + k + \cdots + k^n}
\]
exists. This limit is called the \textit{topological entropy} of the labeled tree \( \tau \). Similarly,
the topological entropy \( h(Z) \) of a tree shift \( Z \) is
\[
\lim_{n \to \infty} \frac{\log p_Z(n)}{1 + k + \cdots + k^n}.
\]

2. The limit in the definition of tree shift entropy is the infimum.

\textbf{Theorem 2.1.} The limit in the definition of tree shift topological entropy is actually
the infimum:
\[
\lim_{n \to \infty} \frac{\log p_\tau(n)}{1 + k + \cdots + k^n} = \inf \left\{ \frac{\log p_\tau(n)}{1 + k + \cdots + k^n} : n \in \mathbb{N} \right\},
\]
and similarly for \( h(Z) \).

\textit{Proof.} In the proof of the existence of the limit for \( h \) in [18], \( \Delta_{jm} \) was decomposed
into a union of shifts of \( \Delta_m \). But these copies of \( \Delta_m \) did not have independent 
entries; in fact they were not disjoint, since the last row of one formed the root 
vertices of the next ones. So here we improve the estimate 2.2 in [18] by making 
them disjoint.

Fix \( n \geq 1 \) and consider \( \Delta_n \). Its last row has \( k^n \) entries. The next row has \( k^{n+1} \)
entries, and we use these as root vertices for new shifts of \( \Delta_n \). This new row ends
with \( k^{n+1} k^n \) entries, so the next row has \( k^{2n+2} \) entries, which we use as vertices of 
new shifts of \( \Delta_n \). The last row now has \( k^{2n+2} \) entries, the next row has \( k^{3n+3} \)
and each of these becomes the root of a new shift of \( \Delta_n \).

We have formed a \( \Delta_{4(n+1)-1} \) out of \( 1 + k^{n+1} + k^{2(n+1)} + k^{3(n+1)} \) disjoint copies 
of \( \Delta_n \). In general, for each \( j \geq 1 \) we have
\[
\log p(j(n+1) - 1) \leq p(n)^{(j^{(n+1)} - 1)/(k^{n+1} - 1)}.
\]
(In formula 2.2 of [18] we find the same estimate, except with \( \leq p(n) \) replaced by
\( \leq p(n+1) \). So this estimate is better.)

Then take logarithms, divide by \( (k^{j(n+1)} - 1)/(k - 1) \), and take the limit as
\( j \to \infty \), to find that
\[
\frac{\log p(n)}{1 + k + k^2 + \cdots + k^n}.
\]

Therefore
\[
\frac{\log p(n)}{1 + k + k^2 + \cdots + k^n}.
\]
A first consequence of this result is a generalization of Theorem 3.3 of [18] for 2-trees from shifts of finite type to arbitrary subshifts. Using the same argument and Theorem 3.7 below, the statement extends to \( k \)-trees. Let \( M \) be a \( d \times d \) matrix with entries from \( \{0, 1\} \). The matrix \( M \) defines a one-step shift of finite type (SFT) \( X_M \) on the alphabet \( A \) and a tree shift \( Z_M \) consisting of all \( k \)-trees labeled by \( A \) with the property that for every \( w = w_0w_1 \ldots w_i \in \Sigma_k^* \), \( M_{w_iw_{i+1}} = 1 \) for all \( 0 \leq i < j \). In [18], Theorem 3.3, it was also proved that \( h(Z_M) = \sup \{ h(\tau) : \tau \in Z_M \} \) dominates the entropy \( h(X_M) \) of the associated shift of finite type.

More generally, given any subshift \( X \subset A^Z \), there is a naturally associated tree shift \( Z(X) \) defined as follows. Denote by \( L(X) \) the language of \( X \), namely the set of all finite words on \( A \) found as subwords of sequences in \( X \). The \( k \) shifts on \( \Sigma_k^* \) are defined by \( \sigma_s(w) = sw, w \in \Sigma_k^* \), \( s \in \Sigma_k \). For \( s_1, s_2, \ldots, s_m \in \Sigma_k \) and \( s = s_1s_2 \ldots s_m \in \Sigma_k^* \), define \( \sigma_s = \sigma_s_1 \ldots \sigma_s_m \). On a labeled tree \( \tau \), define \( (\sigma_s \tau)(w) = \tau(\sigma_s w) = \tau(sw), s \in \Sigma_k \).

We define the one-dimensional language of a labeled tree \( \tau \) to be the set \( L(1)(\tau) \) of words on the alphabet \( A \) found along paths in the tree:

\[
L(1)(\tau) = \epsilon \cup \{ \tau(s) : s \in \Sigma_k^* \} \cup \{ \tau(w)\tau(ws_1)\tau(ws_1s_2)\ldots \tau(ws_1s_2\ldots s_m) : m \in \mathbb{N}; w \in \Sigma_k^*; s_1, s_2, \ldots, s_m \in \Sigma_k \}.
\]

(7)

Given a subshift \( X \subset A^Z \), we define the tree shift associated to \( X \) to be

\[
Z(X) = \{ \tau \in A^{\Sigma_k^*} : L(1)(\tau) \subset L(X) \}.
\]

(8)

Thus a labeled tree is in \( Z(X) \) if and only if along all of its one-dimensional paths we see only words that are in the language of \( X \).

**Corollary 1.** Let \( X \subset A^Z \) be a subshift on a finite alphabet, let \( k = 2 \), and let \( Z(X) \) be the tree shift on the binary tree associated with \( X \). Then

\[
h_{\text{top}}(X) \leq h(Z(X)).
\]

(9)

**Proof.** Given any subshift \( X \), for each \( r \in \mathbb{N} \) let \( X_r \) be the shift of finite type which has the same \( r \)-blocks as \( X \). Then \( X \) is the decreasing intersection of the \( X_r \). Denote by \( p_X, p_X, \) and \( p_Z \) the complexity functions of the subshifts \( X \) and \( X_r \) and the tree shift \( Z \), respectively. Then for \( r \geq n \) we have \( p_X(n) = p_X(n) \), and similarly for \( p_Z(X) \) and \( p_Z(X) \). Thus

\[
h_{\text{top}}(X) \leq \lim_{r \to \infty} h_{\text{top}}(X_r) = \inf_r h_{\text{top}}(X_r)
\]

\[
= \inf_r \inf_n \frac{1}{n} \log p_X(n) = \inf_n \frac{1}{n} \log p_X(n)
\]

(10)

\[
= \inf \frac{1}{n} \log p_X(n) = h_{\text{top}}(X).
\]

(Cf. [8, Prop. 16.12 (i) and Cor.].)

In [18], Theorem 3.3, it was proved that if \( X \) is a one-step SFT, then

\[
h_{\text{top}}(X) \leq h(Z(X)).
\]

(11)

Each \( X_r \) is an \( (r - 1) \)-step SFT and is topologically conjugate to a one-step SFT \( Y_r \) on the alphabet \( A^{(r)}(X) \) of \( r \)-blocks which appear in \( X \). In a labeled tree in \( Z(Y_r) \), we think of the last entry of the \( r \)-block labeling a node as being attached to that node. A labeling by elements of \( A \) of the vertices of the \( k \)-tree is consistent with the restrictions from \( X_r \) if and only if it is consistent with the restrictions from \( Y_r \). Given a labeling of \( \Delta_n \) by \( A \), there may be several corresponding labelings by...
A^{(r)}(X)$, but they all agree after the first $r$ rows. Thus there is a constant $c$ (for example $p_{Y,1}(1)|_{\Delta_n}$) such that $p_{Z(Y,1)} \leq c p_{Z(X,1)}$. Using (10), Theorem 3.3 of [18] applied to $Y_r$ and $Z(Y_r)$, and Theorem 2.1, we then have

$$h_{top}(X) = \inf_r h_{top}(X_r) = \inf_r h_{top}(Y_r) \leq \inf_r h(Z(Y_r)) = \inf_r h(Z(X_r))$$

$$= \inf_r \frac{1}{n} \log p_{Z(X_r)}(n) = \inf_r \frac{1}{n} \log p_{Z(X_r)}(n)$$

$$= \inf_r \frac{1}{n} \log p_{Z(X_r)}(n) = h(Z(X)).$$

Remark 1. A similar statement applies to the entropy of an invariant measure $\mu$ on a tree shift $Z$ and can be proved by a similar argument. Indeed, existence of the limit for random fields was proved already by Berger and Ye [6, 20] (thanks to Tim Austin for the reference). Denote by $\alpha_n$ the partition of $Z$ according to all possible labelings of the nodes of $\Delta_n$ by elements of the alphabet $A$. Then

$$h_{\mu}(Z) = \lim_{n \to \infty} \frac{1}{|\Delta_n|} H_{\mu}(\alpha_n) = \lim_{n \to \infty} \frac{1}{|\Delta_n|} \left[ - \sum_{\Lambda \in \alpha_n} \mu(\Lambda) \log \mu(\Lambda) \right]$$

$$= \inf_n \frac{1}{|\Delta_n|} H_{\mu}(\alpha_n).$$

3. Strict increase with dimension. We use the strip method for the golden mean shift of finite type on the $k$-tree, $k \geq 2$, to show that the entropy $h_{top}(Z)$ is strictly increasing in $k$.

For each $n = 0, 1, 2, \ldots$ we define a 1-dimensional SFT $\Sigma^{(k)}_n$ whose alphabet consists of legal (no adjacent 1’s) labelings of the subtree with $k^n$ nodes consisting of a vertex with $\Delta^{(k)}_{n-1}$ attached. $\Delta^{(k)}_{n-1} = \emptyset$. $\Sigma^{(k)}_n$ consists of allowed labelings of the left edge 0$^*$ (see Section 1.1) of the $k$-tree and is conjugate to the ordinary one-dimensional golden mean SFT. $\Sigma^{(k)}_1$ is the one-dimensional SFT formed by the strip of adjacent pairs of vertices 0, 01 down the left edge. For $k = 2$, thinking of the possible labels on vertices as 0 and 1, we may represent it as the set of all one-sided sequences on an alphabet \{a = 00, b = 01, c = 10\}, with the restriction that the block cc does not occur. The alphabet for $\Sigma^{(2)}_2$ consists of the allowed labelings of a 4-vertex tree as in Figure 1, etc.

Denote by $h(\Sigma^{(k)}_n)$ the topological entropy of the one-dimensional SFT $\Sigma^{(k)}_n$. Recall that the alphabet of $\Sigma^{(k)}_n$ consists of allowed labelings of $k^n$ sites. We will show that the site specific entropies $h_{top}(k) = h(\Sigma^{(k)}_n)/k^n$ of the strip SFT's $\Sigma^{(k)}_n$ have limit $h_{top}(k)$=the entropy of the golden mean labeled $k$-tree. Moreover, with the help of computer algebra, we can prove that for each $k = 2, \ldots, 8$ the $h_{top}(k)$ are strictly increasing, and we believe that this holds for all $k \geq 2$.

For each $n = 0, 1, \ldots$, denote by $B^{(k)}_n$ the number of different labelings of $\Delta^{(k)}_n$ by the alphabet \{0, 1\}, by $B^{(k)}_0(0)$ the number of such labelings that have 0 at the root, and by $r^{(k)}_n = r^{(k)}(0)$ the ratio $B^{(k)}_n(0)/B^{(k)}_n$.

While $k$ is fixed, we will suppress the exponents ($k$).

The labeling counts satisfy the recursions

$$B_0(0) = B_0(1) = 1, B_0 = 2, r_0 = 1/2,$n+1(0) = [B_n(0) + B_n(1)] k = B^{(k)}_n, B_{n+1}(1) = (B_n(0))^k. \tag{14}$$
We define $B_{-1}(0) = 1, B_{-1}(1) = 0$, so that $B_{-1} = 1$. The labels at the nodes $i$ on the left edge respect the SFT restriction no 11. Let $a_i(0)$ denote the number of ways to label the first $i$ levels of the strip of width $n \geq 0$ with 0 labeling node $i$, and define $a_i(1)$ similarly. For $n \geq 1$ the $\Delta_{n-1}$ attached to node $i+1$ can be labeled in $B_{n-1}(0)$ ways if node $i+1$ has label 0, and in $B_{n-1}(0)$ ways if node $i+1$ has label 1. Therefore

$$
\begin{bmatrix}
a_{i+1}(0) \\
 a_{i+1}(1)
\end{bmatrix} =
\begin{bmatrix}
 B_{n-1}^{k-1} & B_{n-1}^{k-1} \\
 B_{n-1}(0)^{k-1} & 0
\end{bmatrix}
\begin{bmatrix}
a_i(0) \\
 a_i(1)
\end{bmatrix}.
$$

(15)

Thus for $n \geq 0$ the number $a_i(0) + a_i(1)$ of $i$-blocks in $\Sigma_{n}^{(k)}$ has the same asymptotic growth rate as the entries of $i$'th powers of the matrix

$$
\begin{bmatrix}
 B_{n-1}^{k-1} & B_{n-1}^{k-1} \\
 B_{n-1}(0)^{k-1} & 0
\end{bmatrix}.
$$

(16)

The latter is given by powers of the Perron-Frobenius eigenvalue $\lambda_{n-1}$, which is the maximum root of the characteristic equation

$$
f_{n-1}(\lambda) = \lambda^2 - B_{n-1}^{k-1} \lambda - [B_{n-1}B_{n-1}(0)]^{k-1},
$$

namely

$$
\lambda_{n-1} = \frac{B_{n-1}^{k-1} + \sqrt{B_{n-1}^{2k-2} + 4[B_{n-1}B_{n-1}(0)]^{k-1}}}{2} = B_{n-1}^{k-1}c_{n-1},
$$

(18)

with

$$
c_n = (1 + \sqrt{1 + 4r_n(0)^{k-1}})/2.
$$

(19)

Thus $c_n = \lambda_n/B_n^{k-1} = c_n(r)$ is the largest root of $x^2 - x - r_n^{k-1}$ and satisfies $c_n^2 - c_n = r_n^{k-1}$.

We have

$$
h_n^{(k)} = \log \frac{\lambda_{n-1}}{k^n}.
$$

(20)

Note that all $h_n^{(k)} \leq \log k$ because $B_n \leq k^n$.

**Theorem 3.1.** For each fixed $k = 2, 3, \ldots$ the site specific strip approximation entropies $h_n^{(k)}$ converge to the entropy $h^{(k)}$ of the golden mean SFT on the $k$-tree.
Proof. Since \( c_n \) is bounded and not less than 1,

\[
\lim h_n^{(k)} = \lim \frac{\log \lambda_{n-1}}{k^n} = \lim \frac{\log B_{n-1}^{k-1}}{k^n} + \lim \frac{\log c_{n-1}}{k^n} \\
= \lim \frac{(k-1) \log B_{n-1}}{k^n-1} = h^{(k)}.
\] (21)

We develop for golden mean tree shifts the analogue of Piantadosi’s [19] infinite series formula for the entropy of the golden mean SFT on a free group. Then we use the formula to prove that the entropy \( h^{(k)} \) of the golden mean SFT on the \( k \)-tree is strictly increasing in \( k \), for all \( k \geq 2 \). (Piantadosi used his infinite series formula and rigorous numerical estimates to prove the strict increase for a range of \( k \).) Our infinite series formula follows from two easy lemmas.

**Lemma 3.2.**

\[
h_n^{(k)} = h_n^{(k)} + \frac{1}{k^{n+1}} \log(1 + r_n^{k})^{k-1} + \frac{1}{k^{n+1}} \log(c_n) - \frac{1}{k^n} \log(c_{n-1}).
\] (22)

Proof.

\[
h_n^{(k)} = \frac{1}{k^{n+1}} \log(\lambda_n) = \frac{1}{k^{n+1}} \log(B_{n-1}^{k-1} c_n) \\
= \frac{1}{k^{n+1}} \log[B_{n-1}^{k(k-1)} (1 + r_n^k)^{k-1} c_n] \\
= \frac{1}{k^{n+1}} \log[B_{n-1}^{k(k-1)} c_n^{k-1} (1 + r_n^{k-1}) c_{n-1}^{k-1}] \\
= \frac{1}{k^n} \log(B_{n-1}^{k-1} c_{n-1}) + \frac{1}{k^{n+1}} \log(1 + r_n^{k})^{k-1} + \frac{1}{k^{n+1}} \log(c_n) - \frac{1}{k^n} \log(c_{n-1}) \\
= h_n^{(k)} + \frac{1}{k^{n+1}} \log(1 + r_n^{k})^{k-1} + \frac{1}{k^{n+1}} \log(c_n) - \frac{1}{k^n} \log(c_{n-1}).
\] (23)

Note

\[
h_1^{(k)} = \frac{1}{k} \log(B_0^{k-1} c_0) \\
= \frac{1}{k} \log(2^{k-1}) + \frac{1}{k} \log(c_0) \\
= \frac{k-1}{k} \log 2 + \frac{1}{k} \log(c_0)
\] (24)

and

\[
h_2^{(k)} = h_1^{(k)} + \frac{1}{k^2} \log(1 + r_0^k)^{k-1} + \frac{1}{k^2} \log(c_1) - \frac{1}{k} \log(c_0) \\
= \frac{k-1}{k} \log 2 + \frac{1}{k^2} \log(c_0) + \frac{1}{k^2} \log(1 + r_0^k)^{k-1} + \frac{1}{k^2} \log(c_1) - \frac{1}{k} \log(c_0) \\
= \frac{k-1}{k} \log 2 + \frac{1}{k^2} \log(1 + r_0^k)^{k-1} + \frac{1}{k^2} \log(c_1).
\] (25)

**Lemma 3.3.**

\[
h_n^{(k)} = \frac{k-1}{k} \log 2 + \sum_{i=1}^{n-1} \frac{1}{k^{i+1}} \log(1 + r_i^{k})^{k-1} + \frac{1}{k^n} \log(c_{n-1}).
\] (26)
Proof. Assume the formula holds for \( n - 1 \). Then
\[
h_{n-1}^{(k)} = \frac{k-1}{k} \log 2 + \sum_{i=1}^{n-2} \frac{1}{k^{i+1}} \log (1 + r_{i-1}^k)^{k-1} + \frac{1}{k^{n-1}} \log (c_{n-2}),
\]
and
\[
h_n^{(k)} = h_{n-1}^{(k)} + \frac{1}{k^n} \log (1 + r_{n-2}^k)^{k-1} + \frac{1}{k^n} \log (c_{n-1}) - \frac{1}{k^{n-1}} \log (c_{n-2})
\]
\[
= \frac{k-1}{k} \log 2 + \sum_{i=1}^{n-1} \frac{1}{k^{i+1}} \log (1 + r_{i-1}^k)^{k-1} + \frac{1}{k^n} \log (c_{n-1}).
\]

\[\square\]

**Theorem 3.4.** For all \( k \geq 2 \) the entropy of the golden mean SFT on the \( k \)-tree is given by the following formula:
\[
h^{(k)} = \frac{k-1}{k} \log 2 + (k-1) \sum_{i=1}^{\infty} \frac{1}{k^{i+1}} \log (1 + r_i^k).
\]

Examination of this infinite series formula provides useful upper and lower bounds for \( h^{(k)} \).

**Lemma 3.5.** For each \( k \geq 2 \), let
\[
L(k) = \frac{k-1}{k} \log 2 + \frac{k-1}{k^2} \log (1 + r_0^k) + \frac{k-1}{k^3} \log (1 + r_1^k)
\]
and
\[
U(k) = \frac{k-1}{k} \log 2 + \frac{k-1}{k^2} \log (1 + r_0^k) + \frac{1}{k^2} \log (1 + r_1^k).
\]

Then
\[
L(k) < h^{(k)} < U(k).
\]

Proof. The first inequality in (32) is immediate by truncation of the series. To prove the second, recall that \( r_0 = 1/2, r_1 = 1/(1 + 1/2^k) \), and note that \( r_i \leq r_1 \) for all \( i \geq 0 \) (see the proof of Theorem 6.1).

These formulas give us
\[
U(k) - L(k) = \frac{1}{k^3} \log (1 + r_1^k).
\]

We also need the estimate \( U(k-1) < L(k) \).

**Lemma 3.6.** For each \( k \geq 2 \),
\[
L(k) = \frac{k-1}{k} \log 2 + \frac{k-1}{k^3} \log[1 + (1 + r_0^k)^k]
\]
and
\[
U(k) = \frac{k-1}{k} \log 2 + \frac{1}{k^2} \log[1 + (1 + r_0^k)^k] - \frac{1}{k^2} \log(1 + r_0^k).
\]

Proof.
\[
L(k) = \frac{k-1}{k} \log 2 + \frac{k-1}{k^2} \log(1 + r_0^k) + \frac{k-1}{k^3} \log(1 + r_1^k)
\]
\[
= \frac{k-1}{k} \log 2 + \frac{k-1}{k^2} \log(1 + r_0^k) + \frac{k-1}{k^3} \log[1 + \frac{1}{(1 + r_0^k)^k}]
\]

We also need the estimate \( U(k-1) < L(k) \).
Recall that the inequality is verified in Corollary 2 by direct computation for \(2 \leq k \leq 5\). Proof. The inequality is satisfied for \(k = 1/2\). It is enough to show that for all \(k > 5\) we have \(L(k) > U(k-1)\), i.e.

\[
\frac{k - 1}{k} \log 2 + \frac{k - 1}{k^2} \log(1 + r_k^1) + \frac{k - 1}{k^3} \log[1 + (1 + x_k^k)^k] = \frac{k - 1}{k} \log 2 + \frac{k - 1}{k^3} \log[1 + (1 + r_k^0)^k].
\]

(36)

\[
U(k) = \frac{k - 1}{k} \log 2 + \frac{k - 1}{k^2} \log(1 + r_k^0) + \frac{1}{k} \log(1 + r_k^1)
\]

\[
= \frac{k - 1}{k} \log 2 + \frac{k - 1}{k^2} \log(1 + r_k^0) + \frac{1}{k^2} \log \left[ \frac{1 + (1 + r_k^0)^k}{(1 + r_k^0)^k} \right]
\]

\[
= \frac{k - 1}{k} \log 2 + \frac{k - 1}{k^2} \log(1 + r_k^0) + \frac{1}{k^2} \log((1 + (1 + r_k^0)^k) - 1) \log(1 + r_k^0)^k
\]

\[
= \frac{k - 1}{k} \log 2 + \frac{1}{k^2} \log[1 + (1 + r_k^0)^k] - \frac{1}{k^2} \log(1 + r_k^0)
\]

(37)

\[
\square
\]

**Theorem 3.7.** The entropy \(h^{(k)}\) of the golden mean SFT on the \(k\)-tree is strictly increasing in \(k\).

**Proof.** The inequality is verified in Corollary 2 by direct computation for \(2 \leq k \leq 5\). Recall that \(r_0 = 1/2\). It is enough to show that for all \(k > 5\) we have \(L(k) > U(k-1)\), i.e.

\[
\frac{k - 1}{k} \log 2 + \frac{k - 1}{k^2} \log[1 + (1 + \frac{1}{2k})^k] > \frac{k - 2}{k - 1} \log 2 + \frac{1}{(k - 1)^2} \log[1 + (1 + \frac{1}{2k-1})^{k-1}] - \frac{1}{(k - 1)^2} \log(1 + \frac{1}{2k-1}).
\]

(38)

Letting \(x_k = 1 + 1/2^k\), exponentiating, simplifying, and taking everything to the power \((k - 1)^2\), (38) is equivalent to

\[
2^{\frac{k-1}{k} + (\frac{k-1}{k})^2} > 1 + x_{k-1}^{k-1},
\]

(39)

We claim that

\[
2^{\frac{k-1}{k} + (\frac{k-1}{k})^2} > 1 + x_{k-1}^{k-1}
\]

for \(k \geq 6\),

(40)

from which (39) will follow, because then

\[
2^{\frac{k-1}{k} + (\frac{k-1}{k})^2} > 2^{\frac{k-1}{k} + (\frac{k-1}{k})^2} > 1 + x_{k-1}^{k-1} > \frac{1 + x_{k-1}^{k-1}}{x_{k-1}}.
\]

(41)

Note that the left side of (40) increases (to 4) in \(k\), while \(1 + x_{k-1}^{k-1}\) decreases (to 2), because, using the alternating series for \(\log(1 + t)\),

\[
\frac{k}{2k} - \frac{1}{2} \frac{k}{2k < \log(1 + \frac{1}{2k})^k < \frac{k}{2k}},
\]

(42)

so that \(x_k^k > x_{k+1}^{k+1}\) follows from

\[
\frac{k + 1}{2^{k+1}} < k \left( \frac{1}{2^k} - \frac{1}{2} \frac{1}{2^{2k}} \right),
\]

(43)

equivalently

\[
k > 1 + \frac{k}{2^k},
\]

(44)

which is true for \(k \geq 2\), because \(k/2^k\) is decreasing in \(k\) and \(2 > 1 + 2/2^2\).
Now (40) holds for \( k = 6 \), because then the left side is 2.6611... while the right side is 2.16633..., and therefore it holds for all \( k \geq 6 \).

4. Monotonicity of the strip approximation entropies for \( k = 2, \ldots, 8 \). Now we turn to the study for fixed \( k \) of the strip approximation entropies \( h_n^{(k)} \) for the entropy of the golden mean \( k \)-tree shift (with labeling alphabet \{0,1\}). Numerical calculations indicate that these approximations increase strictly with \( n \). We can prove this rigorously with the help of computer algebra for \( k = 2, \ldots, 8 \), and we believe that it is so for all \( k \geq 2 \). The strict increase allows for another proof of Theorem 3.7, for rigorous lower bounds for \( h^{(k)} \), and efficient approximations to \( h^{(k)} \), in the range of \( k \) for which strict increase is known to hold. Rigorous upper bounds are easy because the limit in the definition of entropy is the infimum.

**Theorem 4.1.** For \( k = 2, \ldots, 8 \) and \( n = 0, 1, \ldots \), we have
\[
\lambda_n^k < \lambda_{n+1}^k,
\]
and hence
\[
h_n^{(k)} < h_{n+1}^{(k)}.
\]
Moreover
\[
\lim_{k \to \infty} h^{(k)} = \log 2.
\]

**Remark 2.** The analogue of the final statement (47) was conjectured for the golden mean on free groups by Piantadosi [19] and proved already (in more generality, for \( k \)-tree shifts determined by arbitrary irreducible one-dimensional shifts of finite type) by the present authors in a previous paper [18]. Here it follows from Theorem 3.4 as well as (30). For subshifts on the lattices \( \mathbb{Z}^d \), this “asymptotic entropy” was known to be the nonincreasing limit in \( d \) and was proved in [12, 15] to equal the “independence entropy”.

**Proof.** Fix \( k = 2, 3, \ldots \). Note that for \( n \geq 0 \)
\[
r_{n+1} = \frac{1}{1 + r_n^k},
\]
so we define \( T = T_k : [0, 1] \to [0, 1] \) by
\[
Tx = \frac{1}{1 + x^k}, \quad 0 \leq x \leq 1.
\]
Then \( c_{n+1}(r) = c_n(Tr) \). Define also
\[
c = c(r) = \frac{1}{2} \left[ 1 + \sqrt{1 + 4r^{k-1}} \right],
\]
\[
v = v(r) = 1 + r^k.
\]
Note that
\[
c^2 - c = r^{k-1}.
\]
Abbreviate \( r_n = r, c_n(r) = c(r) \), and use \((Tr)^{k-1} = c(Tr)^2 - c(Tr)\). The statements
\[
c(Tr) - (Tr)^{k-1} c(r)^k > 0,
\]
\[
n_k(r) = c^{2k} - (1 + r^k)^{k-1} (c^k + 1) < 0,
\]
\[
c^{2k} - [(c - 1)(c^{k-1} + 1)]^{k/(k-1)} + 1^{k-1}(c^k + 1) < 0,
\]
\[
\left( \frac{c + r^{k-1}}{v} \right)^k < \frac{c^k + 1}{v}.
\]
are equivalent to one another (see Proposition 1). Moreover, when \( r = r_n, n \geq -1 \),
each is equivalent to

\[
\lambda^k_n < \lambda_{n+1} \tag{56}
\]

and implies

\[
\log \lambda_n < \frac{\log \lambda_{n+1}}{k}, \quad \text{so } h_{n+1}^{(k)} = \frac{\log \lambda_n}{k^n} < \frac{\log \lambda_{n+1}}{k^{n+1}} = h_{n+2}^{(k)}. \tag{57}
\]

We can prove by hand in the cases \( k = 2, 3 \) that \( n_k(r) \leq 0 \) for all \( r \in [0, 1] \), with
equality only at the one point where \( p_k(r) = r^{k+1} + r - 1 = 0 \). For \( k = 4, 5, 6, 7, 8 \),
to prove this we need the help of Sage and Mathematica. We believe that given time
and sufficient computer memory, (53) can be verified similarly for each individual
\( k \). So far we have not been able to prove algebraically that (76) has at most one
solution \( r \in [0, 1] \) (which would suffice to prove that the statement holds for all
\( k \)—see below).

The scheme of the calculation for each fixed \( k \) is as follows. (Details for \( k = 2, 3, 4 \)
are given below). Expand \( n_k(r) \) as a function of \( r \), using \( c(r) = (1 + \sqrt{1 + 4r^{k-1}}) / 2 \).
Group into \( B_k(r) \) all the terms that contain a factor of \( \sqrt{1 + 4r^{k-1}} \) to write \( n_k(r) = A_k(r) + B_k(r) \). We verify that \( A_k(r) \) and \( B_k(r) \) have opposite signs on \( [0, 1] \), take
\( B_k(r) \) to the right side of the formula, square both sides and take the difference.
Thus the left side of (53) is nonpositive if and only if \( d_k(r) = A_k(r)^2 - B_k(r)^2 \leq 0 \).
We define

\[
q_k(r) = -\frac{d_k(r)}{p_k(r)^2} \tag{58}
\]

and observe that \( q_k(r) > 0 \) on \( [0, 1] \), concluding the argument. (In some of the cases
a constant is factored out, making no difference in the sign.)

**Remark 3.** It appears that \( q_k(r) \to 1 + 2r + 3r^2 + \cdots = 1/(1 - r)^2 \) as \( k \to \infty \) for
each \( r \in [0, 1] \).

Figure 2 is a graph of \( n_6(r) \), and Figure 3 is a graph of \( q_5(r) \).

Here are the actual calculations for \( k = 2, 3, 4 \); the lengthy formulas in the cases
\( k = 5, 6, 7, 8 \) are omitted here and can be found online [17]. Maybe contemplation
of these formulas can help to find a proof for all \( k \).

\[
\begin{align*}
n_2(r) &= -1 + r - \frac{r^2}{2} - r^3 + r\sqrt{1 + 4r} - \frac{1}{2}r^2\sqrt{1 + 4r} \\
A_2(r) &= -1 + r - \frac{r^2}{2} - r^3 \\
B_2(r) &= (r - r^2/2)\sqrt{1 + 4r} \\
d_2(r) &= 1 - 2r + r^2 - 2r^3 + 2r^4 + r^6 \\
q_2(r) &= 1
\end{align*}
\]  

(59)

\[
\begin{align*}
n_3(r) &= \frac{1}{2}(-2 + 3r^2 - 6r^3 + 9r^4 - 6r^5 - r^6 - 3r^8 + 3r^2\sqrt{1 + 4r} - \\
&\quad 2r^3\sqrt{1 + 4r^2} + 3r^4\sqrt{1 + 4r^2} - 2r^5\sqrt{1 + 4r^2} - r^6\sqrt{1 + 4r^2} - r^8\sqrt{1 + 4r^2}) \\
A_3(r) &= -2 + 3r^2 - 6r^3 + 9r^4 - 6r^5 - r^6 - 3r^8 \\
B_3(r) &= \sqrt{1 + 4r^2}(3r^2 - 2r^3 + 3r^4 - 2r^5 - r^6 - r^8) \\
d_3(r) &= 4 - 12r^2 + 24r^3 - 36r^4 + 36r^5 - 72r^6 + 160r^7 + 8r^8 - 36r^{10} + + \\
&\quad 72r^{11} - 16r^{12} + 24r^{14} - 16r^{15} - 4r^{18} \\
q_3(r)/4 &= 1 + 2r + 4r^3 + r^4 + 4r^6 - 2r^7 - r^{10} \\
&\quad = 1 + 2r + 4r^3 + r^4(1 - r^6) + 2r^6(2 - r) \\
\end{align*}
\]  

(60)

\[
\begin{align*}
n_4(r) &= -1 + 2r^3 - \frac{9r^4}{2} + 9r^6 - 6r^7 - \frac{9r^8}{2} + 8r^9 - 3r^{10} - 6r^{11} - \\
&\quad - \frac{r^{12}}{2} - 3r^{14} - 2r^{15} - r^{18} + 2r^3\sqrt{1 + 4r^3} - \frac{3}{2}r^4\sqrt{1 + 4r^3} + \\
&\quad 5r^6\sqrt{1 + 4r^3} - 3r^7\sqrt{1 + 4r^3} - \frac{3}{2}r^8\sqrt{1 + 4r^3} + \\
&\quad 2r^9\sqrt{1 + 4r^3} - 3r^{11}\sqrt{1 + 4r^3} - \frac{1}{2}r^{12}\sqrt{1 + 4r^3} - r^{15}\sqrt{1 + 4r^3} \\
\end{align*}
\]  

(61)
that here we don’t know that even $h_n^k$ increase with $n$ and that their limit is $h^{(k)}$. In [18] it is proved by linear algebra that, for $k = 2$ and a tree shift with adjacency restrictions given by an irreducible $d \times d$ matrix, $h_0^{(2)} = h^{(1)} \leq h^{(2)}$. But the linear algebra proof does not give the strict inequality $h^{(1)} < h^{(2)}$.

Monotonicity in $n$ of the $h_n^k$, along with knowledge that the limit in the definition is the infimum (Theorem 2.1), allows to show explicitly by rigorous numerical calculation that $h^{(k)}$ increases with $k$, for the values of $k$ where the monotonicity is known to hold, providing another proof of Theorem 3.7 for these cases:

**Corollary 2.** The entropy of the hard square tree shift on the $k$-tree can increase with the dimension $k$ of the tree (in contrast with what is known about the hard square SFT on the lattice $\mathbb{Z}^k$ [9]). In particular, we have

$$h^{(1)} \approx .481 < h^{(2)} \approx .509 < .536 < h^{(3)} < .548 \approx \frac{\log B_2^{(3)}}{1 + 3 + 3^2}$$

Remark 4. We have shown that for $k = 2, \ldots, 8$ the $h_n^k$ increase strictly to a limit, which is $h^{(k)}$, the entropy of the golden mean $k$-tree shift. It is clear that $\lambda_n < \lambda_{n+1}$, but it is not clear that $h_n^k = \log \lambda_{n-1}/k^n < h_{n+1}^k = \log \lambda_n/k^{n+1}$. The $h_n^k$ supply **strict lower bounds** for $h^{(k)}$, and they increase to $h^{(k)}$ very quickly, providing efficient numerical approximations. For $k=2$, we find $h^{(2)} \approx .509 > .481 \approx h^{(1)}$. It is curious

| $n$ | $B_{n-1}$ | $B_{n-1}(0)$ | $r_{n-1}(0)$ | $\log B_{n-1}/(2^n - 1)$ | $\log B_{n-1}/2^n$ | $h_n^{(2)}$ |
|-----|-----------|--------------|-------------|--------------------------|----------------------|-------------|
| 1   | 2         | .5           | .693        | .347                     | .5025                |
| 2   | 5         | .8           | .536        | .402                     | .5078                |
| 3   | 41        | .6098        | .531        | .465                     | .50866               |
| 4   | 2306      | 1681         | .729        | .516                     | .484                 | .50885      |
| 5   | 8,143,397 | 5,317,636    | .653        | .513                     | .497                 | .508889     |

Table 1. Strip entropy estimation for $k = 2$
Proof. Given $k$, by Theorem 2.1 for any $m$ we can approximate $h^{(k)}$ from above by $\log B_m^{(k)} / |\Delta_m^{(k)}|$, and for $k = 2, \ldots, 8$ we can approximate $h^{(k+1)}$ from below by $h^{(k)}$ for any $n$, so checking numerically that $\log B_m^{(k)} / |\Delta_m^{(k)}| < h^{(k+1)}$ ($n = 4$ suffices for the above statements) shows rigorously that $h^{(k)} < h^{(k+1)}$ for that $k$. □

We discuss now a plan to extend Theorem 4.1 to all $k \geq 2$ (and not just $k = 2, \ldots, 8$). It suffices to show that for each fixed $k \geq 2$ equality in any of (52)–(55) can hold for at most one $r \in [0, 1]$, which is irrational (and in fact is the solution of $Tr = r$), and so in particular the inequality holds for all the $r_n - 1, r_1, \ldots$, which are rational.

Note that as $r$ increases from 0 to 1, $c(r)$ increases from 1 to the golden mean $\gamma$, $Tr$ decreases from 1 to 1/2, and so $c(Tr)$ decreases from $\gamma$ to $c(1/2) = (1 + \sqrt{1 + 8/2^k})/2$, while $1 + 1/c(r)^k$ decreases from 2 to $1 + 1/\gamma^k$.

When $r = 0$, $c(Tr) = \gamma \geq 1 = c(r)^k$, so (52) holds.

When $r = 1$, we need to show

$$1 + \sqrt{1 + 2^{-k+3}} > 4 \left(\frac{\gamma}{2}\right)^k.$$  

Taking the 1 from the left side to the right, squaring both sides, and simplifying leaves

$$\frac{\gamma^{2k}}{2k-1} < \gamma^k + 1.$$  

Equivalently (since $\gamma^2 = \gamma + 1$),

$$\left(\frac{\gamma + 1}{2}\right)^k < \frac{\gamma^k + 1}{2},$$  

which is easily proved by induction on $k = 1, 2, \ldots$:

$$\left(\frac{\gamma + 1}{2}\right)^{k+1} < \left(\frac{\gamma^k + 1}{2}\right) \left(\frac{\gamma + 1}{2}\right), \text{ which is } < \frac{\gamma^{k+1} + 1}{2},$$

because

$$\gamma^{k+1} + \gamma^k + \gamma + 1 < 2\gamma^{k+1} + 2 \text{ if and only if }$$

$$\gamma^{k+1} - \gamma^k - \gamma + 1 > 0, \text{ equivalently }$$

$$\gamma^k (\gamma - 1) - \gamma + 1 > 0, \text{ or }$$

$$\gamma^k - 1 > 0,$$

which is true. Thus the desired inequality holds at the extremes $r = 0, 1$.

We want to show that there is at most one value of $r \in [0, 1]$ for which equality can hold in (52). Given this and the continuity of the functions being compared, the inequality 52 holds on all of $[0, 1]$ except at the point where $Tr = r$, where it is an equality.
We have proved (53) above by explicit algebraic calculation in the cases $k = 2, 3, \ldots, 8$, from which it emerges that its left side $n_k(r)$ is divisible by $p_k(r)^2$, where

$$p_k(r) = r^{k+1} + r - 1.$$  \hspace{1cm} (68)

(This polynomial appears also in [13, Example 16].)

It is difficult to make an estimate to prove any of these inequalities. For example, the graph of the left side $n_k(r)$ of inequality (53) touches the $r$ axis with a high degree of tangency, not allowing one to squeeze a simpler expression in between. Figure 4 shows the two terms in (76) of Prop. 1 and their difference for $k = 5$.

If one can obtain a workable expression for $n_k(r)/p_k(r)^2$, however, one could seek to prove, for example,

$$\frac{n_k(r)}{p_k(r)^2} \leq -\frac{1}{2} \text{ for } r \in [1, \gamma],$$  \hspace{1cm} (69)

thus proving (53).

We believe that the irrationality of the single solution will follow from Gauss’ Lemma.

The following observations may be helpful in completing the proof of strict increase of $h_n^{(k)}$ for each $k$.

**Proposition 1.** The following are equivalent:

$$Tr = r,$$  \hspace{1cm} (70)

$$r^{k+1} + r - 1 = 0,$$  \hspace{1cm} (71)

$$cr = 1,$$  \hspace{1cm} (72)

$$c^{k+1} - c^k - 1 = 0.$$  \hspace{1cm} (73)

Moreover, each of (70)–(73) implies

$$\lambda_n^k = \lambda_{n+1},$$  \hspace{1cm} (74)

$$c(Tr) - (Tr)^{k-1}c(r)^k = 0,$$  \hspace{1cm} (75)

$$c^{2k} - (1 + r^k)^{k-1}(c^k + 1) = 0,$$  \hspace{1cm} (76)

$$c^{2k} - [(c(c - 1))^{k/(k-1)} + 1]^{k-1}(c^k + 1) = 0.$$  \hspace{1cm} (77)
\[
\left(\frac{c + r^{k-1}}{v}\right)^k = \frac{c^k}{v} + 1.
\]  
(78)

**Proof.** That (70) is equivalent to (71) is clear.

(70) implies (72) because \(Tr = r\) implies

\[
\frac{1}{r^2} - \frac{1}{r} = r^{k-1} = c^2 - c,
\]
(79)
so that

\[
0 = c^2 - \frac{1}{r^2} - (c - \frac{1}{r}) = ((c - \frac{1}{r}) - 1),
\]
which implies that \(c = 1/r\).

(72) implies (70) because if \(c = 1/r\) we have

\[
r^{k-1} = c^2 - c = \frac{1}{r^2} - \frac{1}{r},
\]
(81)
so that \(Tr = r\).

To see that (72) implies (73), note that we already know that (71) is equivalent to (72), so that \(cr = 1\) implies

\[
c^{k+1} - c^k - 1 = \frac{1}{r^{k+1}} - \frac{1}{r^k} = 1 = r^{k+1}(1 - r - r^{k+1}) = 0.
\]
(82)

Finally, (73) implies (72) because

\[
0 = c^{k+1} - c^k - 1 = c^{k-1}(c^2 - c) - 1 = c^{k-1}r^{k-1} - 1
\]
(83)
implies \(cr = 1\).

Turning now to the second part of the Proposition, if \(Tr = r\) then (75) says \(c(r) = r^{k-1}c(r)^k\), which is true because \(cr = 1\).

We finish the argument by showing that (74)–(78) are equivalent to one another.

Note that \(\lambda_n^k = \lambda_{n+1}\) if and only if

\[
f_{n+1}(\lambda_n^k) = \lambda_n^{2k} - B_{n+1}^{k-1} - [B_{n+1}B_{n+1}(0)]^{k-1} = 0,
\]
(84)
equivalently, since \(c_n = \lambda_n/B_{n+1}^{k-1}\),

\[
\frac{f_{n+1}(\lambda_n^k)}{B_n^{2k(k-1)}} = c_n^{2k} - c_n^k(1 + r_n^k)^{k-1} - (1 + r_n^k)^{k-1} = 0.
\]
(85)
This shows that (74) is equivalent to (76).

Further, \(\lambda_n^k = \lambda_{n+1}\) says that

\[
B_n^{k(k-1)c_n^k} = B_{n+1}^{k-1} = [B_n^k(1 + r_n^k)]^{k-1}c_{n+1},
\]
(86)
equivalently

\[
c_n^k = (1 + r_n^k)^{k-1}c_{n+1}.
\]
(87)
This shows that (74) is equivalent to (75).

That (76) is equivalent to each of (77) and (78) is direct from \(c^2 - c = r^{k-1}\).
6. Map of the simplex. Consider a $k$-tree SFT $Z_M$ with $k \geq 2$, alphabet size $d = 2$, and transition matrix $M = [11, 10]$, so that $Z_M$ consists of all labelings of the $k$-tree with no adjacent nodes both labeled by 1. Consider first the case $k = 2$. In this case $T : K_1 \rightarrow K_1$ is a function of a single variable: $T(x, 1-x) = (1, x^2)/(1+x^2)$. The equations $Tx = x$ reduce to $x^3 + x - 1 = 0$, with root $u = 0.682328$. The derivative at the fixed point is $T'(u) = -0.635345$. The fixed point is attracting: $T^n x \rightarrow u$ for all $x \in [0, 1]$.

In the case of the $k$-tree, $k \geq 2$, $T_k : K_1 \rightarrow K_1$ (again identifying $K_1$ with $[0, 1]$ according to $(x, 1-x) \leftrightarrow x$) is the function $T_k x = 1/(1 + x^k)$. We may consider this function for all $k > 0$, not necessarily an integer. When $k$ is fixed, we will suppress the subscript on $T_k$. Again there is a fixed point $u_k = u \in [0, 1]$, which is unique because $T$ is decreasing. For small $k$ we have $|T'(u)| < 1$, so the fixed point is attracting. But there is a critical value $k_0 \approx 4.141$, the solution of

$$k_0 = 1 + k_0^{k_0/(k_0+1)}, \quad (88)$$

at which $T'(u) = -1$. (See (90)–(93) and replace inequalities by equalities. For $k = k_0$, the fixed point is $u = 1 - 1/k$.) We claim that for $k > k_0$ besides the fixed point $u$ there is also an attracting periodic orbit $\{p_1, p_2\}$, and there are no other periodic points. Figure 5 shows the graphs of $T, T^2$, and the identity function for $k = 7$.

**Theorem 6.1.** (i) For $k < k_0$ the map $T = T_k$ on $K_1$ has a fixed point $u = u_k$ which attracts all of $[0, 1]$, and there are no other periodic orbits.

(ii) For $k > k_0$ the map $T = T_k$ on $K_1$ has a repelling fixed point $u = u_k$, an attracting periodic orbit $\{p_1, p_2\}$ which attracts all of $[0, 1] \setminus \{u\}$, and no other periodic orbits.

**Proof.** Since $T$ is decreasing and $T^2$ is increasing, we have $0 \rightarrow T0 = 1 \rightarrow T1 = 1/2 \rightarrow \ldots$, so that $T^{2n}0 \nearrow p_1$ and $T^{2n+1}0 \searrow p_2$, with $p_1 \leq u \leq p_2$.

Moreover, $T[0, p_1] = [p_2, 1]$, and $T[p_2, 1] = [1/2, p_1]$. Similarly, $T[p_1, u] = [u, p_2], T[u, p_2] = [p_1, u]$.

Computer assisted computation shows that if $k < k_0$ then

$$|DT^2(x)| \leq c_k < 1 \quad \text{for all } x \in [0, 1]. \quad (89)$$
Thus for $k < k_0$, $p_1 = u = p_2$ and all of $[0, 1]$ is attracted to the unique fixed point $u$.

We show now that on the other hand, for $k > k_0$ the periodic cycle $\{p_1, p_2\}$ attracts all of $[0, 1]$. First, $[0, p_1] \cup [p_2, 1]$ is attracted to $\{p_1, p_2\}$, because of the observations above.

Note further that if $k > k_0$, then at the fixed point $u$ of $T_x = 1/(1+x^k)$ we have $|T'(u)| > 1$. Because consider the function $g(k) = k^{k/(k+1)} + 1$ for $k \geq 0$. We have $g(0) = 2, g(k_0) = k_0, g'(k) < 0$ for small $k$, $g'(k) = 0$ when $\log k = k + 1$ ($k = k_1 \approx 0.3$), $g'(k) > 0$ for $k > k_1$. Thus $g(k) > k$ for small $k$, there is a unique solution $k_0 \approx 4.125$ of $g(k) = k$, and $g(5) \approx 4.82362 < 5$, so $g(k) < k$ for $k > k_0$.

See Figure 6.

This implies that for $k > k_0$ we have $|T_k'(u)| > 1$. Because for $k > k_0$ we have

$$1 + k^{k/(k+1)} < k, \quad k - 1 > k^{k/(k+1)} , \quad (k - 1)^{k+1} > k^k ,$$

$$\frac{(k - 1)^{k+1}}{k^{k+1}} > \frac{1}{k} , \quad 1 < 1 - \frac{1}{k} + (1 - \frac{1}{k})^{k+1} = \frac{k - 1}{k} + \left(\frac{k - 1}{k}\right)^{k+1} ,$$

$$\frac{1}{1 + (\frac{k-1}{k})^k} < \frac{k - 1}{k} ;$$

which is to say that

$$T_k \left(\frac{k - 1}{k}\right) < \frac{k - 1}{k} ,$$

so that $(k - 1)/k > u, k - ku > 1$ for $k > k_0$. But

$$T_k'(u) = \frac{-ku^{k-1}}{(1+u^k)^2} = -k u^{k+1}$$

since $u^{k+1} = 1 - u$. Therefore

$$|T_k'(u)| = ku^{k+1} = k - ku > 1 .$$

Since $|T'(u)| > 1$, there are $\delta, \epsilon > 0$ such that $|T_x - u| \geq (1 + \delta)|x - u|$ for $|x - u| \leq \epsilon$. If we take a point $x > u$ close to $u$, we have $T_x < u, T^2 x > u$. We must
have \( u < x < T^2 x \), because otherwise \( u < T^2 x < x \) and

\[
|T^2 x - u| \geq (1 + \delta)|T x - u| \geq (1 + \delta)^2|x - u| \geq |x - u|,
\]

a contradiction. Thus \( T^2 \) moves points near \( u \) away from \( u \), towards either \( p_1 \) or \( p_2 \), depending on whether \( x \in (p_1, u) \) or \( x \in (u, p_2) \).

We show now that the 2-cycle \( \{p_1, p_2\} \) attracts all of \([0, 1] \setminus \{u\}\). This will follow if we can show that (for \( k > k_0 \)) \( T^2 \) has a unique inflection point. Because from above, we have that for \( p_1 < x < u \) and \( x \) near \( u \), \( p_1 < T^2 x < x \). Since \( T^2 \) is increasing, even though maybe no longer is \( T^2 x \approx u \), we may iterate to find

\[
p_1 < T^4 x < T^2 x < x < u,
\]

so that \( T^{2n} \) decreases to a fixed point \( p \geq p_1 \) of \( T^2 \).

But if \( T^2 \) has a unique inflection point in \([0, 1]\), we must have \( p = p_1 \). This is because then the graph of \( T^2 \) hits the line \( y = x \) at least at \( p_1, x, u, T p, p_2 \), and so must change concavity at least twice. This shows that all of \( (p_1, u) \) is attracted to \( p_1 \) under \( T^2 \), and therefore all of \([0, 1] \setminus \{u\}\) is attracted to \( \{p_1, p_2\} \) under \( T \).

We show now that \( T^2 \) has a unique inflection point in \([0, 1]\). With the help of Sage, we find that

\[
T^2 = \frac{(x^k + 1)^k}{(1 + (x^k + 1))^k},
\]

\[
(T^2)' = \frac{k^2(x^k + 1)^{(k-1)}}{(x^k + 1)^2((x^k + 1)^2 + 1)^2},
\]

and \((T^2)''\) is the quotient of

\[
(k^2 \frac{x}{(x^k + 1)} - k^2 x^k - \frac{x}{(x^k + 1)})^k + k \frac{k}{(x^k + 1)^k} + k - x^k - \frac{1}{(x^k + 1)^k - 1}) k^2 \frac{x^{k-2}}{(x^k + 1)^k}
\]

divided by

\[
(x^k + 1)^2(\frac{1}{(x^k + 1)^k + 1})^3.
\]

Let \( y = x^k \) and

\[
n1(x, k) = (T^2)''((x^k + 1)^2(1/(x^k + 1)^k + 1)^3),
\]

\[
n2(x, k) = k^2 \frac{x}{(x^k + 1)^k} - k^2 x^k - \frac{x}{(x^k + 1)^k} + k \frac{k}{(x^k + 1)^k} + k - \frac{1}{(x^k + 1)^k} - 1,
\]

\[
n3(x, k) = n2(x, k)(x^k + 1)^k,
\]

\[
n4(x, k) = k^2 x^k(1 - (x^k + 1)^k - x^k + k + (k - x^k)(x^k + 1)^k - 1 - (x^k + 1)^k), \quad \text{and}
\]

\[
j(y, k) = k^2 y(1 - (y + 1)^k) - y + k + (k - y)(y + 1)^k - 1 - (y + 1)^k,
\]

all of which have the same sign as \((T^2)''\) on \([0, 1]\).

Now \( j(y, k) \) is a polynomial in \( y \) of degree \( k+1 \). The constant term and coefficient of \( y \) are positive, and for \( j > 1 \) the coefficient of \( y^j \) is

\[
(k - 1)\binom{k}{j} - (k^2 + 1)\binom{k}{j - 1},
\]
which is negative for \( k > 2 \). So the second derivative of \( j(y, k) \) is negative, its first derivative starts positive at 0 and later is negative so that \( j \) has a unique zero on the positive axis. Since \( (T^2)'' \) has the same sign as \( j \), \( T \) has a unique inflection point. Therefore, from above, for \( k > k_0 \) \( T \) has a unique fixed point and an attracting cycle of period 2 which attracts everything else, and there are no other periodic points.

7. Intermediate entropy. For a tree labeled by elements of a finite alphabet \( A = \{0, 1, \ldots, d - 1\} \), one can count numbers of possible labelings not just for the fundamental height \( n \) subtrees but for any finite patterns. Focusing in this section on the binary regular tree, we are interested in particular patterns which consist of a translate of \( \Delta_n \) and an ‘initial’ segment of the next row (the final row of the corresponding translate of \( \Delta_{n+1} \)). Let us label the nodes \( \eta \) left to right in each row, continuing downward from row to row: so \( \epsilon \leftrightarrow \eta(1), 0 \leftrightarrow \eta(2), 1 \leftrightarrow \eta(3), 00 \leftrightarrow \eta(4), 01 \leftrightarrow \eta(5), 10 \leftrightarrow \eta(6), 11 \leftrightarrow \eta(7), 000 \leftrightarrow \eta(8), \ldots \). (Recall that nodes of the tree correspond to words in \( \{0, 1\}^* \)—see Section 1.1). Define \( q(n) \) to be the number of labelings of \( \eta(1)\eta(2)\ldots\eta(n) \) in the tree shift. (For a single labeled tree, we define \( q(n) \) to be the number of labelings found among all translates of \( \eta(1)\eta(2)\ldots\eta(n) \).)

For \( n = 0, 1, \ldots \) let \( c_n = |\Delta_n| = 2^{n+1} - 1 \). Thus
\[
q(c_n) = p(n) \quad \text{for all } n = 0, 1, \ldots \tag{106}
\]

We ask whether the intermediate entropy \( h^i = \lim_{n \to \infty} \frac{\log q(n)}{n} \) exists (and hence equals the entropy \( h = \lim_{n \to \infty} \frac{\log q(c_n)}{c_n} \) of the tree shift or labeled tree). For now let us define the upper and lower intermediate entropies of a tree shift or labeled tree to be
\[
h^u = \limsup_{n \to \infty} \frac{\log q(n)}{n} \quad \text{and} \quad h^l = \liminf_{n \to \infty} \frac{\log q(n)}{n}. \tag{107}
\]

**Proposition 2.** The upper intermediate entropy of a 2-tree shift or labeled 2-tree cannot exceed its entropy: \( h^u \leq h \).

**Proof.** Fix \( n \) and \( j = 0, \ldots, 2^{n+1} \) and consider the pattern \( \Delta_n(j) \) that consists of \( \Delta_n \) together with an initial segment of length \( j \) of the next row in \( \Delta_{n+1} \). We wish to estimate from above the number of possible labelings of \( \Delta_n(j) \) that can be found in the tree shift. Consider first the case when \( j = 2^n \), i.e. we use half of the next row; see Figure 7.

\[\Delta_n(2^n)\] consists of the root, a translate of \( \Delta_n \) (the blue nodes in Figure 7), and a translate of \( \Delta_{n-1} \) (the green nodes in Figure 7). There are \( d \) choices for labels.
of the root, \( q(c_n) \) choices for labelings of \( \Delta_n \), and \( q(c_{n-1}) \) choices for labelings of \( \Delta_{n-1} \), so

\[
q(c_n + 2^n) \leq dq(c_n)q(c_{n-1}).
\]  

(108)

Since \( c_n + c_{n-1} = c_n + 2^n - 1 \),

\[
\lim_{n \to \infty} \frac{\log q(c_n)}{c_n} = h,
\]  

(109)

and

\[
\frac{\log q(c_n + 2^n)}{c_n + 2^n} \leq \frac{\log d}{c_n + 2^n} + \frac{c_n \log q(c_n)}{c_n + 2^n} + \frac{c_{n-1} \log q(c_{n-1})}{c_n + 2^n},
\]  

(110)

we have

\[
\limsup_{n \to \infty} \frac{\log q(c_n + 2^n)}{c_n + 2^n} \leq h.
\]  

(111)

Let \( \epsilon > 0 \), choose \( r \) so that \( 1/2^r < \epsilon \), and suppose that \( n >> r \). We divide the \( 2^{n+1} \) nodes on the last row of \( \Delta_{n+1} \) into \( 2^r \) consecutive intervals, each of length \( 2^m \), where \( m = n - r + 1 \). Suppose first that \( j = k2^m \) for some \( k = 0, \ldots, 2^r - 1 \). Write \( k = k_0 + k_12 + \cdots + k_r2^{-1} \) with each \( j_i = 0 \) or 1 and let \( i_0 = \inf \{i : k_i \neq 0\} \). We proceed to decompose \( \Delta_n(j) \) as a subset of the disjoint union of the root and at most \( n \) other nodes, a translate of \( \Delta_{i_0}(2^{i_0}) \), and translates of \( \Delta_{i_0}, \Delta_{i_0+1}, \ldots, \Delta_n \). Note that \( |\Delta_{i_0}(2^{i_0})| + |\Delta_{i_0}| + |\Delta_{i_0+1}| + \cdots + |\Delta_n| \) is within \( n \) of \( |\Delta_n(j)| \). Figure 8 shows an example with \( n = 3, r = 3, j = 10, r = 3, m = 1, k = 5 = 1 + 0 \cdot 2 + 1 \cdot 2^2, i_0 = 0 \). Here the green nodes form a translate of \( \Delta_{i_0}(2^{i_0}) \) (a subset of a translate of \( \Delta_2 \)), the red nodes form a translate of \( \Delta_{i_0} \), the blue nodes form a translate of \( \Delta_n \), and the black nodes are the “free” ones.

![Figure 8. Using 5/8 of the next row.](image)

This decomposition yields the estimate

\[
q(c_n + j) \leq d^n q(c_{i_0})q(c_{i_0-1}) \prod_{t=i_0}^{n-1} q(c_t).
\]  

(112)

This implies that with \( r \) fixed, once \( n \) is large enough that

\[
\frac{\log q(c_t)}{c_t} < h + \epsilon \quad \text{for most } t = i_0, \ldots, n - 1,
\]  

(113)

we will have for all \( j = k2^m \) for some \( k = 0, \ldots, 2^r - 1 \),

\[
\frac{\log q(c_n + j)}{c_n + j} < h + 2\epsilon.
\]  

(114)
Continuing with \( r \) fixed, we consider now values of \( j \) which are not multiples of \( 2^m \). For any such \( j = 1, \ldots, 2^{n+1} - 1 \), let \( j' = \inf\{k2^m > j\} \). Then
\[
\frac{\log q(c_n + j)}{c_n + j} \leq \frac{\log q(c_n + j')}{c_n + j'} \leq (h + \epsilon) \frac{c_n + j + 2^m}{c_n + j} < (h + \epsilon) \frac{c_n + j + 2^m}{c_n + j} \leq (h + \epsilon)(1 + 2^m/c_n) < h + 3\epsilon
\]
for large enough \( n \).

**Corollary 3.** Let \( Z = Z(X) \) be the 2-tree shift corresponding to a 1-dimensional shift of finite type \( X \). Then the intermediate topological entropy of \( Z(X) \) equals its topological entropy: \( h^i(Z(X)) = h(Z(X)) = h^u(Z(X)) \geq h(X) \). More generally, if \( Z \) is a tree SFT (determined by excluding a finite set of blocks), then its intermediate topological entropy equals its topological entropy: \( h^i(Z) = h(Z) \).

**Proof.** In a tree shift corresponding to a one-step one-dimensional shift of finite type, labelings of disjoint translates of basic patterns \( \Delta_n \) are independent given the configurations above them, so the key inequalities in the above proof are actually equalities. For systems with more memory (still bounded), consider the appropriate higher block coding.

For the general case of a tree SFT, labelings of sets of nodes in distinct branches of the tree are again independent, given a fixed allowed configuration above both of them: pairing two such allowed labelings produces an allowed labeling of their union.

**8. Larger alphabets and more general labeling restrictions.** Let \( M \) be an irreducible \( d \times d \) \( \{0, 1\} \) matrix and \( Z_M \) the \( k \)-tree shift consisting of all \( k \)-trees labeled by elements of the alphabet \( A = \{0, \ldots, d - 1\} \) consistently with the transitions allowed by \( M \), as in Section 1.1. We describe briefly the setup for applying the strip method to this general situation to study the topological entropy \( h(Z_M) \) of the system \( Z_M \).

For each \( n = 0, 1, \ldots \) let \( x(n) = (x_0(n), \ldots, x_{d-1}(n)) \) denote the vector of symbol counts in which \( x_i(n) \) is the number of labeled translates of \( \Delta_n \) \((n\text{-blocks})\) with the symbol \( i \) at the root. Let \( |x(n)| = x_0(n) + \cdots + x_{d-1}(n) \) As before, denoting by \( M_i \) the \( i \)’th row of \( M \), we have the recursion
\[
x_i(n + 1) = (Mx(n))^k_i = [M_i \cdot x(n)]^k, \quad i = 0, \ldots, d - 1.
\]
Define
\[
g_k(x) = \sum_j (Mx)^k_j
\]
and the map \( T_k : \mathbb{R}^d \to \mathbb{R}^d \) by
\[
(T_kx)_i = \frac{(Mx)^k_i}{\sum_j (Mx)^k_j} = \frac{(Mx)^k_i}{g_k(x)}, \quad i = 0, 1, \ldots, d - 1,
\]
and as before
\[
r(n) = \frac{x(n)}{|x(n)|}, \quad \text{so that}
\]
\[
r(n + 1) = T_k r(n).
\]

For the 1-dimensional golden mean SFT, the map \( T \) on \( K_1 = \{(x, y) : x \geq 0, y \geq 0, x + y = 1\} \) is \( T(x,y) = ((x+y)^2, x^2)/(x+y)^2 + x^2) \). The equations
Figure 9. The graph of the fixed point equations for the golden mean $T(x, y) = (x, y)$ define an algebraic curve, part of which is shown in Figure 9 along with the simplex $x + y = 1$.

Remark 5. When estimating the entropy $h^{(k)}$ of $Z_M$ numerically, iteration of the mapping $T_k$ combined with a recursion on $\log |x(n)|$ can reduce the size of the numbers involved and give us more accurate estimates more quickly. Since

$$\begin{align*}
x_i(n+1) &= |x(n)|^k (Mr(n))_i^k, i = 0, \ldots, d - 1, \\
|x(n+1)| &= |x(n)|^k g_k(r(n)),
\end{align*}
$$

we may form the $T_k r(n)$ by iterating $T_k$ and use them in the recursion

$$\log |x(n+1)| = k \log |x(n)| + \log g_k(r(n))$$

(121)

to form rapidly improving approximations to

$$h = \lim_{n \to \infty} \frac{\log |x(n)|}{1 + k + \cdots + k^n}.$$ 

(122)

The factors $g_k(r(n))$ are bounded, but they have a cumulative effect on the growth of $|x(n)|$ that may be sufficient to affect the ultimate value of $h$.

Let $C(-1) = M$ and for each $n \geq 0$ define a $d \times d$ matrix $C(n)$ by

$$C(n)_{ij} = M_{ij} (M_i \cdot x(n))^{k-1}.$$ 

(123)

Denote by $\lambda_n$ the maximal eigenvalue of $C(n)$ and define $c_n$ by

$$\lambda_n = |x(n)|^{k-1} c_n.$$ 

(124)
The entropy of the one-dimensional $n$’th strip approximation subshift to $\mathbb{Z}^M$ is
\begin{equation}
    h_n^{(k)} = \frac{\log \lambda_{n-1}}{k^n}.
\end{equation}

**Theorem 8.1.** With notation as above, the entropy of the $k$-tree shift corresponding to an irreducible $d \times d \times 0,1$ matrix $M$ is given by the following infinite series formula:
\begin{equation}
    h^{(k)}(\mathbb{Z}^M) = \frac{k-1}{k} \log d + (k-1) \sum_{i=1}^{\infty} \frac{1}{k^{i+1}} \log g_k(r(i-1)).
\end{equation}

**Proof.** We compute that
\begin{equation}
    h_{n+1}^{(k)} = h_n^{(k)} + \frac{k-1}{k} \log g_k(r(n-1)) + \frac{1}{k^{n+1}} \log c_n - \frac{1}{k^n} \log c_{n-1},
\end{equation}
\begin{equation}
    h_1^{(k)} = \frac{k-1}{k} \log d + \frac{1}{k} \log(c_0),
\end{equation}
\begin{equation}
    h_2^{(k)} = \frac{k-1}{k} \log d + \frac{k-1}{k^2} \log g_k(r(0)) + \frac{1}{k^2} \log c_1,
\end{equation}
and then by induction that
\begin{equation}
    h_n^{(k)} = \frac{k-1}{k} \log d + (k-1) \sum_{i=1}^{n-2} \frac{1}{k^{i+1}} \log g_k(r(i-1)) + \frac{1}{k^n} \log c_{n-1}.
\end{equation}

We showed before that for the golden mean SFT on the $k$-tree, the site specific entropies $h_n^{(k)}$ increase with the strip width to the tree shift entropy $h^{(k)}$ (at least for $k = 2, \ldots, 8$), and that $h^{(k)}$ is strictly increasing in $k$, with limit $\log 2$, in contrast with the situation for the golden mean SFT’s on integer lattices. The same statement for the more general tree shifts considered in this section can be approached by the same techniques, although the formulas and computations will naturally be much more complex. Extensions to pressure and equilibrium states, including measures of maximal entropy, and subshifts on other trees and graphs, are also attractive topics for further research.

**Acknowledgments.** We thank Professors Tim Austin, Henk Bruin, Kevin McGoff, Tom Meyerovitch, Anthony Quas, and Ville Salo for helpful discussions on these topics, and the referees for their comments and corrections.

**REFERENCES**

[1] N. Aubrun and M.-P. Béal, Decidability of conjugacy of tree-shifts of finite type, in *Automata, Languages and Programming Part 1*, Springer, Berlin, Heidelberg, 2009, 132–143.

[2] N. Aubrun and M.-P. Béal, Sofic and almost of finite type tree-shifts, in *Computer Science – Theory and Applications*, Springer, Berlin, 2010, 12–24.

[3] N. Aubrun and M.-P. Béal, Tree-shifts of finite type, *Theoret. Comput. Sci.*, **459** (2012), 16–25.

[4] N. Aubrun and M.-P. Béal, Sofic tree-shifts, *Theory Comput. Syst.*, **53** (2013), 621–644.

[5] N. Aubrun and M.-P. Béal, Tree algebra of sofic tree languages, *RAIRO Theor. Inform. Appl.*, **48** (2014), 431–451.

[6] T. Berger and Z. X. Ye, Entropic aspects of random fields on trees, *IEEE Trans. Inform. Theory*, **36** (1990), 1006–1018.

[7] L. P. Bowen, A brief introduction to sofic entropy theory, in *Proceedings of the International Congress of Mathematicians – Rio de Janeiro 2018*, Vol. 3, World Sci. Publ., Hackensack, NJ, 2019, 1847–1866, arXiv:1711.02062v1.
[8] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics, Vol. 527, Springer-Verlag, Berlin-New York, 1976.
[9] D. Gamarnik and D. Katz, Sequential cavity method for computing free energy and surface pressure, *J. Stat. Phys.*, 137 (2009), 205–232.
[10] M. Hochman and T. Meyerovitch, A characterization of the entropies of multidimensional shifts of finite type, *Ann. of Math.*, 171 (2010), 2011–2038.
[11] D. Kerr and H. Li, *Ergodic Theory*. Independence and Dichotomies, Springer, Cham, 2016.
[12] E. Louidor, B. Marcus and R. Pavlov, Independence entropy of $\mathbb{Z}^d$-shift spaces, *Acta Appl. Math.*, 126 (2013), 297–317.
[13] J. Mairesse and I. Marcovici, Uniform sampling of subshifts of finite type on grids and trees, *Internat. J. Found. Comput. Sci.*, 28 (2017), 263–287.
[14] B. Marcus and R. Pavlov, Approximating entropy for a class of $\mathbb{Z}^2$ Markov random fields and pressure for a class of functions on $\mathbb{Z}^2$ shifts of finite type, *Ergodic Theory Dynam. Systems*, 33 (2013), 186–220.
[15] T. Meyerovitch and R. Pavlov, On independence and entropy for high-dimensional isotropic subshifts, *Proc. Lond. Math. Soc.*, 109 (2014), 921–945.
[16] R. Pavlov, Approximating the hard square entropy constant with probabilistic methods, *Ann. Probab.*, 40 (2012), 2362–2399.
[17] K. Petersen and I. Salama, Entropy on regular trees, preprint, arXiv:1909.05153v1, 2019.
[18] K. Petersen and I. Salama, Tree shift topological entropy, *Theoret. Comput. Sci.*, 743 (2018), 64–71.
[19] S. T. Piantadosi, Symbolic dynamics on free groups, *Discrete Contin. Dyn. Syst.*, 20 (2008), 725–738.
[20] Z. Ye and T. Berger, *Information Measures for Discrete Random Fields*, Science Press Beijing, Beijing, 1998.

Received September 2019; 1st revision December 2019; 2nd revision December 2019.

E-mail address: petersen@math.unc.edu
E-mail address: isalama@nccu.edu