Transfer Operator, Topological Entropy, and Maximal Measure for Cocyclic Subshifts

Jaroslaw Kwapisz

Department of Mathematical Sciences
Montana State University
Bozeman MT 59717-2400
tel: (406) 994 5358
fax: (406) 994 1789
e-mail: jarek@math.montana.edu
web page: http://www.math.montana.edu/~jarek/

August 17, 2001

Abstract

Keywords: Cocyclic subshifts arise as the supports of matrix cocycles over a full shift and generalize topological Markov chains and sofic systems. We compute the topological entropy of a cocyclic subshift as the logarithm of the spectral radius of an appropriate transfer operator and give a concrete description of the measure of maximal entropy in terms of the eigenvectors. Unlike in the Markov or sofic case, the operator is infinite-dimensional and the entropy may be a logarithm of a transcendental number.

Introduction

Cocyclic subshifts, as introduced in [7], are the supports of locally constant matrix cocycles on the full shift over a finite alphabet $\mathcal{A}$. Specifically, given a sequence of square matrices $\mathbf{\Phi} = (\Phi_i)_{i\in\mathcal{A}}$, the associated cocyclic subshift $X_{\mathbf{\Phi}}$ is obtained by disallowing the blocks $\sigma_1\sigma_2\ldots\sigma_n$ (where $\sigma_i \in \mathcal{A}$) for which the matrix product $\Phi_{\sigma_1}\ldots\Phi_{\sigma_n}$ is zero (i.e. of zero rank) — see Definition 1.1. The definition resembles that of sofic systems given in [14] except that the matrix semi-group generated by Φ_i's is not assumed to be finite. As a result, cocyclic subshifts often fail to be sofic and require a new arsenal of tools to analyze them.

[1]Partially supported by NSF grant DMS-9970725.

This is exemplified by [7], where the theory of finite dimensional algebras is used to unravel the dynamical structure of $X_{\mathbf{\Phi}}$ much like the graph theory is used for sofic systems. In particular, $X_{\mathbf{\Phi}}$ decomposes into transitive and intrinsically ergodic components and there is an algorithm for deciding if the topological entropy $h_{\text{top}}(X_{\mathbf{\Phi}})$ is positive. The present paper is mainly devoted to the problem of computing that entropy and to an explicit description of the maximal measure (for transitive $X_{\mathbf{\Phi}}$) in a fashion analogous to Parry's formulas for topological Markov chains. We also construct an example of a cocyclic subshift with $h_{\text{top}}(X_{\mathbf{\Phi}})$ that is a logarithm of a transcendental number. Since $h_{\text{top}}(X_{\mathbf{\Phi}})$ coincides with the growth of periodic orbits, this demonstrates that the dynamical zeta function of $X_{\mathbf{\Phi}}$ may fail to be a rational function.

While formulation of our results requires more definitions and is therefore relegated to the subsequent sections, the overall strategy follows the well established paradigm of computing the entropy via an appropriate *transfer operator*. For a topological Markov chain (or a sofic system), this operator is just the adjacency matrix of the transition graph between states, but for cocyclic subshifts, an infinite dimensional setting is required. Specifically, we shall use the Banach space of all bounded functions defined on the projective space of the linear space upon which the Φ_i's act. Note that if $z \mapsto \phi(z)$ is such a function, the composition $z \mapsto \phi \circ \Phi_i(z)$ makes perfect sense as long as z is not a line in the kernel of Φ_i; and should z be in the kernel, let us define (extend) $\phi(\Phi_i(z))$ to be zero. In this way, each Φ_i determines a bounded linear operator $\mathcal{T}_i : \phi \mapsto \phi \circ \Phi_i$. The transfer operator of $X_{\mathbf{\Phi}}$ is defined as the sum $\mathcal{T}_{\mathbf{\Phi}} := \sum_{i \in \mathcal{A}} \mathcal{T}_i$ (see Section 2 for details). The advantage of this definition is in its simplicity; and the huge *size* of the Banach space is largely offset by a rather penetrable structure of the operator (studied in Sections 3, 4, and 6). Let us outline the main results.

In Section 3, we show that $e^{h_{\text{top}}(X_{\mathbf{\Phi}})}$ coincides with the spectral radius of $\mathcal{T}_{\mathbf{\Phi}}$ (Theorem 2.1). This hinges on approximation of $X_{\mathbf{\Phi}}$ from within by Markov chains (c.f. Theorem 7.2 in [7]). In Section 4, we further characterize $e^{h_{\text{top}}(X_{\mathbf{\Phi}})}$ as a root of a certain *characteristic equation* (Corollary 4.1). Unlike for sofic systems, this equation may be transcendental with transcendental roots. Also in Section 4, we develop Perron-Frobenius type result for $\mathcal{T}_{\mathbf{\Phi}}$ and $\mathcal{T}_{\mathbf{\Phi}}^*$ (Theorem 4.2). Namely, under a natural and easy to secure assumption that the minimal rank of the cocycle is one, i.e. $r_{\mathbf{\Phi}} := \min\{\text{rank}(\Phi_{\sigma_1} \cdots \Phi_{\sigma_n}) : \ \Phi_{\sigma_1} \cdots \Phi_{\sigma_n} \neq 0\} = 1$, we show that the essential spectral radius is strictly smaller than the spectral radius and the later is an eigenvalue corresponding to a positive eigenvector. Dynamical relevance of this eigenvector becomes clear in Section 5 where we use it to recover the measure of maximal entropy. Here the idea may be conveyed again by an analogy with a transitive topological Markov chain. If $A = (a_{ij})_{i,j=1}^m$ is the transition matrix, λ is its spectral radius, and $u = (u_j)_{j=1}^m$ and $v = (v_i)_{i=1}^m$ are the (normalized) eigenvectors of A and its transpose A^T, then the maximal measure is the unique measure on $\mathcal{A}^{\mathbb{N}}$ that assigns to any nonempty cylinder set $\{x \in X_{\mathbf{\Phi}} : \ x_1 =$

$\sigma_1, \ldots, x_n = \sigma_n\}$ a mass

$$\lambda^{-n} v_{\sigma_1} u_{\sigma_n}. \tag{0.1}$$

(In the standard parlance, this is the stationary Markov measure generated by the stochastic vector $(u_i v_i)_{i=1}^m$ and the stochastic matrix $\left(\frac{u_j}{u_i} \frac{a_{ij}}{\lambda}\right)_{i,j=1}^m$ [13].) We give an exact analogue of the above formula for a cocyclic subshift represented as $X_{\mathbf{\Phi}}$ with $r_{\mathbf{\Phi}} = 1$ — see (5.2). Essentially, v_{σ_1} and u_{σ_n} are replaced by the leading eigenvector of $\mathcal{T}_{\mathbf{\Phi}}$ and the leading eigenvector of $\mathcal{T}_{\mathbf{\Phi}^T}$ (where $\mathbf{\Phi}^T := \left(\Phi_i^T\right)_{i \in \mathcal{A}}$ is the transposed cocycle). These eigenvectors (as functions on the projective space) are to be evaluated on $\operatorname{im}(\Phi_{\sigma_1} \ldots \Phi_{\sigma_n})$ and $\operatorname{im}(\Phi_{\sigma_n}^T \ldots \Phi_{\sigma_1}^T)$, respectively, which dictates considering only the *good blocks* $\sigma_1 \ldots \sigma_n$ distinguished by $\operatorname{rank}(\Phi_{\sigma_1} \ldots \Phi_{\sigma_n}) = 1$. In fact, such *good blocks* outweigh the *bad blocks* by a ratio exponentially small in n, and the formula uniquely determines a certain invariant measure that is *nearly-Markov* (Theorem 5.3).

The central role of the good blocks is further explored in Section 6, which foregoes the compact presentation of $X_{\mathbf{\Phi}}$ via the cocycle $\mathbf{\Phi}$ in favor of a certain countable graph with vertices at the points of the projective space associated to $\operatorname{im}(\Phi_{\sigma_1} \ldots \Phi_{\sigma_n})$ where $\sigma_1 \ldots \sigma_n$ is a good block. The advantage of this approach is in reconnecting with the classical theory of countable positive matrices. Indeed, the adjacency matrix of the graph is closely related to $\mathcal{T}_{\mathbf{\Phi}}$, and one can demonstrate some further results on the spectrum and eigenvectors of $\mathcal{T}_{\mathbf{\Phi}}$ (including the simplicity and domination of its Perron eigenvalue — see Theorem 6.4). This point of view also makes clear that transitive cocyclic subshifts are *synchronized coded systems* (as introduced in [3]), which opens a possibility for computing the entropy via a general method based on the so called *loop equation* (see Prop. 2.12 in [3], and the discussion around (7.1) in [10]). Let us also mention that, by the result in [7], mixing cocyclic subshifts have specification and, as such, taken with the maximal measure, are isomorphic (via a finitary isomorphism) with Bernoulli shifts (see Theoreme 2 in [1] and also [11]).

Before concluding this introduction, let us add that, to keep the technicalities to the minimum, we restricted the discussion to the setting in which the cocyclic subshift is presented as $X_{\mathbf{\Phi}}$ for some $\mathbf{\Phi}$. In the remaining sections, we use a slightly more flexible presentation in terms of labeled (colored) directed graphs with matrix multiplicative weights over the edges (which is often preferable in dealing with concrete examples or in devising computer algorithms). The pertaining definitions can be found in Section 1, which should be read before looking over the results of other sections.

1 Preliminaries

We fix notations and recall some key definitions (mainly from [7]). Our attention is confined to the one-sided subshifts; although, a parallel theory exists for two-sided subshifts. As before, $\mathcal{A}$ is a finite alphabet, say $\mathcal{A} = \{1, ..., m\}$. Consider $\mathcal{A}^{\mathbb{N}}$ with the product topology. *The (full) one-sided shift (over A)* is the map $f : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$ such that $f : (\sigma_i)_{i\in\mathbb{N}} \mapsto (\sigma_{i+1})_{i\in\mathbb{N}}$. Any compact $X \subset \mathcal{A}^{\mathbb{N}}$ invariant under f, $f(X) \subset X$, is referred to as a *subshift*. Given a non-zero linear space V over an algebraically closed field, denote by $\mathrm{End}(V)$ all (linear) endomorphisms of V and compose the maps in $\mathrm{End}(V)$ on the right so that V is a right End-module.

Definition 1.1 ([7]) *A cocyclic subshift of* $\mathbf{\Phi} = (\Phi_i)_{i\in\mathcal{A}} \in \mathrm{End}(V)^m$ *is*

$$X_{\mathbf{\Phi}} := \{\sigma \in \mathcal{A}^{\mathbb{N}} : \Phi_{\sigma_1} \cdots \Phi_{\sigma_n} \neq 0,\ \forall n \in \mathbb{N}\}.$$

Somewhat imprecisely we shall refer to a $\mathbf{\Phi}$ in the definition as a cocycle. By a *block* we understand any finite sequence $\sigma \in \mathcal{A}^n$; $n \in \mathbb{N}$ is the *length* of σ, denoted $|\sigma|$. Given $x \in \mathcal{A}^{\mathbb{N}}$, we introduce sub-blocks $[x]_n := (x_1, \ldots, x_n)$ and $[x]_{i,n} := (x_i, \ldots, x_{i+n-1})$ where $i, n \in \mathbb{N}$. Also, the cylinder set associated to σ is $U_\sigma := \{x :\ [x]_{|\sigma|} = \sigma\}$. Having fixed $\mathbf{\Phi}$, a block σ for which $\Phi_\sigma := \Phi_{\sigma_1} \cdots \Phi_{\sigma_n} \neq 0$ is referred to as *allowed*. A block σ *occurs* in $X_{\mathbf{\Phi}}$ iff $\sigma = [x]_n$ for some $x \in X_{\mathbf{\Phi}}$ and $n \in \mathbb{N}$. Clearly, $x \in \mathcal{A}^{\mathbb{N}}$ belongs to $X_{\mathbf{\Phi}}$ iff all its sub-blocks $[x]_n$ are allowed; and $X_{\mathbf{\Phi}} = \mathcal{A}^{\mathbb{N}} \setminus \bigcup\{U_\sigma :\ \sigma \text{ is not allowed}\}$.

Think now of the elements of $\mathcal{A} = \{1, \ldots, m\}$ as encoding colors. By a *directed graph with colored edges* we understand a triple $\mathbf{G} = (\mathbf{Ver}(\mathbf{G}), \mathbf{Edge}(\mathbf{G}), l_{\mathbf{G}})$ where $\mathbf{Ver}(\mathbf{G})$ is a set (of *vertices*), $\mathbf{Edge}(\mathbf{G})$ is a set (of *edges*), and $l_{\mathbf{G}}$ is a (*coloring*) function $l_{\mathbf{G}} : \mathbf{Edge}(\mathbf{G}) \to \mathcal{A}$. Each $e \in \mathbf{Edge}(\mathbf{G})$ has associated unique $e^-, e^+ \in \mathbf{Ver}(\mathbf{G})$ called its *tail* and *head*, respectively. We allow for multiple edges between any two vertices and for self-loops; at this point, we also do not restrict to finite graphs. A sequence of edges $a = (e_i)_{i=1}^n$, where $n \in \mathbb{N} \cup \{\infty\}$, is a *path* in $\mathbf{G}$ iff $e_i^+ = e_{i+1}^-$ for $i = 1, \ldots, n-1$; $|a| := n$ is the *length of* a. We will also regard a single vertex $v \in \mathbf{Ver}(\mathbf{G})$ as a path of length zero, and set $v^+ = v_- = v$. The collection of all paths in $\mathbf{G}$ is denoted by $\mathbf{Path}(\mathbf{G})$. Each path $a = (e_i)_{i=1}^n$ determines a block $\sigma = (l_{\mathbf{G}}(e_i))_{i=1}^n$; we say that σ is *the coloring of* a and write $\sigma = l_{\mathbf{G}}(a)$. The *subshift of a colored directed graph* $\mathbf{G}$ is the subshift $X_{\mathbf{G}} \in \mathcal{A}^{\mathbb{N}}$ given by[1]

$$X_{\mathbf{G}} := \{(x_i)_{i\in\mathbb{N}} :\ \forall_{n\in\mathbb{N}} \exists_{a\in\mathbf{Path}(\mathbf{G})}\ l_{\mathbf{G}}(a) = [x]_n\}. \tag{1.1}$$

Since every subshift can be easily represented as $X_{\mathbf{G}}$ for some countable $\mathbf{G}$, any interesting theory places extra assumptions on $\mathbf{G}$; for instance, finite $\mathbf{G}$ yield *sofic systems* ([5]) and irreducible (i.e. strongly connected) $\mathbf{G}$ yield *coded systems* (Prop 2.1 in [3]).

Now, as in Section 12 of [7], by a *colored graph* $\mathbf{G}$ *with propagation* $\mathbf{\Gamma}$ we understand a pair $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ where $\mathbf{G}$ is a colored directed graph that has each vertex $v \in \mathbf{Ver}(\mathbf{G})$ equipped with a non-zero linear space V_v and each edge $e \in \mathbf{Edge}(\mathbf{G})$ equipped with a linear transformation — called a *propagator* — $\Gamma_e : V_{e^-} \to V_{e^+}$.

[1]For infinite graphs, $X_{\mathbf{G}}$ may strictly contain $\{(l_{\mathbf{G}}(e_i))_{i\in\mathbb{N}} :\ (e_i)_{i\in\mathbb{N}} \text{ is an infinite path in } \mathbf{G}\}$ — see e.g. [10].

Here, $\mathbf{\Gamma} = (\Gamma_e)_{e \in \mathbf{Edge}(\mathbf{G})}$. For a path $a = (e_1, \ldots, e_n)$, write $\Gamma_a := \Gamma_{e_1} \cdots \Gamma_{e_n}$, and say that a *propagates* iff $\Gamma_a \neq 0$. For zero length path, that is $v \in \mathbf{Ver}(\mathbf{G})$, we adopt a convention that $\Gamma_v = \mathbf{Id}$, the identity. By definition, a block $\sigma = (\sigma_1, \ldots, \sigma_n)$ is *allowed* (relative to $\mathbf{P}$) iff it is a coloring of some propagating path a. We also set $\mathbf{Ver}(\mathbf{P}) := \mathbf{Ver}(\mathbf{G})$ and $\mathbf{Edge}(\mathbf{P}) := \mathbf{Edge}(\mathbf{G})$.

Definition 1.2 ([7]) *The subshift of a (finite) colored graph with propagation* $\mathbf{P}$ *is*

$$X_{\mathbf{P}} := \{(l_{\mathbf{P}}(e_i))_{i \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}} : \ (e_i)_{i \in \mathbb{N}} \text{ is an infinite propagating path in } \mathbf{P}\}.$$

For an arbitrary $\mathbf{P}$, $X_{\mathbf{P}}$ is a factor of a cocyclic subshift (see Section 12 of [7]). Whether $X_{\mathbf{P}}$ is actually a cocyclic subshift is open unless one places extra hypotheses on $\mathbf{P}$. Recall that a colored graph $\mathbf{G}$ is *right resolving*, if no two edges with tails at the same vertex coincide in color, i.e. if $e^- = \tilde{e}^-$ and $l(e) = l_{\mathbf{G}}(\tilde{e})$ then $e = \tilde{e}$ for any $e, \tilde{e} \in \mathbf{Edge}(\mathbf{G})$. $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ is *right resolving* iff $\mathbf{G}$ is. *Left resolving* is the analogous notion referring to heads (not tails). We leave to the reader the proof of the following remark that puts at our disposal all the results in [7] that are stated in terms of $X_{\mathbf{\Phi}}$.

Remark 1.1 ([7]) *For a right (or left) resolving* $\mathbf{P}$*,* $X_{\mathbf{P}}$ *is a cocyclic subshift. Specifically, if* $\mathbf{\Phi} = (\Phi_i)_{i \in \mathcal{A}}$ *with* $V := \bigoplus_{v \in \mathbf{Ver}(\mathbf{G})} V_v$ *and* $x\Phi_i := \sum_{e \in \mathbf{Edge}(\mathbf{G}):\ e^- = v} x\Gamma_e$ *for all* $x \in V_v$*,* $v \in \mathbf{Ver}(\mathbf{G})$*, then a block* σ *is allowed with respect to* $\mathbf{P}$ *iff it is allowed with respect to* $\mathbf{\Phi}$*. In particular,* $X_{\mathbf{P}} = X_{\mathbf{\Phi}}$*.*

In order to deal with the iterates of f on $X_{\mathbf{P}}$, it is convenient to introduce together with any $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ its powers $\mathbf{P}^n$. For $n \in \mathbb{N}$, by definition, the graph of $\mathbf{P}^n$ is the power graph $\mathbf{G}^n$, i.e. $\mathbf{Ver}(\mathbf{G}^n) := \mathbf{Ver}(\mathbf{G})$ and $\mathbf{Edge}(\mathbf{G}^n) := \{a \in \mathbf{Path}(\mathbf{G}) : \ |a| = n\}$ with the obvious head and tail assignment. The coloring $l_{\mathbf{P}^n} : \mathbf{Edge}(\mathbf{G}^n) \to \mathcal{A}^n$ is just the (restriction of) Cartesian power $l_{\mathbf{P}} \times \ldots \times l_{\mathbf{P}}$, and the propagator of $a \in \mathbf{Ver}(\mathbf{G}^n)$ is simply Γ_a. The subshift $X_{\mathbf{P}^n} \subset (\mathcal{A}^n)^{\mathbb{N}}$ is easily seen to be *the* n^{th} *power* of the subshift $X_{\mathbf{P}}$, and as such is naturally conjugated to the n^{th} iterate of f on $X_{\mathbf{P}}$ (see [7, 8]). Clearly, if $\mathbf{P}$ is right (left) resolving so is $\mathbf{P}^n$.

Finally, we define $\mathbf{P}^T$ as $\mathbf{P}^T = (\mathbf{G}^T, \Gamma^T)$ where $\mathbf{G}^T$ is obtained by reversing the edges of $\mathbf{G}$, and if $e \in \mathbf{Edge}(\mathbf{G})$ and $\overline{e} \in \mathbf{Edge}(\mathbf{G}^T)$ is its reverse then $\Gamma^T_{\overline{e}} := (\Gamma_e)^T : V^*_{e^+} \to V^*_{e^-}$ is the linear dual[2]. Clearly, $\mathbf{P}^T$ is right resolving iff $\mathbf{P}$ is left resolving.

Fact 1.1 *The cocyclic subshifts* $X_{\mathbf{P}}$ *and* $X_{\mathbf{P}^T}$ *have the same topological entropy.*

Proof. The natural extension of $X_{\mathbf{P}}$ can be identified with a two sided subshift

$$\tilde{X}_{\mathbf{P}} := \{x \in \{1, \ldots, m\}^{\mathbb{Z}} : \ \forall_{i \in \mathbb{Z},\, k \in \mathbb{N}} \exists_{\text{a propagating path } a \text{ in } \mathbf{P}} \ [x]_{i,k} = l_{\mathbf{P}}(a)\},$$

acted upon by the shift map $\tilde{f}$, $\tilde{f}(x_i)_{i \in \mathbb{Z}} = (x_{i+1})_{i \in \mathbb{Z}}$. The natural extension of $X_{\mathbf{P}^T}$ is then identified with $\tilde{X}_{\mathbf{P}^T}$ acted upon by the inverse shift $\tilde{f}^{-1}$. The entropies of $\tilde{f}$ and $\tilde{f}^{-1}$ are the same. □

[2] Unlike in [7], we use here superscript T not the asterisk to avoid the confusion with the dual of the transfer operator.

2 Transfer operator (with an example)

Fix a colored graph with propagation $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$. For each $v \in \mathbf{Ver}(\mathbf{G})$, consider the usual projective space $\mathbb{P}(V_v)$ of all lines in V_v passing through 0 and a Banach space $\mathbf{B}(\mathbb{P}(V_v))$ of all bounded real valued functions on $\mathbb{P}(V_v)$ with the sup-norm, $\|\phi\|_{\mathbf{B}(\mathbb{P}(V_v))} := \sup_{z \in V_v} |\phi(z)|$. For a path $a \in \mathbf{Path}(\mathbf{G})$, the linear map $\Gamma_a : V_{a^+} \to V_{a^-}$ induces an operator

$$\mathcal{T}_a : \mathbf{B}(\mathbb{P}(V_{a^+})) \to \mathbf{B}(\mathbb{P}(V_{a^-}))$$

that is defined on $\phi \in \mathbf{B}(\mathbb{P}(V_{a^+}))$ by

$$(\mathcal{T}_a \phi)(z) := \begin{cases} \phi(z\Gamma_a), & \text{if } z \not\subset \ker(\Gamma_a), \\ 0, & \text{if } z \subset \ker(\Gamma_a). \end{cases} \tag{2.1}$$

Clearly, if $\ker(\Gamma_a) \neq 0$, then $\mathcal{T}_a \phi$ may have a discontinuity at $z \subset \ker(\Gamma_a)$ even if ϕ is continuous. Consider now the product space

$$\mathbf{B}_{\mathbf{P}} := \prod_{v \in \mathbf{Ver}(\mathbf{G})} \mathbf{B}(\mathbb{P}(V_v))$$

with the sup-norm $\| \cdot \|_{\mathbf{B}_{\mathbf{P}}} := \max_v \| \cdot \|_{\mathbf{B}(\mathbb{P}(V_v))}$. Given $\phi \in \mathbf{B}_{\mathbf{P}}$, we write $\phi(v, \cdot)$ for its component in $\mathbf{B}(\mathbb{P}(V_v))$.

Definition 2.3 *The transfer operator $\mathcal{T}_{\mathbf{P}} : \mathbf{B}_{\mathbf{P}} \to \mathbf{B}_{\mathbf{P}}$ of a colored graph with propagation $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ is given on $\phi \in \mathbf{B}_{\mathbf{P}}$ by*

$$(\mathcal{T}_{\mathbf{P}} \phi)(v, \cdot) := \sum_{e \in \mathbf{Edge}(\mathbf{G}):\ e^- = v} \mathcal{T}_e(\phi(e^+, \cdot)). \tag{2.2}$$

To cast $\mathcal{T}_{\mathbf{P}}$ differently, associate with $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ a colored directed graph with (uncountably many) vertices of the form (v, z) where v is a vertex of $\mathbf{G}$ and $z \in \mathbb{P}(V_v)$. Place an edge from (v, z) to $(\tilde{v}, \tilde{z})$ for any edge e joining v to $\tilde{v}$ in $\mathbf{G}$ with $z\Gamma_e = \tilde{z}$; and denote this instance by

$$(v, z) \xrightarrow[\mathbf{P}]{} (\tilde{v}, \tilde{z}).$$

With the obvious coloring inherited from $\mathbf{G}$, the subshift of this graph is equal to $X_{\mathbf{P}}$ and, for $\phi \in \mathbf{B}_{\mathbf{P}}$, $v \in \mathbf{Ver}(\mathbf{G})$ and $z \in \mathbb{P}(V_v)$,

$$(\mathcal{T}_{\mathbf{P}} \phi)(v, z) = \sum_{(v,z) \xrightarrow[\mathbf{P}]{} (\tilde{v}, \tilde{z})} \phi(\tilde{v}, \tilde{z}) \tag{2.3}$$

where the summation is over all edges emanating from (v, z). Thus $\mathcal{T}_{\mathbf{P}}$ is formally analogous to the adjacency matrix of the uncountable graph (c.f. Section 6).

For $k \in \mathbb{N}$, the k^{th}-iterate of $\mathcal{T}_{\mathbf{P}}$ on ϕ is given by

$$(\mathcal{T}_{\mathbf{P}}^k \phi)(v, \cdot) = \sum_{a \in \mathbf{Path}(\mathbf{G}):\ |a|=k,\ a^-=v} \mathcal{T}_a \phi(a^+, \cdot), \tag{2.4}$$

which is to say that $\mathcal{T}_{\mathbf{P}}^k = \mathcal{T}_{\mathbf{P}^k}$. Since only the propagating paths contribute to the above sum and the operator norm $\|\mathcal{T}_e\| \le 1$, we conclude that

$$\|\mathcal{T}_{\mathbf{P}}^k\| \le \#\{\text{propagating paths in } \mathbf{P} \text{ of length } k\}. \tag{2.5}$$

In particular, $\mathcal{T}_{\mathbf{P}}$ is bounded and has a finite spectral radius, which we denote

$$\lambda_{\mathbf{P}} := \rho(\mathcal{T}_{\mathbf{P}}).$$

Now, the topological entropy of $X_{\mathbf{P}}$ is given by

$$h_{\text{top}}(X_{\mathbf{P}}) = \lim_{n\to\infty} \frac{1}{n} \log \#\{\sigma \in \mathcal{A}^n :\ \sigma \text{ occurs in } X_{\mathbf{P}}\} \tag{2.6}$$

$$= \lim_{n\to\infty} \frac{1}{n} \log \#\{\sigma \in \mathcal{A}^n :\ \sigma \text{ is allowed}\}, \tag{2.7}$$

where (2.7) follows from a general Proposition 1.4 in [3]. If $\mathbf{P}$ is right resolving then, for a fixed vertex v and any block σ, there is at most one path a colored σ that starts at v; and combining (2.5) and (2.7) yields

$$\lambda_{\mathbf{P}} = \lim_{k\to\infty} \|\mathcal{T}_{\mathbf{P}}^k\|^{1/k} \le e^{h_{\text{top}}(X_{\mathbf{P}})}. \tag{2.8}$$

We shall prove the following theorem in the next section.

Theorem 2.1 *For a right resolving colored graph with propagation* $\mathbf{P}$*, we have*

$$\lambda_{\mathbf{P}} = e^{h_{\text{top}}(X_{\mathbf{P}})}. \tag{2.9}$$

To illustrate the theorem, let us compute $\lambda_{\mathbf{P}}$ for a concrete (non-sofic) example[3].

Example: Consider a right resolving $\mathbf{P}$ (over $\mathbb{C}$) depicted in Figure 2.1. Here $\mathbf{B_P} = \mathbf{B}(\mathbb{CP}^1) \times \mathbf{B}(\mathbb{CP}^1)$. If we identify $\mathbb{CP}^1$ with the Riemann sphere $\overline{\mathbb{C}}$ by using the standard map

$$\mathbb{C}^2 \ni (x, y) \mapsto z := y/x \in \mathbb{C} \cup \{\infty\} = \overline{\mathbb{C}},$$

then $\mathbf{B_P} = \mathbf{B}(\overline{\mathbb{C}}) \times \mathbf{B}(\overline{\mathbb{C}})$, and $(\tilde{\phi}, \tilde{\psi}) = \mathcal{T}_{\mathbf{P}}(\phi, \psi)$ iff

$$\tilde{\phi}(z) = \phi(z/2) + \begin{cases} \psi(1) & \text{if } z \ne 1, \\ 0 & \text{if } z = 1. \end{cases} \tag{2.10}$$

$$\tilde{\psi}(z) = \phi(z) + \psi(z+3). \tag{2.11}$$

[3] A very similar left-resolving example can be found in [7]; in fact, the two cocyclic subshifts coincide on their non-wandering sets and so have the same entropy.

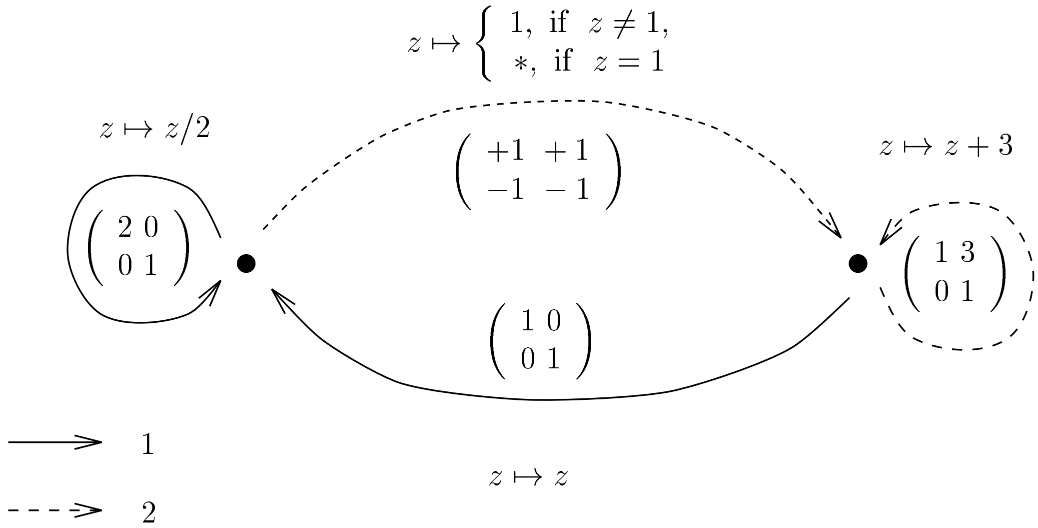

Figure 2.1: Graph with propagation of a nonsofic cocyclic subshift. (The transformations of z reflect the projective action.)

Proposition 2.1 *A complex number λ with $|\lambda| > 1$ is in the spectrum of $\mathcal{T}_{\mathbf{P}}$ iff it satisfies the equation*

$$\sum_{k=0}^{\infty} \lambda^{-2(k+1)-(4^k-1)/3} = \frac{1}{(1-\lambda^2)} - 1. \tag{2.12}$$

Moreover, any such λ is an eigenvalue of $\mathcal{T}_{\mathbf{P}}$.

Proof. For any $\lambda \in \mathbb{C}$, we can write

$$(\mathcal{T}_{\mathbf{P}} - \lambda \mathbf{Id}) \begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} D_\lambda & S \\ \mathbf{Id} & \Delta_\lambda \end{pmatrix} \begin{pmatrix} \phi \\ \psi \end{pmatrix} \tag{2.13}$$

where Δ_λ, D_λ, and S are bounded operators acting on $\phi \in \mathbf{B}(\overline{\mathbb{C}})$ as follows

$$(\Delta_\lambda \phi)(z) = \phi(z+3) - \lambda\phi(z),$$

$$(D_\lambda \phi)(z) = \phi(z/2) - \lambda\phi(z),$$

$$(S\phi)(z) = \phi(1)(\mathbf{1} - \chi_1(z)),$$

where $\mathbf{1}$ is a constant function equal to 1 and $\chi_1(z)$ is the characteristic function of the singleton $\{1\} \subset \overline{\mathbb{C}}$. Since $\mathcal{T}_{\mathbf{P}}$ is a bounded linear operator on a Banach space, $\lambda \in \mathbb{C}$ is in its resolvent iff $(\mathcal{T}_{\mathbf{P}} - \lambda \mathbf{Id})$ is one-to-one and onto, i.e. when for any $\tilde{\phi}, \tilde{\psi} \in \mathbf{B}(\overline{\mathbb{C}})$ there is a unique solution $(\phi, \psi) \in \mathbf{B}(\overline{\mathbb{C}})^2$ to the system

$$\begin{cases} D_\lambda \phi + S\psi = \tilde{\phi} \\ \phi + \Delta_\lambda \psi = \tilde{\psi} \end{cases} . \tag{2.14}$$

This happens exactly when the determinant $(D_\lambda \Delta_\lambda - S)$ is invertible, i.e. for any $\hat{\psi} \in \mathbf{B}(\overline{\mathbb{C}})$ there is a unique solution $\psi \in \mathbf{B}(\overline{\mathbb{C}})$ to

$$(D_\lambda \Delta_\lambda - S)\psi = \hat{\psi}. \tag{2.15}$$

For $|\lambda| > 1$, both D_λ and Δ_λ are invertible and so (2.15) is equivalent to

$$\psi = (D_\lambda \Delta_\lambda)^{-1}\hat{\psi} + \psi(1)(D_\lambda \Delta_\lambda)^{-1}(\mathbf{1} - \chi_1), \tag{2.16}$$

which is solvable unless

$$((D_\lambda \Delta_\lambda)^{-1}(\mathbf{1} - \chi_1))(1) = 1,$$

or, equivalently,

$$((D_\lambda \Delta_\lambda)^{-1}(\chi_1))(1) = ((D_\lambda \Delta_\lambda)^{-1}(\mathbf{1}))(1) - 1. \tag{2.17}$$

We claim that (2.12) is equivalent to (2.17). By using the Neumann power series for the resolvents, $D_\lambda^{-1} = \sum_{n=0}^{\infty} \lambda^{-n-1} D_0^n$ and $\Delta_\lambda = \sum_{n=0}^{\infty} \lambda^{-n-1} \Delta_0^n$, we see that the right sides of (2.12) and (2.17) coincide. Also, the left hand side of (2.17) equals

$$((D_\lambda \Delta_\lambda)^{-1}(\chi_1))(1) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \lambda^{-n-m-2} \Delta_0^n D_0^m(\chi_1)(1) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \lambda^{-n-m-2} \chi_1\left(\frac{1+3n}{2^m}\right).$$

which matches the left side of (2.12) since $2^m \equiv 1 \pmod 3$ iff m is even.

Thus we know that λ's with $|\lambda| > 1$ that satisfy (2.12) are in the spectrum. We leave it to the reader to verify that, for $\psi_\lambda := (D_\lambda \Delta_\lambda)^{-1}(1 - \chi_1)$ and $\phi_\lambda := -\Delta_\lambda \psi_\lambda$, $(\phi_\lambda, \psi_\lambda)$ is an eigenvector of $\mathcal{T}_\mathbf{P}$ with eigenvalue λ whenever λ satisfies (2.17). □

Fact 2.2 *Any root λ of (2.12) is a transcendental number.*

Proof. This is an immediate consequence of Theorem 2 in [9] on the transcendence of $f(z)$ for functions f given by a lacunary power series. Precisely, by taking $u_k := 4^k$, $a_k(z) := z^{6(k+1)} - 1$, that theorem implies that $\sum_{k=0}^{\infty} z^{6(k+1)+(4^k-1)}$ is transcendental at algebraic z with $0 < |z| < 1$, which readily implies the fact. □

By combining the proposition with Theorem 2.1, we conclude that the topological entropy of the example is a logarithm of a transcendental number, which precludes rationality of the dynamical zeta function.

Before leaving this section, note that the proof of the proposition hinged on the decomposition of the transfer operator $\mathcal{T}_\mathbf{P}$ into a sum of a truly infinite-dimensional part represented by

$$\begin{pmatrix} D & 0 \\ \mathbf{Id} & \Delta \end{pmatrix}$$

and a finite-dimensional (actually, one-dimensional) operator

$$\begin{pmatrix} 0 & S \\ 0 & 0 \end{pmatrix}.$$

In Section 4, we shall see that this is a key general feature of $\mathcal{T}_\mathbf{P}$.

3 Proof of Theorem 2.1

In order to prove Theorem 2.1, we shall use a presentation of $X_{\mathbf{P}}$ as a subshift of a certain countable colored graph that allows for approximation from within (in the sense of entropy) by finite graphs. Let $\mathbf{Grass}(V_v)$ be the collection of all linear subspaces of V_v. With a colored graph with propagation $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$, we associate a new colored graph $\mathbf{G_P}$ with vertex set

$$\mathbf{Ver}(\mathbf{G_P}) :=$$

$$\{(v, Y) : \ v \in \mathbf{Ver}(\mathbf{P}), \ Y \in \mathbf{Grass}(V_v) \setminus \{0\}, \text{ such that } Y = \mathrm{im}(\Gamma_a) \\ \text{for some } a \in \mathbf{Path}(\mathbf{P}) \text{ with } v = a^+\},$$

and an edge $\hat{e}$ from (v, Y) to $(\tilde{v}, \tilde{Y})$, denoted by $(v, Y) \underset{\mathbf{G_P}}{\longrightarrow} (\tilde{v}, \tilde{Y})$, for any edge e joining v to $\tilde{v}$ in $\mathbf{G}$ with $Y\Gamma_e = \tilde{Y}$. The coloring of $\mathbf{G_P}$ is that induced from $\mathbf{G}$ (i.e. $l_{\mathbf{G_P}}(\hat{e}) := l_{\mathbf{P}}(e)$). (Note that each zero length path $a = v \in \mathbf{Ver}(\mathbf{P})$ contributes (v, V_v) to $\mathbf{Ver}(\mathbf{G_P})$.) From the construction, a propagating path in $\mathbf{P}$ determines a path in $\mathbf{G_P}$ and vice versa, so that

$$X_{\mathbf{P}} = X_{\mathbf{G_P}}.$$

The following approximation lemma is closely related to Theorem 7.1 in [7].

Lemma 3.1 *Suppose that $\mathbf{P}$ is right resolving. For any $\epsilon > 0$, there exists a finite subgraph $\mathbf{G}_1$ of $\mathbf{G_P}$ such that $X_{\mathbf{G}_1} \subset X_{\mathbf{P}}$ and $h_{\mathrm{top}}(X_{\mathbf{G}_1}) \geq h_{\mathrm{top}}(X_{\mathbf{P}}) - \epsilon$.*

In fact, $\mathbf{G}_1$ will be found as the graph of a certain graph with propagation whose propagators are all linear isomorphisms obtained by taking suitable restrictions of the propagators in $\mathbf{P}$.

Proof of Lemma 3.1. Because of the right resolving hypothesis, Remark 1.1 secures a cocycle $\mathbf{\Phi} \in \mathrm{End}(V)^m$, $V = \bigoplus_{v \in \mathbf{Ver}(\mathbf{P})} V_v$, such that $X_{\mathbf{P}} = X_{\mathbf{\Phi}}$. Given a vertex v and a block σ, $x\Phi_\sigma = x\Gamma_a$ for all $x \in V_v$ taking a as the only path starting at v with coloring σ (and we suppress the natural embedding of V_v into V).

By Theorem 5.1 in [7], $X_{\mathbf{P}}$ is a union of finitely many topologically transitive components, so the entropy must be realized on a single such component. Furthermore, by Theorems 6.1 and 7.1 in [7], this component is a union of finitely many sets, each fixed by a certain iterate f^d and satisfying specification under f^d. We can pick one such set, $\Lambda_1 \subset X_{\mathbf{\Phi}}$, and then the topological entropy of f^d restricted to Λ_1 is $d \cdot h_{\mathrm{top}}(X_{\mathbf{P}})$.

As in the proof of Theorem 7.1 in [7], we consider the minimal rank,

$$r := r_{\mathbf{\Phi}} = \min\left\{\mathrm{rank}\left(\Phi_{[\zeta]_{nd}}\right) : \ \zeta \in \Lambda_1\right\}.$$

Fix $k \in d\mathbb{N}$ and $\zeta \in \Lambda_1$ so that $\mathrm{rank}(\Phi_{[\zeta]_k}) = r$. We may require that ζ is periodic and $\zeta = ([\zeta]_k)^\infty$ since one can always use specification to find ν with $\tilde{\zeta} := ([\zeta]_k \nu)^\infty \in \Lambda_1$ and replace ζ by $\tilde{\zeta}$. Set $\eta := [\zeta]_k$ and $W := \mathrm{im}(\Phi_\eta)$.

From the definition of the topological entropy, there is (arbitrarily large) $n \in d\mathbb{N}$ for which some $N \in \mathbb{N}$ with

$$\ln N/n \geq h_{\rm top}(X_{\mathbf{P}}) - \epsilon/2 \tag{3.1}$$

is a cardinality of a $(n,1)$-separated set $\{\sigma_1, ..., \sigma_N\} \subset \Lambda_1$ — i.e., the blocks $\{[\sigma_i]_n : 1 \leq i \leq N\}$ are pairwise different. By specification, there is $l \in d\mathbb{N}$ (that depends only on $\mathbf{P}$) such that one can find blocks α_i and β_i for which $\eta\alpha_i[\sigma_i]_n\beta_i\eta$ occurs in Λ_1 and has length $|\eta\alpha_i[\sigma_i]_n\beta_i\eta| = n + 2l$, $i = 1, ..., N$.

Consider the blocks $\mu_i := \alpha_i[\sigma_i]_n\beta_i\eta$, $i = 1, ..., N$. Their key property is that $V\Phi_{\mu_i} = W$ and $W\Phi_{\mu_i} = W$, which follows from the definition (minimality) of r and the fact that $\dim(W) = r$.

Claim: There is a (unique) set of vertices $I \subset \mathbf{Ver}(\mathbf{P})$ and spaces $W_v \subset V_v$, $v \in I$, such that $W = \bigoplus_{v\in I} W_v$. Moreover, for $1 \leq i \leq N$, Φ_{μ_i} permutes W_v's (i.e. for any $v \in I$ there is a unique $\tilde{v} \in I$ so that $W_v\Phi_{\mu_i} = W_{\tilde{v}}$), $\Phi_{\mu_i}|_{W_v} : W_v \to W_{\tilde{v}}$ is an isomorphism, and there is $q \in \mathbb{N}$ so that $W_v = W_v\Phi_{\mu_i^q} = V_v\Phi_{\mu_i^q}$ for any $v \in I$.

To verify the claim, fix i and set $\phi = \Phi_{\mu_i}$. By right resolving, for any $v \in \mathbf{Ver}(\mathbf{P})$, there is a unique path a with $l_{\mathbf{G}}(\mu_i) = a$ and $a^- = v$; set $\tilde{v} := a^+$ and let $g : v \mapsto \tilde{v}$ be the resulting map on the vertices. By the construction of $\mathbf{\Phi}$, $V_v\phi \subset V_{\tilde{v}}$. Since $\mathbf{Ver}(\mathbf{P})$ is finite, there is $q \in \mathbb{N}$ such that g^q is a retraction onto some $I \subset \mathbf{Ver}(\mathbf{P})$, i.e. $g^q(\mathbf{Ver}(\mathbf{P})) = I$ and $g^q|_I = \mathbf{Id}_I$. This means that $\mathrm{im}(\phi^q)$ is contained in $\bigoplus_{v\in I} V_v$ and ϕ^q restricted to $\bigoplus_{v\in I} V_v$ is a direct product $\bigoplus_{v\in I}\psi_v$ for certain $\psi_v \in \mathrm{End}(V_v)$. Thus $W = \mathrm{im}(\phi^q) = \bigoplus_{v\in I} W_v$ where $W_v := \mathrm{im}(\psi_v) = W \cap V_v$. Moreover, from the definition of g, $W_v\phi \subset W_{g(v)} = W \cap V_{g(v)}$. Actually, $\phi|_{W_v} : W_v \to W_{g(v)}$ is an isomorphism for all $v \in I$; indeed, $\psi_v := \phi^q|_{W_v} : W_v \to W_v$ is an isomorphism (due to $W\phi^q = W$) so that we have all isomorphisms in $W_v \xrightarrow{\phi} W_{g(v)} \xrightarrow{\phi} ... \xrightarrow{\phi} W_{g^{q-1}(v)} \xrightarrow{\phi} W_v$. This establishes the claim, and we can complete the proof of the lemma now.

Consider for a moment a single block μ_i and any path a with coloring μ_i. Note that $s := |a| = n + 2l - k$. Let $v_j \in \mathbf{Ver}(\mathbf{P})$, $j = 0, \ldots, s$, be the vertices visited by a, i.e. $v_0 = a^-$ and $v_j = ([a]_j)^+$ for $j > 0$. The claim guarantees that each W_v is an image of a propagator, and so we have the following edges in $\mathbf{G_P}$,

$$(v_{j-1}, W_{v_0}\Gamma_{[a]_{j-1}}) \underset{\mathbf{G_P}}{\longrightarrow} (v_j, W_{v_0}\Gamma_{[a]_j}),\ j = 1, \ldots, s. \tag{3.2}$$

Also by the claim, $W_{v_0}\Gamma_a = W_{v_0}\Phi_{\mu_i} = W_{v_s}$ so that the terminal vertex is simply (v_s, W_{v_s}); and all the linear transformations $\Gamma_{[a]_{j,1}} : W_{v_0}\Gamma_{[a]_{j-1}} \to W_{v_0}\Gamma_{[a]_j}$ are isomorphisms. Let $\mathbf{G}_1$ be the subgraph of $\mathbf{G_P}$ obtained by keeping only the edges given by (3.2) where a runs over all the paths a in $\mathbf{G_P}$ with coloring of some μ_i, $i = 1, ..., N$. We claim that $X_{\mathbf{G}_1}$ contains any *free* concatenation $\mu_{i_1}\mu_{i_2}...$ where $i_j \in \{1, \ldots, N\}$ for $j \in \mathbb{N}$. Indeed, for any $q \in \mathbb{N}$, $W\Phi_{\mu_{i_1}\mu_{i_2}...\mu_{i_q}} = W$ and so there must be a path a in $\mathbf{G}$ with coloring $\mu_{i_1}\mu_{i_2}...\mu_{i_q}$ and $\Gamma_a \neq 0$. All the edges along a are in $\mathbf{G}_1$ by the construction of $\mathbf{G}_1$.

Finally, in view (3.1), we can estimate entropy as follows

$$h_{\mathrm{top}}(X_{\mathbf{G}_1}) \geq \ln N^q/(qs) = \ln N/(n+2l-k) \geq (1-2l/n)\ln N/n \geq \tag{3.3}$$
$$(1-2l/n)h_{\mathrm{top}}(X_{\mathbf{P}}) \geq h_{\mathrm{top}}(X_{\mathbf{P}}) - \epsilon, \tag{3.4}$$

where n needs to be taken sufficiently large compared with l. □

Proof of Theorem 2.1. Fix for a moment $\epsilon > 0$. Let $\mathbf{G}_1 \subset \mathbf{G_P}$ be as in Lemma 3.1. Since $\mathbf{G}_1$ is a finite colored right resolving graph the entropy of the sofic system $X_{\mathbf{G}_1}$ is $\ln\lambda \geq h_{\mathrm{top}}(X_{\mathbf{P}}) - \epsilon$ where λ is the Perron eigenvalue of the adjacency matrix and corresponds to some non-negative eigenvector ν (see e.g. [8]). In our notation, $\nu : \mathbf{Ver}(\mathbf{G}_1) \to \mathbb{R}^+ \cup \{0\}$ with

$$\lambda\nu(v,U) = \sum_{(v,U)\xrightarrow[\mathbf{G}_1]{}(\tilde v,\tilde U)} \nu(\tilde v,\tilde U), \quad (v,U) \in \mathbf{Ver}(\mathbf{G}_1).$$

Given (v,z), $v \in \mathbf{Ver}(\mathbf{P})$ and $z \in \mathbb{P}(V_v)$, we set

$$\mu(v,z) := \max\{\nu(v,U) : \ z \subset U,\ (v,U) \in \mathbf{Ver}(\mathbf{G}_1)\},$$

where $\mu(v,z) = 0$ if the maximum is taken over an empty set. We claim that thus defined $\mu \in \mathbf{B_P}$ satisfies, for $v \in \mathbf{Ver}(\mathbf{P})$ and $z \in \mathbb{P}(V_v)$,

$$(\mathcal{T}_{\mathbf{P}}\mu)(v,z) \geq \lambda\mu(v,z). \tag{3.5}$$

If $\mu(v,z) = 0$, there is nothing to prove. Otherwise fix (v,z) and (v,U) so that $z \subset U$. From the construction of $\mathbf{G}_1$, $\Gamma_e|_U$ is an isomorphism so that $z \not\subset \ker(\Gamma_e)$ for any edge $e \in \mathbf{Edge}(\mathbf{G}_1)$ with $e^- = (v,U)$. By summing over such edges (each of which corresponds to a different edge of $\mathbf{P}$), we get

$$\sum_{(v,U)\xrightarrow[\mathbf{G}_1]{}(\tilde v,\tilde U)} \nu(\tilde v,\tilde U) \leq \sum_{(v,z)\xrightarrow[\mathbf{P}]{}(\tilde v,\tilde z)} \begin{cases} \mu(\tilde v, z\Gamma_e), & \text{if } z \not\subset \ker(\Gamma_e), \\ 0, & \text{if } z \subset \ker(\Gamma_e). \end{cases}$$

Since the right side equals $(\mathcal{T}_{\mathbf{P}}\mu)(v,z)$, we see that $(\mathcal{T}\mu)(v,z) \geq \lambda\nu(v,U)$; and (3.5) follows by the arbitrariness of U.

Applying (3.5) repeatedly yields $\mathcal{T}_{\mathbf{P}}^n\mu \geq \lambda^n\mu$ so that $\|\mathcal{T}_{\mathbf{P}}^n\mu\| \geq \lambda^n\|\mu\|$. It follows that $\lambda_{\mathbf{P}} = \rho(\mathcal{T}_{\mathbf{P}}) \geq \lambda$, which proves the theorem due to the arbitrariness of ϵ. □

4 Eigenvalue problem

The main goal of this section is to show how $\lambda_{\mathbf{P}}$ can be determined as a solution of a certain finite dimensional (albeit transcendental) eigenvalue problem and to prove that, after perhaps passing to a suitable exterior power of $\mathbf{P}$, the essential spectral radius is strictly less than $\lambda_{\mathbf{P}}$ and $\lambda_{\mathbf{P}}$ is an eigenvalue that corresponds to a positive eigenvector (for both $\mathcal{T}_{\mathbf{P}}$ and its dual $\mathcal{T}_{\mathbf{P}}^*$). Further spectral properties are discussed in Section 6.

We begin with a *rank reduction* analogous to the considerations of Section 8 in [7]. Together with $\mathbf{P}$, consider its r^{th}-exterior power, denoted $\mathbf{P}^{\wedge r}$, that (by definition) has the same underlying colored graph $\mathbf{G}^{\wedge r} := \mathbf{G}$ but the propagators are replaced by their r^{th}-exterior powers so that $\Gamma^{\wedge r} := (\Gamma_e^{\wedge r})_{e\in\mathbf{Edge_G}}$.

Definition 4.4 *A graph with propagation* $\mathbf{P}$ *is* rank reduced *iff*

$$h_{top}(X_{\mathbf{P}^{\wedge 2}}) < h_{top}(X_{\mathbf{P}}) \tag{4.1}$$

(By convention, $h_{\text{top}}(\emptyset) = -\infty$.) Note that for any $\mathbf{P}$, we have a finite filtration of cocyclic subshifts $X_1 \supset X_2 \supset \ldots \supset X_d$ where $X_r := X_{\mathbf{P}^{\wedge r}}$, $r = 1, 2, 3, \ldots, d$, $d := \max_{v\in\mathbf{Ver(G)}}\{\dim(V)\}$. For $r_0 := \max\{r \in \{1, \ldots, d\} :\ h_{\text{top}}(X_r) = h_{\text{top}}(X_1)\}$, $\mathbf{P}^{\wedge r_0}$ is rank reduced and $h_{\text{top}}(X_{\mathbf{P}}) = h_{\text{top}}(X_{\mathbf{P}^{\wedge r_0}})$. We warn however that possibly $X_{\mathbf{P}} \neq X_{\mathbf{P}^{\wedge r_0}}$. This fact does not affect our further arguments, but let us nevertheless explain. First suppose $X_{\mathbf{P}}$ is topologically transitive (e.g. $\mathbf{P}$ is irreducible — see Section 6). Then $X_{\mathbf{P}}$ is a union of a cycle of subshifts with specification (see [7]); and, therefore, $h_{\text{top}}(Y) < h_{\text{top}}(X_{\mathbf{P}})$ whenever Y is a subshift properly contained in $X_{\mathbf{P}}$. It follows that $r_0 = \max\{r :\ X_r = X_1\}$ so that $X_{\mathbf{P}} = X_{\mathbf{P}^{\wedge r_0}}$. Put differently, if we define the rank of a block σ as

$$\text{rank}_{\mathbf{P}}(\sigma) := \max\{\text{rank}(\Gamma_a) :\ \sigma = l_{\mathbf{P}}(a),\ a \in \mathbf{Path(G)}\}$$

and the eventual rank of $x \in \{1, ..., m\}^{\mathbb{N}}$ as $\text{rank}_{\mathbf{P}}(x) := \lim_{n\to\infty} \text{rank}([x]_0^n)$, then $r_0 = \min\{\text{rank}_{\mathbf{P}}(x) :\ x \in X_{\mathbf{P}}\}$. (The limit exists since $\text{rank}_{\mathbf{P}}(\sigma\eta) \leq \min\{\text{rank}_{\mathbf{P}}(\sigma), \text{rank}_{\mathbf{P}}(\eta)\}$.) In the general case, when $X_{\mathbf{P}}$ is not necessarily transitive, $X_{\mathbf{P}^{\wedge r_0}}$ is a union of the transitive components of $X_{\mathbf{P}}$ that carry the full entropy $h_{\text{top}}(X_{\mathbf{P}})$. On the other hand, by using the spectral decomposition (Section 5 in [7]), one can see that every cocyclic subshift with no *transient points* can be presented as $X_{\mathbf{P}}$ for a rank reduced and right resolving $\mathbf{P}$.

Consider pointwise ordering of the functions in $\mathbf{B_P}$, i.e. ϕ is *positive*, denoted $\phi \geq 0$, iff $\phi(v, z) \geq 0$ for all (v, z). The induced order on $\mathbf{B}_{\mathbf{P}}^*$ is such that $\mu \geq 0$ iff $\mu(\phi) \geq 0$ for all $\phi \geq 0$, $\phi \in \mathbf{B_P}$.

Theorem 4.2 *If* $\mathbf{P}$ *is a rank reduced right resolving graph with propagation then, for both* $\mathcal{T}_{\mathbf{P}}$ *and its adjoint* $\mathcal{T}_{\mathbf{P}}^*$*, the essential spectral radius is strictly less than the spectral radius* $\lambda_{\mathbf{P}}$*, and* $\lambda_{\mathbf{P}}$ *is an eigenvalue with a positive eigenvector.*

The relevance of the rank reduced hypothesis is demonstrated by a simple example: take $\#\mathcal{A} = 1$ and $\Phi_1 : \mathbb{C}^2 \to \mathbb{C}^2$ given by $\Phi_1(w_1, w_2) = (e^{2\pi i\sqrt{2}}w_1, w_2)$, then 1 is the spectral radius of $\mathcal{T}_{\mathbf{\Phi}}$ but not an isolated eigenvalue. The proof of the theorem occupies the rest of this section. The arguments for $\mathcal{T}_{\mathbf{P}}$ and $\mathcal{T}_{\mathbf{P}}^*$ follow the same path.

We fix a rank reduced right resolving graph $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$. Consider the power $\mathbf{P}^n$ of $\mathbf{P}$ and partition the set of edges of $\mathbf{G}^n$ (i.e. n-paths in $\mathbf{G}$) according to the rank of their propagator:

$$\mathbf{Edge}_-(\mathbf{G}^n) := \{a \in \mathbf{Edge}(\mathbf{G}^n) : \ \mathrm{rank}(\Gamma_a) > 1\},$$

$$\mathbf{Edge}_+(\mathbf{G}^n) := \{a \in \mathbf{Edge}(\mathbf{G}^n) : \ \mathrm{rank}(\Gamma_a) = 1\}.$$

By allowing only the paths through the edges in $\mathbf{Edge}_-(\mathbf{G}^n)$, we obtain

$$X_-^{(n)} := \{(l(e_i))_{i=1}^{\infty} \in X_{\mathbf{P}} : \ (e_{in+1}, ..., e_{in+n}) \in \mathbf{Edge}_-(\mathbf{G}^n) \text{ for all } i \geq 0\}.$$

Note that $f^n(X_-^{(n)}) = X_-^{(n)}$.

Lemma 4.2 *If h_n is the topological entropy of the restriction $f^n|_{X_-^{(n)}}$, then*

$$\lim_{n\to\infty} \frac{1}{n} h_n = h_{top}(X_{\mathbf{P}^{\wedge 2}}) \tag{4.2}$$

Proof. Consider $Y_-^{(n)} := X_-^{(n)} \cup \ldots \cup f^{n-1} X_-^{(n)}$ so that $f(Y_-^{(n)}) = Y_-^{(n)}$ and $\frac{1}{n} h_n = h_{\mathrm{top}}(Y_-^{(n)})$. If $\tilde{n}|n$ then $X_-^{(n)} \subset X_{\tilde{n}}$ and $Y_-^{(n)} \subset Y_{\tilde{n}}$. Therefore, $Y_-^{(\infty)} := \bigcap_{n\in\mathbb{N}} Y_-^{(n)}$ is an nonempty subshift, and $h_{\mathrm{top}}(Y_-^{(\infty)}) = \lim_{n\to\infty} h_{\mathrm{top}}(Y_-^{(n)})$. (Indeed, $h_{\mathrm{top}}(Y_-^{(\infty)}) \leq \liminf_{n\to\infty} h_{\mathrm{top}}(Y_-^{(n)})$ is clear; and $h_{\mathrm{top}}(Y_-^{(\infty)}) \geq \limsup_{n\to\infty} h_{\mathrm{top}}(Y_-^{(n)})$ by upper semi-continuity of entropy — c.f. Proposition 4.4.6, p 122, in [8].) It is left to show that $h_{\mathrm{top}}(Y_-^{(\infty)}) = h_{\mathrm{top}}(X_{\mathbf{P}^{\wedge 2}})$. One inequality follows from $X_{\mathbf{P}^{\wedge 2}} \subset Y_-^{(\infty)}$. We claim that $Y_-^{(\infty)} \subset \bigcup_{q\geq 0} f^{-q}(X_{\mathbf{P}^{\wedge 2}})$, which already assures $\bigcap_{q\geq 0} f^q(Y_-^{(\infty)})) \subset X_{\mathbf{P}^{\wedge 2}}$ and yields the other inequality: $h_{\mathrm{top}}(Y_-^{(\infty)}) = h_{\mathrm{top}}(\bigcap_{q\geq 0} f^q(Y_-^{(\infty)})) \leq h_{\mathrm{top}}(X_{\mathbf{P}^{\wedge 2}})$. To show the claim, consider $x \in Y_-^{(\infty)}$. For each $n \in \mathbb{N}$, $x \in Y_-^{(n)}$ implies that $\mathrm{rank}_{\mathbf{P}}([x]_{q_n}) > 1$ and $\mathrm{rank}_{\mathbf{P}}([x]_{q_n,n}) > 1$ for some $q_n \in \mathbb{N}$. If $q_* := \sup_{n\in\mathbb{N}} q_n = \infty$, then it is not hard to see that $\mathrm{rank}_{\mathbf{P}}(x) > 1$ and so $x \in X_{\mathbf{P}^{\wedge 2}}$. If $q_* < \infty$, then there is $q \geq q_*$ such that $\mathrm{rank}_{\mathbf{P}}(f^q(x)) > 1$ and so $f^q(x) \in X_{\mathbf{P}^{\wedge 2}}$. $\square$

Because $\mathbf{P}$ is rank reduced, the lemma assures that there is $n_0 \in \mathbb{N}$ such that we have

$$\frac{1}{n} h_n < h_{\mathrm{top}}(X_{\mathbf{P}}) \qquad \text{for all } n \geq n_0. \tag{4.3}$$

Consider now a fixed $n \geq n_0$. Let $\mathbf{P}_-^n$ and $\mathbf{P}_+^n$ be the colored graphs with propagation obtained by removing from $\mathbf{P}^n$ all the edges in $\mathbf{Edge}_+(\mathbf{G}^n)$ and $\mathbf{Edge}_-(\mathbf{G}^n)$, respectively, so that $\mathbf{Edge}(\mathbf{P}_+^n) = \mathbf{Edge}_+(\mathbf{G}^n)$ and $\mathbf{Edge}(\mathbf{P}_-^n) = \mathbf{Edge}_-(\mathbf{G}^n)$. For brevity, denote the transfer operators of $\mathbf{P}$, $\mathbf{P}_-^n$, and $\mathbf{P}_+^n$ by $\mathcal{T}$, $\mathcal{T}_-^n$, and $\mathcal{T}_+^n$, respectively. Observe that they all act on the same space $\mathbf{B}_{\mathbf{P}} = \mathbf{B}_{\mathbf{P}^n} = \mathbf{B}_{\mathbf{P}_-^n} = \mathbf{B}_{\mathbf{P}_+^n}$ and the construction assures that

$$\mathcal{T}^n = \mathcal{T}_-^n + \mathcal{T}_+^n. \tag{4.4}$$

Here is the key feature of the above splitting:

Fact 4.3 *The operator* $\mathcal{T}_+^n$ *and its dual* $\mathcal{T}_+^{n*}$ *have finite-dimensional images.*

Proof. Indeed, if $a \in \mathbf{Edge}_+(\mathbf{G}^n)$ then $\mathrm{im}(\Gamma_a)$ is a single line in V_{a^+} and we set

$$K := \{(v, z) : z = \mathrm{im}(\Gamma_a), v = a^+, a \in \mathbf{Edge}(\mathbf{P}_+^n)\}.$$

We claim that the image $\mathrm{im}(\mathcal{T}_+^n)$ is spanned by $\#K$ functions $\{\phi_{(v,z)}\}_{(v,z)\in K}$ given by

$$\phi_{(v,z)}(\tilde{v}, \tilde{z}) := \sum_{(\tilde{v},\tilde{z}) \xrightarrow[\mathbf{P}_+^n]{} (v,z)} 1, \quad \tilde{z} \in V_{\tilde{v}}.$$

Indeed, for $\phi \in \mathbf{B_P}$, we see that (c.f. (2.3))

$$\mathcal{T}_+^n \phi(\tilde{v}, \tilde{z}) = \sum_{(\tilde{v},\tilde{z}) \xrightarrow[\mathbf{P}_+^n]{} (v,z)} \phi(v, z) = \sum_{(v,z)\in K} \phi(v, z)\phi_{(v,z)}(\tilde{v}, \tilde{z}). \tag{4.5}$$

Similarly, $\mathrm{im}(\mathcal{T}_+^{n*})$ is contained in the linear span of the evaluation functionals $\{\delta_{(v,z)}\}_{(v,z)\in K}$ defined by $\delta_{(v,z)}(\phi) := \phi(v, z)$ for $\phi \in \mathbf{B_P}$. Indeed, for $\mu \in \mathbf{B_P^*}$ and $\phi \in \mathbf{B_P}$, via (4.5), we obtain

$$\mathcal{T}_+^{n*}\mu(\phi) = \mu(\mathcal{T}_+^n\phi) = \mu\left(\sum_{(v,z)\in K} \phi(v, z)\phi_{(v,z)}\right) = \sum_{(v,z)\in K} \mu\left(\phi_{(v,z)}\right)\phi(v, z), \tag{4.6}$$

which shows that $\mathcal{T}_+^{n*}\mu = \sum_{(v,z)\in K} \mu\left(\phi_{(v,z)}\right)\delta_{(v,z)}$. $\square$

Now, observe that $X_{\mathbf{P}_-^n}$ as a subshift of $(\mathcal{A}^n)^{\mathbb{N}}$ is naturally conjugated with $f^n : X_-^{(n)} \to X_-^{(n)}$. Thus (4.3) and Theorem 2.1 yield

$$\lambda_-^{(n)} := \rho(\mathcal{T}_-^n)^{1/n} = \rho(\mathcal{T}_-^{n*})^{1/n} = e^{\frac{1}{n}h_n} < e^{h_{\mathrm{top}}(X_\mathbf{P})} = \lambda_\mathbf{P}. \tag{4.7}$$

In this way, for $|\lambda| > \lambda_-^{(n)}$, we have at our disposal the Neumann series

$$(\mathcal{T}_-^n - \lambda^n)^{-1} = \sum_{k=0}^{\infty} -\lambda^{-n(k+1)}\mathcal{T}_-^{nk}, \qquad (\mathcal{T}_-^{n*} - \lambda^n)^{-1} = \sum_{k=0}^{\infty} -\lambda^{-n(k+1)}(\mathcal{T}_-^{n*})^k. \tag{4.8}$$

This enables us to solve the eigenvalue problem for $\mathcal{T}$ via the following simple lemma, which also holds after replacing $\mathcal{T}$, $\mathcal{T}_+^n$, $\mathcal{T}_-^n$ by $\mathcal{T}^*$, $\mathcal{T}_+^{n*}$, $\mathcal{T}_-^{n*}$, respectively.

Lemma 4.3 *Suppose that* $|\lambda| > \lambda_-^{(n)}$*. Given* $\beta \in \mathbf{B_P}$*, the solutions* $\phi \in \mathbf{B_P}$ *of*

$$\mathcal{T}^n\phi = \lambda^n\phi + \beta \tag{4.9}$$

are in bijective correspondence with the solutions $\alpha \in \mathrm{im}(\mathcal{T}_+^n)$ *of*

$$\alpha = \mathcal{T}_+^n(\lambda^n - \mathcal{T}_-^n)^{-1}(\alpha - \beta). \tag{4.10}$$

Moreover, if α *and* ϕ *are the solutions of their respective equations, then*

$$\alpha = \mathcal{T}_+^n\phi \qquad \text{and} \qquad \phi = (\lambda^n - \mathcal{T}_-^n)^{-1}(\alpha - \beta). \tag{4.11}$$

Proof of Lemma 4.3. Suppose (4.10) holds. To satisfy $\mathcal{T}_+^n\phi = \alpha$, set $\phi := (\lambda^n - \mathcal{T}_-^n)^{-1}(\alpha - \beta)$; and then apply $(\lambda^n - \mathcal{T}_-^n)$ to both sides (of the definition) to obtain $(\lambda^n - \mathcal{T}_-^n)\phi = \alpha - \beta = \mathcal{T}_+^n\phi - \beta$, which rearranges into (4.9). Likewise, (4.9) is equivalent to $\mathcal{T}_+^n\phi - \beta = (\lambda^n - \mathcal{T}_-^n)\phi$. If it holds, substituting $\alpha := \mathcal{T}_+^n\phi$ and applying $(\lambda^n - \mathcal{T}_-^n)^{-1}$ to both sides yields the second equation in (4.11), which further becomes (4.10) after application of $\mathcal{T}_+^n$ to both sides. □

The lemma motivates the following definitions.

Definition 4.5 *Suppose that* $\mathbf{P}$ *is a rank reduced colored graph with propagation and* $n \geq n_0$ *(as in (4.3)). The small transfer operator of* $\mathbf{P}$ *(associated to* n*) is*

$$\mathcal{Q}_n(\lambda) := \mathcal{T}_+^n(\lambda^n - \mathcal{T}_-^n)^{-1},$$

where λ *is a complex parameter with* $|\lambda| \geq \lambda_-^{(n)}$*. Likewise, the small adjoint transfer operator of* $\mathbf{P}$ *is*

$$\mathcal{P}_n(\lambda) := \mathcal{T}_+^{n*}(\lambda^n - \mathcal{T}_-^{n*})^{-1}.$$

Proposition 4.2 *For* $n \geq n_0$*, a complex number* λ *with* $|\lambda| > \lambda_-^{(n)}$ *is in the spectrum of* $\mathcal{T}^n$ *(or* $\mathcal{T}^{*n}$*) iff*

$$\det\left(\mathbf{Id} - \mathcal{Q}_n(\lambda)\right) = 0, \tag{4.12}$$

or, equivalently, iff

$$\det\left(\mathbf{Id} - \mathcal{P}_n(\lambda)\right) = 0. \tag{4.13}$$

Every such λ *is an eigenvalue and the eigenvectors can be found from (4.11) by solving (4.10) (with* $\beta = 0$*).*

Note that both $\mathcal{Q}_n(\lambda)$ and $\mathcal{P}_n(\lambda)$ are finite-dimensional and so are the determinants in (4.12) and (4.13).

Proof of Proposition 4.2. By the Banach open mapping theorem, λ^n is in the resolvent of $\mathcal{T}_{\mathbf{P}}^n$ iff (4.9) has a unique solution for any $\beta \in \mathbf{B}_{\mathbf{P}}$. By the lemma, this is exactly when (4.10), i.e. $(\mathbf{Id} - \mathcal{Q}_n(\lambda))\alpha = -\mathcal{Q}_n(\lambda)\beta$ is uniquely solvable for $\alpha \in \mathrm{im}(\mathcal{T}_+^n) = \mathrm{im}(\mathcal{Q}_n(\lambda))$. Thus λ is in the resolvent iff $\mathbf{Id} - \mathcal{Q}_n(\lambda)$ is non-singular on $\mathrm{im}(\mathcal{Q}_n(\lambda))$. A similar argument (based on the analogue of the lemma for $\mathcal{T}^*$) holds for $\mathcal{P}_n(\lambda)$. □

Proposition 4.2 establishes the assertion of Theorem 4.2 regarding the essential spectral radii. It may be worth to also note that Proposition 4.2 enables one to recover all eigenvalues λ of $\mathcal{T}$ satisfying $|\lambda| > \lambda_-^{(n_0)}$. This rests on the following general observation.

Lemma 4.4 *Let* T *and* S *be two operators and* $A^n := \{z \in \mathbb{C} : r^n \leq |z| \leq R^n\}$ *for some* $r, R \geq 0$*. If* $\sigma(T) \cap A^1$ *and* $\sigma(S) \cap A^1$ *are finite and there is* $n_0 \in \mathbb{N}$ *such that* $\sigma(T^n) \cap A^n = \sigma(S^n) \cap A^n$ *for all* $n \geq n_0$*, then* $\sigma(T) \cap A^1 = \sigma(S) \cap A^1$*.*

Proof. Clearly, if λ is an eigenvalue of T then λ^n is an eigenvalue of T^n; and if μ is an eigenvalue of T^n, then one of its n^{th}-roots is an eigenvalue of T. Likewise for S. Therefore, we are dealing with an elementary property of the circle group: of concern is only the argument of the eigenvalues and it suffices to argue for $r = R > 0$. Fix $\lambda \in \sigma(T) \cap A^1$. For $\mu \in \sigma(S) \cap A^1$, let $I_\mu := \{n \geq n_0 : \ \mu^n = \lambda^n\}$ and let d_μ be the greatest common divisor of I_μ. Since $\bigcup_{\mu \in \sigma(S) \cap A^1} I_\mu = \{n \geq n_0\}$ and $\mu \in \sigma(S) \cap A^1$ is finite, there exists μ with $I_\mu \neq \emptyset$ and $d_\mu = 1$. It follows that the set of equations for z, $z^n = \lambda^n$, $n \in I_\mu$, has a unique solution $z = \lambda$. Thus $\lambda = \mu \in \sigma(S) \cap A^1$, which shows $\sigma(T) \cap A^1 \subset \sigma(S) \cap A^1$. Swapping T and S yields the opposite inclusion. □

To finish the proof of Theorem 4.2, it remains to obtain a positive eigenvector corresponding to $\lambda_{\mathbf{P}}$.

Proposition 4.3 *There are positive $\alpha_{\mathbf{P}} \in \mathbf{B_P}$ and $\xi_{\mathbf{P}} \in \mathbf{B}^*_{\mathbf{P}}$ such that $\mathcal{Q}_n(\lambda_{\mathbf{P}})\alpha_{\mathbf{P}} = \alpha_{\mathbf{P}}$ and $\mathcal{P}_n(\lambda_{\mathbf{P}})\xi_{\mathbf{P}} = \xi_{\mathbf{P}}$.*

Before proving the proposition, we observe that it combines with Theorem 2.1 and Proposition 4.2 to yield the following:

Corollary 4.1 (entropy computation) *For a rank reduced graph with propagation $\mathbf{P}$, the entropy $h_{\text{top}}(X_{\mathbf{P}})$ is the logarithm of the largest positive root λ of the equation (4.12) (or (4.13).*

Both $\mathbf{B_P}$ and $\mathbf{B}^*_{\mathbf{P}}$ are Banach lattices (see e.g. [12]); and the proof of Proposition 4.3 will depend on the following basic properties. For any positive bounded linear operator L on $\mathbf{B_P}$ (or $\mathbf{B}^*_{\mathbf{P}}$) and any ϕ_1, ϕ_2, we have: $|L(\phi_1)| \leq L(|\phi_1|)$, $\|\phi_1\| \leq \| \, |\phi_1| \, \|$, and $0 \leq \phi_1 \leq \phi_2$ implies $\|\phi_1\| \leq \|\phi_2\|$. Moreover, $\mathcal{Q}_n$ and $\mathcal{P}_n$ are positive operators for any $\lambda > 0$ as it is clear from their representations as series of positive operators obtained from (4.8):

$$\mathcal{Q}_n(\lambda) = \sum_{k=0}^{\infty} \lambda^{-n(k+1)} \mathcal{T}^n_+ (\mathcal{T}^n_-)^k \quad \text{and} \quad \mathcal{P}_n(\lambda) = \sum_{k=0}^{\infty} \lambda^{-n(k+1)} \mathcal{T}^{n*}_+ (\mathcal{T}^{n*}_-)^k. \tag{4.14}$$

Fact 4.4 *(i) If $\lambda_2 \geq \lambda_1 \geq \lambda_-^{(n)}$, then $\mathcal{Q}_n(\lambda_2) \leq \mathcal{Q}_n(\lambda_1)$ (i.e. $\mathcal{Q}_n(\lambda_2)\phi \leq \mathcal{Q}_n(\lambda_1)\phi$ for $\phi \geq 0$). In particular, $\rho(\mathcal{Q}_n(\lambda_2)) \leq \rho(\mathcal{Q}_n(\lambda_1))$.*

(ii) for $|\lambda| \geq \lambda_-^{(n)}$, $|\mathcal{Q}_n(\lambda)\phi| \leq \mathcal{Q}_n(|\lambda|)|\phi|$. In particular, $\rho(\mathcal{Q}_n(\lambda)) \leq \rho(\mathcal{Q}_n(|\lambda|))$.

Proof. (i) That $\mathcal{Q}_n(\lambda_2) \leq \mathcal{Q}_n(\lambda_1)$ follows immediately from the power expansion. Also, for arbitrary $\phi \in \mathbf{B}$ and $j \in \mathbb{N}$, $|\mathcal{Q}_n(\lambda_2)^j(\phi)| \leq \mathcal{Q}_n(\lambda_2)^j(|\phi|) \leq \mathcal{Q}_n(\lambda_1)^j(|\phi|)$ and so $\|\mathcal{Q}_n(\lambda_2)^j(\phi)\| \leq \|\mathcal{Q}_n(\lambda_1)^j(|\phi|)\|$, which yields the inequality between the spectral radii.

(ii) For any fixed k, $L := \mathcal{T}^n_+(\mathcal{T}^n_-)^k$ is positive and so $|L(\phi)| \leq L(|\phi|)$. Thus the inequality $|\mathcal{Q}_n(\lambda)\phi| \leq \mathcal{Q}_n(|\lambda|)|\phi|$ follows immediately by taking absolute value of the the power expansion (4.14). Again, $|\mathcal{Q}_n(\lambda)^j\phi| \leq \mathcal{Q}_n(|\lambda|)^j|\phi|$ so that that $\|\mathcal{Q}_n(\lambda)^j(\phi)\| \leq \|\mathcal{Q}_n(|\lambda|)^j(|\phi|)\|$, which yields the inequality between the spectral radii. □

Proof of Proposition 4.3. First let us show that $\rho(\mathcal{Q}_n(\lambda_{\mathbf{P}})) = 1$. By the definition of $\lambda_{\mathbf{P}}$ as the spectral radius and by Corollary 4.1,

$$\lambda_{\mathbf{P}} = \sup\{|\lambda| : 1 \in \text{ spectrum}(\mathcal{Q}_n(\lambda))\}. \tag{4.15}$$

Therefore, if $1 \in$ spectrum$(\mathcal{Q}_n(\lambda))$, then $\lambda_{\mathbf{P}} \geq |\lambda|$; and Fact 4.4 yields $\rho(\mathcal{Q}_n(\lambda_{\mathbf{P}})) \geq \rho(\mathcal{Q}_n(|\lambda|)) \geq \rho(\mathcal{Q}_n(\lambda)) \geq 1$. In particular, $\rho(\mathcal{Q}_n(\lambda_{\mathbf{P}})) \geq 1$. Now, observe that $\rho(\mathcal{Q}_n(\lambda_{\mathbf{P}})) > 1$ would imply existence of $\lambda_* > \lambda_{\mathbf{P}}$ with $1 \in \sigma(\mathcal{Q}_n(\lambda_*))$ thus contradicting (4.15). Indeed, as a function of $\lambda \in [\lambda_{\mathbf{P}}, \infty)$, $\rho(\mathcal{Q}_n(\lambda))$ is continuous (since $\mathcal{Q}_n(\lambda)$ is analytic in λ), non-increasing (due to (i) of Fact 4.4), and diminishes to zero as $\lambda \to \infty$ (since (4.14) yields $\mathcal{Q}_n(\lambda) \leq 1/(\lambda - C)$ for some $C > 0$ and large $\lambda > 0$).

Thus we have shown that $\rho(\mathcal{Q}_n(\lambda_{\mathbf{P}})) = 1$; and the same argument yields $\rho(\mathcal{P}_n(\lambda_{\mathbf{P}})) = 1$. Now, $\mathcal{Q}_n(\lambda_{\mathbf{P}})$ and $\mathcal{P}_n(\lambda_{\mathbf{P}})$ are positive and compact (actually finite-dimensional), and $\mathbf{B}$ and $\mathbf{B}^*$ are ordered real Banach spaces with a total positive cone (i.e. positive vectors span the whole space); hence, the Krein-Rutman theorem (p 265 in [12]) yields positive eigenvectors $\alpha_{\mathbf{P}}$ and $\xi_{\mathbf{P}}$ with $\mathcal{Q}_n(\lambda_{\mathbf{P}})(\alpha_{\mathbf{P}}) = \alpha_{\mathbf{P}}$ and $\mathcal{P}_n(\lambda_{\mathbf{P}})(\xi_{\mathbf{P}}) = \xi_{\mathbf{P}}$. □

Finally, we assemble our findings into a proof of Theorem 4.2.

Proof of Theorem 4.2. If $\alpha_{\mathbf{P}}$ is as in the Proposition 4.3, the function

$$\phi := (\lambda^n - \mathcal{T}_-^n)^{-1}(\alpha_{\mathbf{P}}), \tag{4.16}$$

satisfies the eigenvalue equation $\mathcal{T}^n\phi = \lambda^n\phi$ by Lemma 4.3. In this way,

$$\psi_{\mathbf{P}} := (\phi + \lambda^{-1}\mathcal{T}\phi + ... + \lambda^{-n+1}\mathcal{T}^{n-1}\phi) \tag{4.17}$$

is a positive eigenvector of $\mathcal{T}$,

$$\mathcal{T}\psi_{\mathbf{P}} = \lambda_{\mathbf{P}}\psi_{\mathbf{P}}.$$

That the essential spectral radius of $\mathcal{T}$ does not exceed $\lambda_-^{(n)} < \lambda_{\mathbf{P}}$ is guaranteed by Proposition 4.2. The arguments for $\mathcal{T}^*$ are completely analogous. □

5 Measure of maximal entropy

A topologically transitive $X_{\mathbf{P}}$ has a unique measure of maximal entropy (see [7]). In this section, we shall give a concrete description of this measure in terms of the transfer operator. The construction does not require transitivity of $X_{\mathbf{P}}$, but we have to make the following standing hypothesis:

(A) $\mathbf{P}$ is a right-resolving and left-resolving rank reduced colored graph with propagation.

We note that the requirement that $\mathbf{P}$ is both left and right resolving is natural already in the sofic context (i.e. for colored graphs without propagation) and is dictated by the fact that we shall apply the results of the previous section to the transfer operator of both $\mathbf{P}$ and $\mathbf{P}^T$. By Fact 1.1 and Theorem 2.1, $\lambda_{\mathbf{P}^T} = \lambda_{\mathbf{P}}$ and throughout this section we set

$$\lambda := \lambda_{\mathbf{P}} = \lambda_{\mathbf{P}^T}.$$

Also, given a right resolving $\mathbf{P}$, it is easy to construct $\mathbf{P}_1$ that is both right and left resolving and $X_{\mathbf{P}}$ is a finite-to-one factor of $X_{\mathbf{P}_1}$, which guarantees that the maximal entropy measure of $X_{\mathbf{P}_1}$ pushed forward to the maximal entropy measure of $X_{\mathbf{P}}$. (For instance, $\mathbf{P}_1$ can be taken as $\mathbf{P}$ with a richer labeling that assigns to an edge e the label $(e_+, l_{\mathbf{P}}(e))$; or one can use Remark 1.1.)

Let us fix eigenvectors ϕ and ϕ^T of $\mathcal{T}_{\mathbf{P}}$ and $\mathcal{T}_{\mathbf{P}^T}$, respectively, corresponding to the eigenvalue λ (as supplied by Theorem 4.2). For a block σ with $\operatorname{rank}_{\mathbf{P}}(\sigma) = 1$ (and thus also $\operatorname{rank}_{\mathbf{P}^T}(\sigma) = 1$), *the weight of* σ is defined by

$$\nu_{\mathbf{P}}(\sigma) := \sum_a \lambda^{-|\sigma|} \phi(a^+, \operatorname{im}(\Gamma_a)) \phi^T(a^-, \operatorname{im}(\Gamma_a^T)), \tag{5.1}$$

where the summation extends over all propagating paths a in $\mathbf{P}$ with $l_{\mathbf{P}}(a) = \sigma$. For convenience, we also set $\nu_{\mathbf{P}}(\sigma) = 0$ if $\operatorname{rank}_{\mathbf{P}}(\sigma) = 0$ (i.e. σ is dissalowed). In the "one vertex case", when the cocyclic subshift is presented as $X_{\mathbf{\Phi}}$, the formula (5.1) becomes a perfect analogue of *the Parry weight* (0.1) discussed in the introduction:

$$\nu_{\mathbf{\Phi}}(\sigma) = \lambda^{-|\sigma|} \phi(\operatorname{im}(\Phi_\sigma)) \phi^T(\operatorname{im}(\Phi_\sigma^T)). \tag{5.2}$$

Precisely, given a colored right and left resolving graph $\mathbf{G}$, one naturarly takes Φ_i to be the adjacency of the subgraph with the edges of color i, and then $X_{\mathbf{G}} = X_{\mathbf{\Phi}}$ and the two formulas coincide. We also comment that the restriction to the rank one blocks only in (5.1) is offset by the fact that such blocks dominate: $\#\{\sigma \in \mathcal{A}^n : \operatorname{rank}_{\mathbf{P}}(\sigma) > 1\} / \#\{\sigma \in \mathcal{A}^n : \operatorname{rank}_{\mathbf{P}}(\sigma) = 1\}$ converges to zero exponentially in n because $X_{\mathbf{P}^{\wedge 2}} = \{x \in \mathcal{A}^{\mathbb{N}} : \operatorname{rank}_{\mathbf{P}}(x) > 1\}$ is a cocyclic subshift of lower entropy than $X_{\mathbf{P}}$ (see Definition 4.4).

Theorem 5.3 *Under hypothesis (A), there exists a unique shift invariant Borel probability measure μ on $X_{\mathbf{P}}$ that is proportional to the weights $\nu_{\mathbf{P}}$, i.e. for some constant $C > 0$, if $\operatorname{rank}_{\mathbf{P}}(\sigma) = 1$ then $\mu(U_\sigma) = C\nu_{\mathbf{P}}(\sigma)$ (where $U_\sigma = \{x : [x]_{|\sigma|} = \sigma\}$). The entropy of μ under the shift map is $\ln \lambda_{\mathbf{P}} = h_{top}(X_{\mathbf{P}})$.*

The proof of the theorem will proceed through a number of lemmas. We start with the analog of the Markov condition.

Lemma 5.5 *If* $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$*, then*

$$\nu_{\mathbf{P}}(\sigma) = \sum_{i=1}^{m} \nu_{\mathbf{P}}(\sigma i) \qquad \textit{and} \qquad \nu_{\mathbf{P}}(\sigma) = \sum_{i=1}^{m} \nu_{\mathbf{P}}(i\sigma).$$

Proof. We prove only the first equality as the proof of the second one is completely analogous. The right side equals

$$R = \sum_{i=1}^{m} \sum_{a \in A_i} \lambda^{-|\sigma|-1} \phi(a^+, \mathrm{im}(\Gamma_a)) \phi^T(a^-, \mathrm{im}(\Gamma_a^T)),$$

where $A_i := \{a : \ \Gamma_a \neq 0, l(a) = \sigma i\}$. Any $a \in A_i$ has the form $\tilde{a}e$ where $l(\tilde{a}) = \sigma$, $l(e) = i$, and $\tilde{a} \in \tilde{A} := \{\tilde{a} : \Gamma_{\tilde{a}} \neq 0, \ l(\tilde{a}) = \sigma\}$. Moreover, for a and $\tilde{a}$ as above, $a^- = \tilde{a}^-$, and $\mathrm{im}(\Gamma_a^T) = \mathrm{im}(\Gamma_{\tilde{a}}^T)$ because $\mathrm{im}(\Gamma_a^T) = \mathrm{im}(\Gamma_e^T)\Gamma_{\tilde{a}}^T \subset \mathrm{im}(\Gamma_{\tilde{a}}^T)$ and the dimension of both sides is 1 from $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$. Therefore,

$$R = \sum_{\tilde{a} \in \tilde{A}} \sum_{i} \lambda^{-|\sigma|-1} \phi(\tilde{a}e_i^+, \mathrm{im}(\Gamma_{\tilde{a}e_i})) \phi^T(\tilde{a}^-, \mathrm{im}(\Gamma_{\tilde{a}}^T)),$$

where the second sum is over i for which there is (a unique, by the right resolving) edge e_i colored i with $\tilde{a}^+ = e^-$. Finally, by using that ϕ is an eigenvector of $\mathcal{T}_{\mathbf{P}}$, we obtain

$$\begin{aligned} R = \lambda^{-|\sigma|-1} \sum_{\tilde{a} \in \tilde{A}} \phi^T(\tilde{a}^-, \mathrm{im}(\Gamma_{\tilde{a}}^T)) \sum_{i} \phi(e_i^+, \mathrm{im}(\Gamma_{\tilde{a}})\Gamma_{e_i}) = \\ \lambda^{-|\sigma|-1} \sum_{\tilde{a} \in \tilde{A}} \phi^T(\tilde{a}^-, \mathrm{im}(\Gamma_{\tilde{a}}^T)) \ \mathcal{T}_{\mathbf{P}} \phi(\tilde{a}^+, \mathrm{im}(\Gamma_{\tilde{a}})) = \\ \lambda^{-|\sigma|} \sum_{\tilde{a} \in \tilde{A}} \phi^T(\tilde{a}^-, \mathrm{im}(\Gamma_{\tilde{a}}^T)) \phi(\tilde{a}^+, \mathrm{im}(\Gamma_{\tilde{a}})). \end{aligned}$$

The last sum equals $\nu_{\mathbf{P}}(\sigma)$, which finishes the proof. □

Corollary 5.2 *If* $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$ *and* $U_\sigma = \bigcup_{k=1}^{N} U_{\sigma_k}$ *where the sum is disjoint, then*

$$\nu_{\mathbf{P}}(\sigma) = \sum_{k=1}^{N} \nu_{\mathbf{P}}(\sigma_k).$$

Proof. Since $\mathrm{rank}_{\mathbf{P}}(\sigma_k) = 1$, this amounts to repeated application of the lemma. □ Set, for $n \in \mathbb{N}$,

$$\nu_n := \sum_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1,\ |\sigma|=n} \nu_{\mathbf{P}}(\sigma).$$

In the Markov case ν_n is constant. Here, we introduce

$$\nu_\infty := \lim_n \nu_n < \infty$$

where the limit exists on the force of the following lemma.

Lemma 5.6 *For $n > 1$, we have*

(i) $\nu_n \geq \nu_{n-1}$;

(ii) $\nu_n \leq \sum_{v\in\mathbf{Ver}(\mathbf{G})} \dim(V_v)\ \|\phi\|\|\phi^T\|$.

Proof. (i) By factoring $\sigma = \eta i$ where $i \in \mathcal{A}$, we can use Lemma 5.5 to obtain

$$\begin{aligned}\nu_n &= \sum_{\sigma=\eta i:\ \mathrm{rank}_{\mathbf{P}}(\eta)=1} \nu_{\mathbf{P}}(\eta i) + \sum_{\sigma=\eta i:\ \mathrm{rank}_{\mathbf{P}}(\eta)>1} \nu_{\mathbf{P}}(\eta i) &\quad (5.3)\\ &= \nu_{n-1} + \sum_{\sigma=\eta i:\ \mathrm{rank}_{\mathbf{P}}(\eta)>1} \nu_{\mathbf{P}}(\eta i) \geq \nu_{n-1}. &\quad (5.4)\end{aligned}$$

(ii) Pick a basis B_v of V_v for each $v \in \mathbf{Ver}(\mathbf{G})$. We unravel the definitions of $\nu_{\mathbf{P}}(\sigma)$ and ν_n and then estimate:

$$\begin{aligned}\nu_n = \lambda^{-n} &\sum_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1,\ |\sigma|=n}\ \sum_{a:\ l(a)=\sigma,\ \Gamma_a\neq 0} \phi(a^+, \mathrm{im}(\Gamma_a))\ \phi^T(a^-, \mathrm{im}(\Gamma_a^T)) \leq \\ &\lambda^{-n}\|\phi^T\| \sum_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1,\ |\sigma|=n}\ \sum_{a:\ l(a)=\sigma,\ \Gamma_a\neq 0} \phi(a^+, \mathrm{im}(\Gamma_a)) \leq \\ &\lambda^{-n}\|\phi^T\| \sum_{\sigma:\ |\sigma|=n}\ \sum_{a:\ l(a)=\sigma,\ \Gamma_a\neq 0}\ \sum_{z\in B_{a_-}:\ z\Gamma_a\neq 0} \phi(a^+, z\Gamma_a) = \\ &\lambda^{-n}\|\phi^T\| \sum_{v\in\mathbf{Ver}(\mathbf{G})}\ \sum_{z\in B_v}\ \sum_{a:\ a^-=v,\ |a|=n,\ z\Gamma_a\neq 0} \phi(a^+, z\Gamma_a) = \\ &\lambda^{-n}\|\phi^T\| \sum_{v\in\mathbf{Ver}(\mathbf{G})}\ \sum_{z\in B_v} \mathcal{T}_{\mathbf{P}}^n\phi(v,z) = \\ &\|\phi^T\| \sum_{v\in\mathbf{Ver}(\mathbf{G})}\ \sum_{z\in B_v} \phi(v,z) \leq \\ &\|\phi^T\| \sum_{v\in\mathbf{Ver}(\mathbf{G})}\ \sum_{z\in B_v} \|\phi\|.\end{aligned}$$

□

Consider the family $\mathcal{R}_0$ of all finite unions of the cylinders to which we assigned weights, that is

$$\mathcal{R}_0 := \left\{ \bigcup_{k=1}^{N} U_{\sigma_k} :\ \mathrm{rank}_{\mathbf{P}}(\sigma_k) = 1,\ k = 1, \ldots, N,\ N \in \mathbb{N} \right\} \cup \{\emptyset\}.$$

It is easy to see that $\mathcal{R}_0$ is a ring (i.e. is closed under differences and finite sums).

Corollary 5.3 *There exists a unique finite and (finitely) additive measure μ_0 on $\mathcal{R}_0$ such that if* $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$ *then* $\mu_0(U_\sigma) = \nu_{\mathbf{P}}(\sigma)$.

Proof. Set $\mu_0(\emptyset) = 0$. Any set in $\mathcal{R}_0$ can be represented as a disjoint union of cylinders in $\mathcal{R}_0$; set then

$$\mu_0\left(\bigcup_{k=1}^{N} U_{\sigma_k}\right) := \sum_{k=1}^{N} \nu_{\mathbf{P}}(\sigma_k).$$

By Corollary 5.5, this definition does not depend on the decomposition. The additivity is manifest. From Lemma 5.6, $\mu(A) \leq \nu_\infty < +\infty$ for any $A \in \mathcal{R}_0$. $\square$

To extend μ_0 to μ_1 defined on the ring of all finite unions of cylinders,

$$\mathcal{R}_1 := \left\{\bigcup_{k=1}^{N} U_{\sigma_k} : \; N \in \mathbb{N}\right\} \cup \{\emptyset\},$$

we mimic the usual construction of *inner measure* and set, for $B \in \mathcal{R}_1$,

$$\mu_1(B) := \sup\{\mu_0(A) : \; A \subset B, \; A \in \mathcal{R}_0\}.$$

Lemma 5.7 *There is a unique (countably additive) Borel measure μ_2 such that $\mu_2(B) = \mu_1(B)$ for $B \in \mathcal{R}_1$. Moreover, μ_2 is concentrated on $X_{\mathbf{P}}$ and is finite with $\mu_2(X_{\mathbf{P}}) = \mu\left(\bigcup_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1} U_\sigma\right) = \nu_\infty$.*

Proof. The measure μ_2 is supplied by the standard theorem about unique extension of a measure from a ring to a σ-ring (see e.g. [2]). We only have to verify that μ_1 is finite and countably additive on $\mathcal{R}_1$. To check countable additivity, let $B = \bigcup_{n=1}^{\infty} B_n \in \mathcal{R}_1$ where B_n are pairwise disjoint and $B_n \in \mathcal{R}_1$. Since the sets of $\mathcal{R}_1$ are open and compact, $B = \bigcup_{n=1}^{N} B_n \in \mathcal{R}_1$ for some $N \in \mathbb{N}$. Suppose that $A \subset B$ and $A \in \mathcal{R}_0$. Observe that the intersection of A with any cylinder is in $\mathcal{R}_0$ so that $A_n := A \cap B_n \in \mathcal{R}_0$. Hence, from the definition of μ_1 and additivity of μ_0, $\mu_0(A) = \sum_{n=1}^{N} \mu_0(A_n) \leq \sum_{n=1}^{N} \mu_1(B_n)$, which yields $\mu_1(B) \leq \sum_{n=1}^{N} \mu_1(B_n)$ by arbitrariness of A. The opposite inequality is equally clear since if $A_n \subset B_n$ and $A_n \in \mathcal{R}_0$, then $\bigcup_{n=1}^{N} A_n \in \mathcal{R}_0$ and $\bigcup_{n=1}^{N} A_n \subset B$.

To see that μ_1 is concentrated on $X_{\mathbf{P}}$, consider σ that is not allowed. Then, for any η, $\sigma\eta$ is also not allowed and so $\nu_{\mathbf{P}}(\sigma\eta) = 0$ by definition. It follows that $\mu_0(A) = 0$ for any $A \in \mathcal{R}_0$ with $A \subset U_\sigma$, and so $\mu_1(U_\sigma) = 0$ by the definition of μ_1.

Finally, since any $A \in \mathcal{R}_0$ is a disjoint union of cylinders U_σ with constant length $|\sigma| = n$ for some $n \in \mathbb{N}$, we have $\mu_0(A) \leq \nu_n \leq \nu_\infty$. On the other hand, by the countable additivity of μ_2,

$$\mu_2\left(\bigcup_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1} U_\sigma\right) = \lim_{n\to\infty} \mu_2\left(\bigcup_{\sigma:\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1,\ |\sigma|=n} U_\sigma\right) = \lim_{n\to\infty} \nu_n = \nu_\infty.$$

$\square$

From Lemma 5.7, there is a probability measure μ on $X_{\mathbf{P}}$ given on any Borel measurable set B by

$$\mu(B) := \mu_2(B)/\nu_\infty. \tag{5.5}$$

In view of Theorem 2.1, the following lemma identifies μ as the measure of maximal entropy.

Lemma 5.8 *The measure μ is shift invariant and*

$$h_\mu(f) = \ln \lambda,$$

where $h_\mu(f)$ denotes the measure theoretic entropy of μ under the shift map.

Proof. Lemma 5.5 yields, for $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$,

$$\mu_0(f^{-1}U_\sigma) = \mu_0\left(\bigcup_{i=1}^m U_{i\sigma}\right) = \mu_0(U_\sigma).$$

Also, whenever $A \in \mathcal{R}_0$ then $f^{-1}(A) \in \mathcal{R}_0$, so that the equality above implies $\mu_0(A) = \mu_0(f^{-1}A)$. As a consequence, $\mu_1(B) \leq \mu_1(f^{-1}B)$ by the definition of μ_1. The reverse inequality follows from passing to the complements (which are also in $\mathcal{R}_1$). Thus μ_1 is invariant and so is its extension μ_2.

The partition of the full shift space $\mathcal{A}^{\mathbb{N}}$ into cylinders U_i, $i = 1, ..., m$, is a generator (c.f. [4]); therefore,

$$h_\mu(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{\sigma:\ |\sigma|=n} -\mu(U_\sigma) \ln \mu(U_\sigma).$$

For $\mathrm{rank}_{\mathbf{P}}(\sigma) = 1$, we can use the definition of $\nu_{\mathbf{P}}(\sigma)$ and that there are no more than $\#\mathbf{Ver}(\mathbf{G})$ paths a with $l(a) = \sigma$ (by right resolving) to obtain a bound

$$\mu(U_\sigma)\nu_\infty = \nu_{\mathbf{P}}(\sigma) \leq \lambda^{-|\sigma|}\|\phi\|\|\phi^T\|\#\mathbf{Ver}(\mathbf{G}),$$

which we use to estimate:

$$\begin{aligned}
\frac{1}{n}\sum_{\sigma:\ |\sigma|=n} -\mu(U_\sigma)\ln\mu(U_\sigma) &\geq \\
\frac{1}{n}\sum_{\sigma:\ |\sigma|=n,\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1} -\mu(U_\sigma)\ln\mu(U_\sigma) &\geq \\
\sum_{\sigma:\ |\sigma|=n,\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1} -\ln\left(\|\phi\|\|\phi^T\|\#\mathbf{Ver}(\mathbf{G})\lambda^{-n}\nu_\infty^{-1}\right)\mu(U_\sigma) &\geq \\
\left(\lambda + \ln(\|\phi\|\|\phi^T\|\#\mathbf{Ver}(\mathbf{G}))/n\right)\sum_{\sigma:\ |\sigma|=n,\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1}\mu(U_\sigma).
\end{aligned}$$

Via Lemma 5.7, $\sum_{\sigma:\ |\sigma|=n,\ \mathrm{rank}_{\mathbf{P}}(\sigma)=1}\mu(U_\sigma) = \nu_n/\nu_\infty \to 1$, so it follows that $h_\mu(f) \geq \lambda = h_{\mathrm{top}}(X_{\mathbf{P}})$. Of course, $h_\mu(f) \leq h_{\mathrm{top}}(X_{\mathbf{P}})$, as for any invariant measure. □

Conclusion of the proof of Theorem 5.3. The union of Lemma 5.7, (5.5), and Lemma 5.8 constitutes the theorem. □

6 Further properties

We concentrate in this section on topologically transitive cocyclic subshifts. Our goal is two-fold. First, we want to present $X_{\mathbf{P}}$ as a subshift of a certain countable colored graph $\hat{\mathbf{G}}_{\mathbf{P}}$ that is irreducible and recurrent — which makes $X_{\mathbf{P}}$ *a coded system* (as studied in [3]). Second, we want to supply the missing piece of the Perron-Frobenius theory to the effect that $\lambda_{\mathbf{P}}$ is a simple eigenvalue for *irreducible* $\mathbf{P}$ and the only eigenvalue on the circle $|\lambda| = \lambda_{\mathbf{P}}$ for *primitive* (irreducible and aperiodic) $\mathbf{P}$. Here irreducibility and primitivity of $\mathbf{P}$ correspond to transitivity and mixing of $X_{\mathbf{P}}$, a connection we barely touch on since it was studied thoroughly in [7] in the context of cocycles $\mathbf{\Phi}$.

Definition 6.6 *For a rank reduced* $\mathbf{P}$*, the transition graph of* $\mathbf{P}$ *is defined as the colored graph* $\hat{\mathbf{G}}_{\mathbf{P}}$ *whose vertices are of the form* (v, z) *where* $v \in \mathbf{Ver}(\mathbf{P})$ *and* $z = \mathrm{im}(\Gamma_a)$ *for some* $a \in \mathbf{Path}(\mathbf{P})$ *with* $a^+ = v$ *and* $\mathrm{rank}(\Gamma_a) = 1$*; and there is an edge, call it* $\hat{e}$*, joining* (v, z) *to* $(\tilde{v}, \tilde{z})$ *for each edge* $e \in \mathbf{Edge}(\mathbf{P})$ *such that* $e^- = v$*,* $e^+ = \tilde{v}$ *and* $z\Gamma_e = \tilde{z}$*. The coloring of* $\mathbf{P}$ *is that induced from* $\mathbf{P}$*, that is* $l(\hat{e}) := l(e)$*.*

Note that $\hat{\mathbf{G}}_{\mathbf{P}}$ is a subgraph of the graph $\mathbf{G}_{\mathbf{P}}$ introduced in Section 3. In particular, if $\mathbf{P}$ is right (left) resolving then so is $\hat{\mathbf{G}}_{\mathbf{P}}$. Also, if (v, z) is a vertex of $\hat{\mathbf{G}}_{\mathbf{P}}$, $e \in \mathbf{Edge}(\mathbf{P})$ with $v = e^-$, $\tilde{v} = e^+$, and $\tilde{z} = z\Gamma_e$, then $(\tilde{v}, \tilde{z})$ is automatically a vertex of $\hat{\mathbf{G}}_{\mathbf{P}}$. Hence, for any $\phi \in \mathbf{B}_{\mathbf{P}}$ and $(v, z) \in \mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}})$, we have

$$(\mathcal{T}_{\mathbf{P}}\phi)(v, z) = \sum_{(v,z) \xrightarrow[\hat{\mathbf{G}}_{\mathbf{P}}]{} (\tilde{v},\tilde{z})} \phi(\tilde{v}, \tilde{z}) \tag{6.6}$$

This is to say that if $A_{\mathbf{P}}$ is the countable adjacency matrix of $\hat{\mathbf{G}}_{\mathbf{P}}$ acting on the space $\mathbf{B}(\mathbf{Ver}(\mathbf{P}))$ of bounded functions on $\mathbf{Ver}(\mathbf{P})$ and $R : \mathbf{B}_{\mathbf{P}} \to \mathbf{B}(\mathbf{Ver}(\mathbf{P}))$ is the natural restriction operator, then

$$R \circ \mathcal{T}_{\mathbf{P}} = A_{\mathbf{P}} \circ R. \tag{6.7}$$

Furthermore, $\mathcal{T}_{\mathbf{P}}^*$ leaves invariant the closed linear subspace $D_{\mathbf{P}}$ of $\mathbf{B}_{\mathbf{P}}^*$ spanned by the evaluation functionals $\delta_{(v,z)}$ where $(v, z) \in \mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}})$. Indeed, consider $\eta \in D_{\mathbf{P}}$ of the form $\eta = \sum_{(v,z)} w(v, z)\delta_{(v,z)}$ where $w \in l^1(\mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}}))$, then one computes that $\mathcal{T}_{\mathbf{P}}^*\eta = \sum_{(v,z)} u(v, z)\delta_{(v,z)}$ where

$$u(\tilde{v}, \tilde{z}) = \sum_{(v,z) \xrightarrow[\hat{\mathbf{G}}_{\mathbf{P}}]{} (\tilde{v},\tilde{z})} w(v, z).$$

The right hand side above corresponds to the action of the transposed matrix $A_{\mathbf{P}}^T$, i.e., upon identifying $D_{\mathbf{P}}$ with $l^1(\mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}}))$, we obtain

$$\mathcal{T}_{\mathbf{P}}^*|_{D_{\mathbf{P}}} = A_{\mathbf{P}}^T. \tag{6.8}$$

(Because the in- and out-degrees of the vertices of $\hat{\mathbf{G}}_\mathbf{P}$ are uniformly bounded by $m = \#\mathcal{A}$, $A_\mathbf{P} : \mathbf{B}(\mathbf{Ver}(\mathbf{P})) \to \mathbf{B}(\mathbf{Ver}(\mathbf{P}))$ and $A^T_\mathbf{P} : l^1(\mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P})) \to l^1(\mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P}))$ are bounded linear operators; and $(A^T_\mathbf{P})^* = A_\mathbf{P}$.)

The results of Section 4 on the discrete nature of the outer spectrum of $\mathcal{T}_\mathbf{P}$ and $\mathcal{T}^*_\mathbf{P}$ enable one to recover that part of the spectrum from $A_\mathbf{P}$.

Fact 6.5 *Suppose that $n_0 \in \mathbb{N}$ and $\lambda^{(n)}_- \geq 0$ are as in (4.3) and (4.7), $n \geq n_0$, and $|\lambda| > \lambda^{(n)}_-$.*

(i) For $\phi \in \mathbf{B}_\mathbf{P}$, if $\mathcal{T}^n_\mathbf{P}\phi = \lambda^n\phi$, then $A^n(R\phi) = \lambda^n R\phi$; and $\phi \neq 0$ if $R(\phi) \neq 0$.

(ii) For $\eta \in \mathbf{B}^*_\mathbf{P}$, if $\mathcal{T}^{*n}_\mathbf{P}\eta = \lambda^n\eta$, then $\eta \in D_\mathbf{P}$ and $(A^T)^n\eta = \lambda^n\eta$.

Proof. (i) The first implication follows from (6.7). Now, suppose that $R\phi = 0$, i.e. $\phi(v, z) = 0$ for all $(v, z) \in \mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P})$. Recalling the definition of $\mathcal{T}^n_+$, we see that $\mathcal{T}^n_+\phi$ is obtained by evaluating (and summing) $\phi(v, z)$ over certain vertices $(v, z) \in \mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P})$; therefore, $\alpha := \mathcal{T}^n_+\phi = 0$ and we conclude that $\phi = 0$ by Lemma 4.3.
(ii) By the analogue of Lemma 4.3 for $\mathcal{T}^*_\mathbf{P}$, we see that $\eta = (\lambda^n - \mathcal{T}^{n*}_-)^{-1}\alpha$ for some $\alpha \in \mathrm{im}(\mathcal{T}^{n*}_+)$; and $\mathrm{im}(\mathcal{T}^{n*}_+) \subset D_\mathbf{P}$ from (the proof of) Fact 4.3. That $\eta \in D_\mathbf{P}$ follows from $(\lambda^n - \mathcal{T}^{n*}_-)^{-1}D_\mathbf{P} \subset D_\mathbf{P}$, which can be seen by using the Neumann series (4.8) for $(\lambda^n - \mathcal{T}^{n*}_-)^{-1}$. □

Corollary 6.4 *The parts of the spectra of $A_\mathbf{P}$ and $\mathcal{T}_\mathbf{P}$ in $\{\lambda : |\lambda| > \lambda^{(n_0)}_-\}$ coincide and consist exclusively of eigenvalues.*

Proof. On the force of Lemma 4.4, it suffices to show that the corresponding parts of the spectra coincide for all powers $A^n_\mathbf{P}$ and $\mathcal{T}^n_\mathbf{P}$ where $n \geq n_0$. Fix $n \geq n_0$ and consider λ such that $|\lambda| > \lambda^{(n_0)}_-$. Suppose that $\lambda^n \in \sigma(A^n_\mathbf{P})$. Then $\lambda^n \in \sigma((A^T_\mathbf{P})^n)$ (since $(A^T_\mathbf{P})^* = A_\mathbf{P}$) and (6.8) yields $\lambda^n \in \sigma((\mathcal{T}^*_\mathbf{P})^n) = \sigma(\mathcal{T}^n_\mathbf{P})$. On the other hand, any such $\lambda^n \in \sigma(\mathcal{T}^n_\mathbf{P})$ is an eigenvalue by Proposition 4.2. Hence, by (i) of Fact 6.5, $\lambda^n \in \sigma(A^n_\mathbf{P})$, and λ^n is an eigenvalue of $A^n_\mathbf{P}$. □

Recall that a graph $\mathbf{G}$ is *irreducible* iff every vertex connects to any other vertex via a path; and $\mathbf{G}$ is *primitive* (also called *irreducible aperiodic*) if all the power graphs $\mathbf{G}^n$, $n \in \mathbb{N}$, are irreducible.

Definition 6.7 *A colored graph with propagation $\mathbf{P} = (\mathbf{G}, \mathbf{\Gamma})$ with $\mathbf{\Gamma} \neq 0$ is irreducible iff the graph $\mathbf{G}$ is irreducible and, for any two vertices $v, \tilde{v} \in \mathbf{G}$, the linear hull*

$$\mathrm{lin}\{\Gamma_a : a^- = v, a^+ = \tilde{v}, a \in \mathbf{Path}(\mathbf{P})\} = \mathbf{L}(V_v, V_{\tilde{v}}) \tag{6.9}$$

where $\mathbf{L}(V_v, V_{\tilde{v}})$ stands for all the linear operators $V_v \to V_{\tilde{v}}$. Also, $\mathbf{P}$ is called primitive if all its powers $\mathbf{P}^n$, $n \in \mathbb{N}$, are irreducible.

Observe that Theorem 8.1 from [7] assures that every topologically transitive cocyclic subshift can be presented by an irreducible rank reduced one-vertex $\mathbf{P}$ with the extra property that $\mathbf{P}$ is primitive if the subshift is mixing.

Proposition 6.4 *If $\mathbf{P}$ is an irreducible rank reduced colored graph with propagation, then*

(i) $X_{\hat{\mathbf{G}}_\mathbf{P}} = X_\mathbf{P}$*;*

(ii) $\hat{\mathbf{G}}_\mathbf{P}$ *is an irreducible graph and* $X_\mathbf{P}$ *is transitive;*

(iii) if $\mathbf{P}$ *is primitive, then* $\hat{\mathbf{G}}_\mathbf{P}$ *is primitive.*

(iv) $\hat{\mathbf{G}}_\mathbf{P}$ *is positively recurrent.*

Proof. (ii) Consider $(v, z), (\tilde{v}, \tilde{z}) \in \mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P})$ where $z = \mathrm{im}(\Gamma_a)$ and $\tilde{z} = \mathrm{im}(\Gamma_b)$ for some paths a and b with $a^+ = v$ and $b^+ = \tilde{v}$. Since $\Gamma_a, \Gamma_b \neq 0$, there is $C \in \mathbf{L}(V_v, V_{\tilde{v}})$ such that $\Gamma_a C \Gamma_b \neq 0$. By irreducibility of $\mathbf{P}$, there is then a path c (from v to $\tilde{v}$) such that $\Gamma_a \Gamma_c \Gamma_b \neq 0$ (c.f. Lemma 4.1 in [7]). In particular, $z\Gamma_c = \tilde{z}$, which means that there is a path in $\hat{\mathbf{G}}_\mathbf{P}$ from (v, z) to $(\tilde{v}, \tilde{z})$. In a similar way, one shows that, given two occurring blocks σ and η, there is γ such that $\sigma\gamma\eta$ occurs in $X_\mathbf{P}$, which readily implies transitivity of $X_\mathbf{P}$.

(i) If σ is a block that is a coloring of some path in $\hat{\mathbf{G}}_\mathbf{P}$, then it is a coloring of the corresponding path in $\mathbf{P}$, so $X_{\hat{\mathbf{G}}_\mathbf{P}} \subset X_\mathbf{P}$. Vice versa, if σ is a coloring of a path a in $\mathbf{P}$, then to see that σ is a coloring of some path in $\hat{\mathbf{G}}_\mathbf{P}$ it suffices to find a path b in $\mathbf{P}$ such that $\mathrm{rank}(\Gamma_b) = 1$, $b^+ = a^-$, and $\Gamma_b \Gamma_a \neq 0$. Because $\mathbf{\Gamma} \neq 0$ and $\mathbf{P}$ is irreducible, $X_\mathbf{P} \neq \emptyset$. Being also rank reduced, $X_\mathbf{P}$ contains then a path $\tilde{b}$ with $\mathrm{rank}(\Gamma_{\tilde{b}}) = 1$; and $\tilde{b}$ can be extended (as in the proof of (ii)) to get a suitable b.

(iii) First note that $\hat{\mathbf{G}}_\mathbf{P}^j = \hat{\mathbf{G}}_{\mathbf{P}^j}$. (Here, that every vertex of $\hat{\mathbf{G}}_\mathbf{P}^j$ is a vertex of $\hat{\mathbf{G}}_{\mathbf{P}^j}$ hinges on extension of paths via irreducibility of $\mathbf{P}$.) Thus, if the powers $\mathbf{P}^j$, $j \in \mathbb{N}$, are all irreducible so are $\hat{\mathbf{G}}_\mathbf{P}^j$ by (ii). Hence, $\hat{\mathbf{G}}_\mathbf{P}$ is primitive.

(iv) Since $\hat{\mathbf{G}}_\mathbf{P}$ is irreducible by (ii), its *Perron value* can be defined as the growth rate of the entries of the matrices $A_\mathbf{P}^j$ or $(A_\mathbf{P}^T)^j$; namely, after fixing a vertex $i = (v, z)$, $\lambda_{\hat{\mathbf{G}}_\mathbf{P}} := \limsup_{j\to\infty} \left((A_\mathbf{P}^j)_{ii}\right)^{1/j} = \limsup_{j\to\infty} \left((A_\mathbf{P}^T)_{ii}^j\right)^{1/j}$ (see page 223 in [6]), where the subscript ii indicates taking the appropriate entry of the matrix. In view of (6.8), $(A_\mathbf{P}^T)_{ii}^j = ((\mathcal{T}_\mathbf{P}^*)^j \delta_i)(I_i)$ where $\delta_i \in \mathbf{B}_\mathbf{P}^*$ is the evaluation functional $\delta_i(\phi) := \phi(v, z)$ and $I_i \in \mathbf{B}_\mathbf{P}$ is the indicator function of the singleton $\{(v, z)\}$. Therefore, $\lambda_{\hat{\mathbf{G}}_\mathbf{P}} \leq \rho(\mathcal{T}_\mathbf{P}^*) = \lambda_\mathbf{P}$. Also, $\lambda_{\hat{\mathbf{G}}_\mathbf{P}} \geq \lambda_\mathbf{P}$ because $\lambda_\mathbf{P}$, being an eigenvalue of $\mathcal{T}_\mathbf{P}$, is an eigenvalue of $A_\mathbf{P}$ by Corollary 6.4.

Having shown that $\lambda_{\hat{\mathbf{G}}_\mathbf{P}} = \lambda_\mathbf{P}$, positive recurrence of $\hat{\mathbf{G}}_\mathbf{P}$ will follow from Lemma 7.1.38 in [6], if we find positive $x \in \mathbf{B}(\mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P}))$ and $y \in l^1(\mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P}))$ such that $A_\mathbf{P} x = \lambda_\mathbf{P} x$ and $A_\mathbf{P}^T y = \lambda_\mathbf{P} y$ (since $x \cdot y < \infty$ is automatic). Let ϕ be a positive eigenvector of $\mathcal{T}_\mathbf{P}$ corresponding to $\lambda_\mathbf{P}$ as secured by Theorem 4.2. By (6.7), $x := R\phi \in \mathbf{B}(\mathbf{Ver}(\hat{\mathbf{G}}_\mathbf{P}))$ satisfies $A_\mathbf{P} x = \lambda_\mathbf{P} x$; and $x \neq 0$ by (i) of Fact 6.5.

Likewise, if η is a positive eigenvector of $\mathcal{T}^*_{\mathbf{P}}$ corresponding to $\lambda_{\mathbf{P}}$, then (ii) of Fact 6.5 secures $\eta \in D_{\mathbf{P}} = l^1(\mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}}))$ and $(A^T)^n = \lambda^n_{\mathbf{P}}\eta$ for some $n \in \mathbb{N}$. Thus $y := \eta + \ldots \lambda_{\mathbf{P}}^{-n+1}(A^T)^{n-1}\eta \in l^1(\mathbf{Ver}(\hat{\mathbf{G}}_{\mathbf{P}}))$ is the sought after eigenvector of $A^T_{\mathbf{P}}$. $\square$

We are ready to exploit the connection between $\mathcal{T}_{\mathbf{P}}$ and $A_{\mathbf{P}}$ to show the following result.

Theorem 6.4 *If* $\mathbf{P}$ *is an irreducible rank reduced colored graph with propagation, then the Perron eigenvalue* $\lambda_{\mathbf{P}}$ *is simple (for both* $\mathcal{T}_{\mathbf{P}}$ *and* $\mathcal{T}^*_{\mathbf{P}}$*) and the corresponding eigenvector is strictly positive. Moreover, if* $\mathbf{P}$ *is primitive, then* $\lambda_{\mathbf{P}}$ *is the only point of the spectrum on the circle* $|\lambda| = \lambda_{\mathbf{P}}$*.*

Proof. If $\mathbf{P}$ is irreducible and rank reduced, then $\hat{\mathbf{G}}_{\mathbf{P}}$ is irreducible and positively recurrent by Proposition 6.4. By the Perron-Frobenius theorem for countable Markov chains (Theorem 7.1.3 in [6]), $A_{\mathbf{P}}$ and $A^T_{\mathbf{G}}$ have one dimensional space of eigenvalues corresponding to the Perron value $\lambda_{\hat{\mathbf{G}}_{\mathbf{P}}} = \lambda_{\mathbf{P}}$. By Fact 6.5, the same holds for $\mathcal{T}_{\mathbf{P}}$ and $\mathcal{T}^*_{\mathbf{P}}$, i.e. $\lambda_{\mathbf{P}}$ is simple.

If $\mathbf{P}$ is also primitive, then so is $\hat{\mathbf{G}}_{\mathbf{P}}$ by Proposition 6.4. By (f) of Theorem 7.1.3 in [6], $A_{\mathbf{P}}$ has no eigenvalues of modulus $\lambda_{\mathbf{P}}$ besides $\lambda_{\mathbf{P}}$. By (i) of Fact 6.5, no such eigenvalues exist for $\mathcal{T}_{\mathbf{P}}$. We could consider only eigenvalues due to Proposition 4.2. $\square$

Finally let us comment on the connection with the *coded systems.* This is easiest done in the context of the presentation via cocycles (as adapted in the introduction). Consider then a transitive cocyclic subshift X. By [7], one can present X as $X_{\mathbf{\Phi}}$ for some irreducible cocycle of minimal rank one. The associated graph with propagation $\mathbf{P}$ has only one vertex v_0 and is irreducible and right (and left) resolving. Thus $\hat{\mathbf{G}}_{\mathbf{P}}$ is an irreducible and right (and left) resolving colored graph. Such graphs correspond to *irreducible Fischer automata*, which makes $X = X_{\hat{\mathbf{G}}_{\mathbf{P}}}$ a *coded system* (as defined in [3]). Moreover, if σ is a block with $\mathrm{rank}_{\mathbf{P}}(\Phi_\sigma) = 1$, then one easily checks that σ is the so called synchronizing word (see e.g. Section 3 in [3], or page 3 in [1]). Thus X is a *synchronized coded system.*

References

[1] A. Bertrand. Spécification et synchronisation dans les systémes dynamiques symboliques ou expansifs. preprint, ?

[2] P. Billingsley. *Probability and Measure.* John Wiley & Sons, 1979.

[3] F. Blanchard and G. Hansel. Systemes Codes. *Theoretical Computer Science*, 44:17–50, 1986.

[4] M. Denker, C. Grillenberger, and K. Sigmund. *Ergodic Theory on Compact Spaces*, volume 527 of *Lecture Notes in Mathematics.* Springer-Verlag, New York, 1976.

[5] R. Fisher. Sofic systems and graphs. *Monatshefte fur Mathematik*, 80:179–186, 1975.

[6] B. Kitchens. *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts.* Universitext. Springer-Verlag, Berlin, 1998.

[7] J. Kwapisz. Cocyclic Subshifts. *Math. Z.*, 234 no. 2:255–290, 2000.

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

[9] J.H. Loxton and A.J. van der Poorten. Arithmetic properties of certain functions in several variables III. *Bull. Austral. Math. Soc.*, 16:15–47, 1977.

[10] K. Petersen. Chains, entropy, coding. *Erg. Th. and Dyn. Sys.*, 6:415–448, 1986.

[11] D. Rudolph. A mixing Markov chain with exponentially decaying return time is finitary Bernoulli. *Erg. Th. and Dyn. Sys.*, 2:85–97, 1982.

[12] H.H. Schaefer. *Topological Vector Spaces*, volume 3 of *Graduate texts in mathematics.* Springer-Verlag, New York, 1971.

[13] P. Walters. *An Introduction to Ergodic Theory*, volume 79 of *Graduate texts in mathematics.* Springer-Verlag, New York, 1982.

[14] B. Weiss. Subshifts of finite type and sofic systems. *Monatshefte fur Mathematik*, 77:462–474, 1973.