A COMBINATORIAL APPROACH TO STUDY SUBSHIFTS ASSOCIATED WITH MULTIGRAPHS

NIKITA AGARWAL, HARITHA CHERIYATH, AND SHARVARI NEETIN TIKEKAR

Abstract. A subshift of finite type over finitely many symbols can be described as a collection of all infinite walks on a digraph with at most a single edge from a vertex to another. The associated finite set $F$ of forbidden words is a constraint which determines the language of the shift entirely. In this paper, in order to describe infinite walks on a multigraph, we introduce the notion of multiplicity of a word (finite walk) and define repeated words as those having multiplicity at least 2. In general, for given collections $F$ of forbidden words and $R$ of repeated words with pre-assigned multiplicities, we define notion of a generalized language which is a multiset. We obtain a subshift associated with $F$ and $R$ such that its entropy is calculated using the generalized language. We also study the relationship between the language of this subshift and the generalized language. We then obtain a combinatorial expression for the generating function that enumerates the number of words of fixed length in this generalized language. This gives the Perron root and eigenvectors of the adjacency matrix with integer entries associated to the underlying multigraph. Using this, the topological entropy and an alternate definition of Parry measure for the associated edge shift are obtained. We also discuss some properties of Markov measures on this subshift.

Keywords: Subshift of finite type, edge shift, Perron root and eigenvectors, Markov measure, correlation polynomial, generating function

2020 Mathematics Subject Classification: 37B10 (Primary); 68R15, 05A16 (Secondary)

1. Introduction

Symbolic dynamics is mainly used as a tool to study dynamical systems that exhibit some kind of hyperbolicity such as Anosov diffeomorphisms or Axiom A maps [15, 25, 29, 31]. It involves studying the collection of infinite sequences on a given set of symbols known as a shift space and is quite combinatorial in nature. A shift space which is characterized in terms of a finite collection of words that are forbidden, is popularly known as a subshift of finite type. For instance, the collection of all infinite paths on a given digraph describes a subshift of finite type. The complexity of this subshift is given by the connectivity of this graph, which is therefore, given by the spectral properties of the associated adjacency matrix. When a graph does not have multiple edges from a vertex to another, any path can be represented in terms of its vertices as well as its edges. Such shifts are, in particular, known as the vertex and edge shifts, respectively. Shift spaces commonly appear, albeit with different terminologies, in other areas of sciences as well. For instance, in coding and information theory, the finite collection of words that are forbidden is called a constraint and the corresponding subshift is known as the constrained coding, refer [10, 11, 18, 19].

A directed graph containing multiple edges between some pair of initial and terminal vertices is known as a multigraph and its associated adjacency matrix has non-negative integer entries. In this case, two distinct paths may have identical representation in terms of vertices. Hence, to avoid such a discrepancy, distinct labels are assigned to
each edge. Each path is then described in terms of labelled edges and associated edge shifts are considered. Some dynamical properties of the edge shifts are studied in [22, 33]. For instance, Williams [33], using the concept of Markov partitions, models a class of dynamical systems as edge shifts associated with non-negative integer matrices via conjugacy. They also provide a complete characterization of these systems in terms of their adjacency matrices. Although any edge shift can be thought of as a vertex shift with respect to some other digraph, it is more convenient and useful to consider edge shift since it makes the associated adjacency matrices closed under multiplication. It also often reduces the size of the adjacency matrix significantly, but at the expense of adding multiple edges. Hence the study of edge shifts is important to understand several other complex dynamical systems.

An important invariant of shift spaces under conjugacy is topological entropy, which measures the growth rate of the number of words of finite length in the shift, and can be computed for a wide class of shifts. Computing entropy also gives rise to some interesting combinatorial problems, as it involves counting the number of finite words. However, there is not much literature on counting the finite words in an edge shift, especially when the edge shift is associated to some multigraph. This constitutes one of the problems that we aim to address in this work. Recent studies have proved that multigraphs provide a more suitable and efficient framework to model real life systems, including financial risk models [34] and software architecture [8]. A simple example would be of a world wide web multigraph model, where a vertex represents a webpage and a directed edge represents a hyperlink (arrowhead matrices). A certain class of directed multigraphs associated with arrowhead matrices are studied in [24].

In this paper, we focus on multigraphs and present an alternate interpretation of edge shifts as follows. Instead of assigning distinct labels to all the directed edges between a fixed pair of initial and terminal vertices in a multigraph, we consider only a single edge and associate a number with it, called as multiplicity of an edge. The multiplicity represents the number of all the edges between the fixed pair of initial and terminal vertices. In this way, each edge can be represented uniquely by its initial and terminal vertices. Since we keep track of the multiple occurrence of each edge through its multiplicity, information on the total number of paths is preserved. At the same time, the collection of symbols is reduced which makes the computations easier.

One of the fundamental results in the study of spectral properties of matrices is the well-known Perron Frobenius theorem. It guarantees the existence of a maximal real simple eigenvalue of any irreducible non-negative matrix. This result has plenty of applications in various branches of mathematics, as well as in other disciplines such as network theory and engineering, see [4, 5, 9, 17, 20, 26–28]. The Perron Frobenius theorem for matrices is useful in studying spectral and thermodynamic properties of shift spaces with associated binary adjacency matrix (a real matrix with each entry either 0 or 1). Guibas and Odlyzko [13] give formulas for enumeration of words of given length over some symbol set which avoid certain set of words in terms of correlation polynomials of these words. Using these combinatorial techniques, an expression for the Perron root (the largest eigenvalue) and Perron eigenvectors of the binary adjacency matrix are obtained in [7]. These results have several independent applications in other areas including symbolic dynamics, game theory, information theory and network theory, some of which are given in [1, 3, 6]. Building upon the work of Guibas and Odlyzko, in this paper, we develop similar combinatorial tools by considering a modified set of rules in terms of multiple occurrence and non-occurrence of certain words. We then employ these results to study the spectral and thermodynamic properties of edge shifts associated with a non-negative integer matrix.
Organization of the paper. In Section 2, we discuss some preliminaries on subshifts of finite type (SFT) associated to a binary adjacency matrix and some known results in this setup. In Section 3, we consider a collection $F$ of forbidden words and $R$ of repeated words having multiplicities at least 2, with symbols from a finite set $\Sigma$. We introduce the concept of generalized language denoted by $\Lambda^p$ associated to $F$ and $R$ where multiplicity of any word is computed in terms of multiplicities of words from $R$. Here $p$ is the length of the longest word in $F \cup R$. We derive a necessary and sufficient condition on $F$ and $R$ such that $\Lambda^p$ becomes the language of an edge shift associated to some non-negative integer matrix. Moreover, in case when $\Lambda^p$ is not a language, there exists a non-negative integer matrix such that the language of the edge shift associated to this matrix is maximally contained in $\Lambda^p$. Interestingly, the entropy of this edge shift can be computed in terms of the number of words of length $n$, say $f(n)$, in $\Lambda^p$ with multiplicities.

In Section 4, a formula for generating function $F(z)$ of $f(n)$ is obtained. It is observed that the Perron root of the adjacency matrix associated to $F$ and $R$ can be expressed as the largest real pole of $F(z)$. In Section 5, we obtain an expression for the left and right Perron eigenvectors of the adjacency matrix associated to $F$ and $R$ in terms of the correlation of words from $F \cup R$. In Section 6, we derive a simple expression for the normalization factor for the Perron eigenvectors. The theory has been well-studied in [7, 14], when the collection $R$ is empty, that is, when there are no repeated words, or equivalently, no multiple edges in the associated graph.

In Section 7, we observe that for a fixed $F$, the SFT associated to $F$ is a factor of the edge shift associated to $F$ and $R$ for any given collection $R$. We then study the relation between the Markov measures on the respective shifts and find that the Parry measure on the SFT associated to $F$ can be obtained as a push forward of the Shannon-Parry measure on the edge shift associated to $F$ and $R$ for some collection $R$. Section 8 consists of some concluding remarks and future directions of study.

2. Preliminaries

Let $\Sigma$ be a finite set of symbols and $\Sigma^\mathbb{N}$ be the set of all one-sided sequences with symbols from $\Sigma$. A finite sequence with symbols from $\Sigma$ is called a word. For a word $w$, let $|w|$ denote its length. A finite collection of words is said to be reduced if for any two distinct words in the collection, one is not a subword of the other. Let $F$ be a finite reduced collection of words with symbols from $\Sigma$ having length at least 2. Define the subshift of finite type, denoted as $\Sigma_F$, consisting of one-sided sequences which do not contain any word from $F$ as a subword. A word is said to be allowed if it is a subword of some sequence in $\Sigma_F$. A word is said to be forbidden if it is not allowed in $\Sigma_F$.

Definition 2.1 (Language). The language of the subshift $\Sigma_F$ is defined as the collection $L = \bigcup_{n \geq 1} L_n$ of words, where for each $n \geq 1$, $L_n$ denotes the collection of all allowed words of length $n$ in $\Sigma_F$.

A given subshift of finite type can equivalently be thought as the collection of infinite paths on certain digraph. Hence a given SFT corresponds to a matrix which is the adjacency matrix associated to this digraph. We will soon discuss this correspondence. For now, in the following subsection, we will focus on subshifts of finite type corresponding to binary adjacency matrices. The subshifts of finite type corresponding to general (non-negative integer) adjacency matrices will be discussed in the next section.

2.1. A subshift of finite type associated to a binary matrix. Let $\Sigma$ be a finite set of symbols. Let $A$ be a binary matrix whose rows and columns are indexed by $\Sigma$. Also, $A$ is assumed to be irreducible (a square non-negative matrix $A = (A_{xy})$ is called irreducible
if for every \( x, y \), there exists \( m = m(x, y) \) such that \( A_{xy}^m > 0 \). Consider the digraph \( \mathcal{G}_A = (\mathcal{V}_A, \mathcal{E}_A) \) associated to \( A \) defined as \( \mathcal{V}_A = \Sigma \) and \( \mathcal{E}_A = \{(xy) : A_{xy} = 1\} \). An infinite path of vertices on \( \mathcal{G}_A \) is of the form \( x_1x_2 \ldots \) where \( A_{x_ix_{i+1}} = 1 \) for all \( i \). Similarly, an infinite path of edges on \( \mathcal{G}_A \) is of the form \( (x_1y_1)(x_2y_2) \ldots \) where \( A_{x_iy_i} = 1 \) and \( y_i = x_{i+1} \) for all \( i \).

**Definition 2.2** (Vertex and Edge Shifts). The collection of all infinite paths of vertices on \( \mathcal{G}_A \) is called as the vertex shift associated to \( A \) and is denoted as \( \Sigma_A \). That is,

\[
\Sigma_A = \{(x_1x_2)\ldots : x_i \in \mathcal{V}_A, \, A_{x_ix_{i+1}} = 1\}
\]

The edge shift associated to \( A \), denoted by \( \Sigma_A \), is defined as the collection of all infinite paths of edges on \( \mathcal{G}_A \), that is,

\[
\Sigma_A = \{(x_1x_2)(x_2x_3)\ldots : x_i \in \mathcal{V}_A, \, A_{x_ix_{i+1}} = 1\}
\]

Note that \( \Sigma_A = \Sigma_F \) where \( F = \{xy : A_{xy} = 0\} \). Here all words in \( F \) are of length 2. Further, the number of allowed words of length \( n \) in \( \Sigma_F \), given by \( |\mathcal{L}_n| \), equals the sum of entries of \( A^{n-1} \). Also, \( \Sigma_A \) and \( \Sigma_A \) are conjugates where the conjugacy \( \mathcal{F} \) is given by \( x_1x_2x_3\cdots \mapsto (x_1x_2)(x_2x_3)\cdots \).

In short, for a given binary matrix, both vertex and edge shifts are conjugate to an SFT where the forbidden collection consists of words of length 2. Conversely, if we are given a forbidden collection \( F \) consisting of words of length 2, we can obtain a digraph such that the vertex shift and the edge shift associated to it are conjugate to \( \Sigma_F \). In the following section, we consider a forbidden collection \( F \) that contains words of arbitrary lengths and analyze properties similar to those just discussed.

### 2.2 A binary adjacency matrix for a general \( F \).

Let \( F \) be a finite reduced collection of forbidden words with symbols from \( \Sigma \). Let \( \Sigma_F \) be the associated SFT. Throughout this paper, we assume that symbols in \( \Sigma \) are not forbidden, otherwise we can remove them from the collection of symbols \( \Sigma \). Let \( \mathcal{L} = \bigcup_{n \geq 1} \mathcal{L}_n \) be its language. For \( m \geq 1 \), we define a conjugacy from \( \Sigma_F \) to an \( m \)-step subshift \( \Sigma_F^{[m]} \) as

\[
x_1x_2x_3\cdots \mapsto (x_1\ldots x_m)(x_2\ldots x_{m+1})(x_3\ldots x_{m+2})\ldots,
\]

where the subshift \( \Sigma_F^{[m]} \) has symbols from \( \mathcal{L}_m \) (refer higher block presentation in [21]).

**Definition 2.3.** Let \( m \geq 2 \) and \( X = x_1x_2\ldots x_m, Y = y_1y_2\ldots y_m \in \mathcal{L}_m \). Then \( X \ast Y \) is defined only if \( x_2\ldots x_m = y_1\ldots y_{m-1} \), and

\[
X \ast Y := x_1x_2\ldots x_my_m.
\]

If \( X = x_1, Y = y_1 \in \mathcal{L}_1 \), then \( X \ast Y = x_1y_1 \).

Let \( p \) be the length of the longest word in \( F \). Now we define an adjacency matrix \( A = (A_{XY}) \) indexed by words from \( \mathcal{L}_{p-1} \) (the indexing of rows and columns are in lexicographic order for convenience), where

\[
A_{XY} = \begin{cases} 1, & \text{if } X \ast Y \text{ is defined and } X \ast Y \in \mathcal{L}_p \\ 0, & \text{otherwise.} \end{cases}
\]

This adjacency matrix gives a digraph \( \mathcal{G}_A = (\mathcal{V}_A, \mathcal{E}_A) \) where \( \mathcal{V}_A = \mathcal{L}_{p-1} \) and \( \mathcal{E}_A = \{(XY) : A_{XY} = 1\} \). We label \((XY) \in \mathcal{E}_A\) as \( X \ast Y \). Then the vertex shift \( \Sigma_A \) consists of infinite paths of vertices on \( \mathcal{G}_A \) that are of the form \( X_1X_2X_3\ldots \), where \( A_{X_iX_{i+1}} = 1 \) for all \( i \).

That is \( \Sigma_A = \Sigma_F^{[p-1]} \). Similarly, the edge shift \( \Sigma_A = \Sigma_F^{[p]} \). Note that \( \Sigma_A \) and \( \Sigma_A \) are

---

Here conjugacy is simply a “recoding” of sequences (we refer to [21] Definition 1.5.9]). In symbolic dynamics, this type of conjugacy is known as sliding block code.
conjugates to $\Sigma$. Since the words from $L_{p-1}$ and $L_p$ completely describe the subshift, the words in $L_{p-1} \cup L_p$ can be regarded as the building blocks of $\Sigma$. Moreover, $|L_n|$ is given by the sum of entries of $A^{n-p+1}$. A vertex (edge) shift is said to be irreducible, if the adjacency matrix associated to it is irreducible. In this case, the graph $G_A$ associated to $A$ is strongly connected.

2.3. Subshift of finite type as a dynamical system. For a given $F$ with longest word of length $p$, consider the adjacency matrix $A$ indexed by words from $L_{p-1}$, as defined in the previous subsection. In this subsection, we discuss some properties of $\Sigma_A$ from a dynamical viewpoint. Since $\Sigma_F$ or $\Sigma_A$ (in fact any $\Sigma^{[m]}_F$, $m \geq 1$) is conjugate to $\Sigma_A^n$, they exhibit the same properties.

The left shift map $\sigma : \Sigma_A^u \to \Sigma_A^v$ is defined as $\sigma(X_1X_2X_3\ldots) = X_2X_3\ldots$. A $\sigma$-invariant probability measure $\mu$ on $\Sigma_A^v$, known as the Parry measure, is defined as follows.

**Definition 2.4** (The Parry measure). For an allowed word $w = X_1X_2\ldots X_n$, let $C_w$ denote the cylinder based at $w$, that is, $C_w$ consists of all sequences in $\Sigma_A^v$ that begin with the word $w$. Then

$$\mu(C_w) = \frac{U_{X_1}V_{X_n}}{\theta^{n-1}},$$

where $\theta$ is the largest eigenvalue of $A$ (known as the Perron root), which exists, and is positive by the Perron Frobenius theorem. The vectors $U$ and $V$, with indexing of rows same as that of $A$, are the left and right Perron (column) eigenvectors corresponding to the Perron root satisfying $U^2V = 1$. Extend $\mu$ to the sigma-algebra generated by all cylinders. Since $\Sigma_A^u$ and $\Sigma_A$ are conjugates, this measure can be extended to give a measure on $\Sigma_A$, which will also be called the Parry measure.

It is a fact that the topological entropy of $\Sigma_F$ (or $\Sigma_A^u$ or $\Sigma_A$) which is defined as

$$h_{\text{top}}(\Sigma_F) = \lim_{n \to \infty} \frac{1}{n} \ln(|L_n|),$$

is given by $\ln(\theta)$, where $\theta$ is the Perron root of $A$. The Parry measure is the unique measure of maximal entropy according to the variational principle, that is, the measure theoretic entropy with respect to $\mu$ is equal to $\ln(\theta)$, see [12]. In [11], the authors obtain a combinatorial expression for the Perron root and the associated eigenvectors of $A$, for a given forbidden collection $F$. We now recall their results.

2.4. Summary of existing results. The concept of correlation between two words which determines the overlap between them, is described below.

**Definition 2.5** (Correlation). The correlation of two words $u$ and $v$, denoted by $(u,v)$, is a binary string $(c_1,\ldots, c_{|u|})$ of length $|u|$, defined by the following algorithm. The $i^{th}$ bit $c_i$ is determined by placing $v$ below $u$ such that the leftmost symbol in $v$ lies under the $i^{th}$ symbol from left in $u$. Set $c_i = 1$ if and only if the overlapping segments of $u$ and $v$ are identical, else set $c_i = 0$. The correlation of $u$ and $v$ can also be interpreted as a polynomial in some variable $z$ as, $(u,v)_z = \sum_{i=1}^{|u|} c_i z^{|u|-i}$. The polynomial $(u,v)_z$ is termed as the correlation polynomial.

Let $F = \{a_1, a_2, \ldots, a_s\}$ be a reduced collection of forbidden words with symbols from $\Sigma$, $\Sigma_F$ be the associated subshift and $L = \bigcup_{n \geq 1} L_n$ be its language. Let $f(n) = |L_n|$ and $f_i(n)$ be the number of words of length $n$ ending with $a_i$ with symbols from $\Sigma$ not containing any of the words from $F$ except the single appearance of $a_i$ at the end. Let $F(z) = \sum_{n=0}^{\infty} f(n)z^{-n}$ and $F_i(z) = \sum_{n=0}^{\infty} f_i(n)z^{-n}$, $1 \leq i \leq s$, be their respective generating
functions. The following result gives an expression for these generating functions in terms of the correlation between the forbidden words.

**Theorem 2.6.** [14, Theorem 1] The generating functions \(F(z), F_i(z)\) satisfy the linear system of equations

\[
K(z) \begin{pmatrix} F(z) \\ F_1(z) \\ \vdots \\ F_s(z) \end{pmatrix} = \begin{pmatrix} z \\ 0 \\ \vdots \\ 0 \end{pmatrix},
\]

where \(K(z) = \begin{pmatrix} z - q & z \mathbf{1}^T \\ 1 & -z \mathbf{M}(z) \end{pmatrix}\), \(\mathbf{M}(z) = ((a_j, a_i)_{z})_{1 \leq i,j \leq s}\) is the correlation matrix for the collection \(\mathcal{F}\), \(\mathbf{1}\) denotes the column vector of size \(s\) with all 1’s and \(q\) is the size of \(\Sigma\). In particular,

\[
F(z) = \frac{z}{z - q + R(z)},
\]

where \(R(z)\) is the sum of the entries of \(\mathbf{M}^{-1}(z)\).

In [7, Theorem 1], the Perron root of the associated adjacency matrix \(A\) is proved to be the largest real pole of \(F(z)\) and hence we have the following result.

**Theorem 2.7.** With the notations as above, the Perron root \(\theta > 0\) is the largest (in modulus) zero of the rational function \(z - q + R(z)\).

In [7], the authors also obtain an expression for the left and right Perron eigenvectors corresponding to the Perron root, in terms of the correlation polynomials. This is then used to obtain a combinatorial expression for the Parry measure on \(\Sigma^r\).

3. **Subshift associated with a non-negative integer matrix**

In the preceding section, we obtained an SFT associated to a given binary matrix and vice-a-versa. In this section, we generalize this concept to a subshift of finite type associated to a non-negative integer matrix. Here the corresponding digraph may have multiple edges between its vertices.

### 3.1. A subshift of finite type from a non-negative integer matrix.

Let \(\Sigma\) be a finite symbol set. Let \(A = (A_{xy})\) be a non-negative integer matrix indexed by \(\Sigma\). Consider the directed multigraph \(G_A = (\mathcal{V}_A, \mathcal{E}_A)\) associated to the matrix \(A\) where \(\mathcal{V}_A = \Sigma\) and the number of edges from the vertex \(x\) to \(y\) is given by \(A_{xy}\). Let us visualize \(G_A\) labelled with all the edges of initial vertex \(x\) and terminal vertex \(y\) distinctly as, \((xy)_1, \ldots, (xy)_{A_{xy}}\). Using the edge set \(\mathcal{E}_A\) as a new symbol set, we consider the edge shift associated with \(A\), denoted by \(\Sigma_A\), as

\[
\Sigma_A = \{ (x_1x_2i_1 (x_2x_3i_2) \ldots : x_k \in \mathcal{V}_A, 1 \leq i_k \leq A_{x_kx_{k+1}}, k \geq 1 \}.
\]

Let \(\mathcal{L}_n\) be the set of all allowed words of length \(n\) in \(\Sigma_A\), and let \(\mathcal{L} = \bigcup_{n \geq 1} \mathcal{L}_n\) be the language of \(\Sigma_A\). Note here that the label for the language is different than that in Definition 2.1.

**Definition 3.1** (Shannon-Parry measure). The Shannon-Parry measure (refer to [21,30]) on \(\Sigma_A\) is defined as

\[
\mu(C_W) = \frac{U_{x_1}V_{x_{n+1}}}{\theta^n},
\]

where \(\theta\) is the Perron root of \(A\).
where $C_W$ is the cylinder based at $W = (x_1 x_2)_{i_1} (x_2 x_3)_{i_2} \ldots (x_n x_{n+1})_{i_n} \in L_n$, $\theta$ is the Perron root of $A$ and $U$ and $V$ are the left and right Perron (column) eigenvectors such that $U^T V = 1$.

The Shannon-Parry measure $\mu$ is invariant with respect to the left shift map $\sigma$ on $\Sigma_A$. It is the unique measure of maximal entropy where the topological entropy of $\Sigma_A$ is given by $h_{\text{top}}(\Sigma_A) = \lim n \to \infty \frac{1}{n} \ln(|L_n|)$ and is the same as $\ln(\theta)$.

Let $F = \{xy : A_{xy} = 0\}$, $\Sigma_F$ be the associated SFT and $L = \bigcup_{n \geq 1} L_n$ be its language. Note that $\Sigma_F$ consists of all infinite paths of vertices on $G_A$. However, unlike in Section 2, $\Sigma_F$ and $\Sigma_A$ need not be conjugates as $\Sigma_F$ does not differentiate the multiple edges. To incorporate this, we define a multiplicity for each word in $\Sigma_F$.

**Definition 3.2** (Multiplicity of an allowed word). Let $w = x_1 x_2 \ldots x_n \in L_n$. The multiplicity of $w$ in $G_A$, denoted as $m(w)$, is defined as the number of paths (of edges) of length $n - 1$ in $G_A$ with fixed vertices $x_1, \ldots, x_n$ (in order). In other words, $m(w) = \prod_{i=1}^{n-1} A_{x_i x_{i+1}}$.

A word $w \in L_n$ corresponds to $m(w)$ many words of the form $(x_1 x_2)_{i_1} (x_2 x_3)_{i_2} \ldots (x_{n-1} x_n)_{i_{n-1}}$ in $L_{n-1}$. Hence

$$|L_{n-1}| = \sum_{w \in L_n} m(w),$$

which equals the sum of entries of the matrix $A^{n-1}$.

A word is called repeated if its multiplicity is greater than 1. For $x, y \in V_A$, $m(xy) = A_{xy}$. Define $R$ to be the collection of repeated words given by $R = \{xy : A_{xy} > 1\}$. Note that the multiplicity of any allowed word can be computed in terms of multiplicities of words from $R$. Here words from $F$ and $R$ have length 2 and they uniquely determine $\Sigma_A$.

3.2. A non-negative integer matrix for general collections $F$ and $R$. Now we consider a general situation where a finite reduced collection $F$ of forbidden words is given. Let $\Sigma_F$ be the associated subshift of finite type. Assume that $\Sigma_F$ is irreducible with language given by $L = \bigcup_{n \geq 1} L_n$, where $L_n$ is the collection of allowed words of length $n$ in $\Sigma_F$.

Let $R = \{r_1, r_2, \ldots, r_\ell\}$ be a reduced collection of allowed words in $\Sigma_F$. For each word $r_i \in R$, we assign a number $m_i$, at least 2, which we call the multiplicity of the word $r_i$. For convenience of notation, we denote $R$ as $R = \{r_1(m_1), r_2(m_2), \ldots, r_\ell(m_\ell)\}$, where the multiplicity $m_i$ of the word $r_i$ is mentioned in the corresponding parentheses. Here the words from $F$ or $R$ can be of arbitrary lengths. Given a collection $R$, we extend the concept of multiplicity to all finite words as follows.

**Definition 3.3** (Multiplicity of a word). Let $R = \{r_1(m_1), \ldots, r_\ell(m_\ell)\}$ be a given collection consisting of allowed words in $\Sigma_F$. For any $w \in \Sigma^n$, we define its multiplicity, denoted by $m(w)$, as,

$$m(w) = \begin{cases} \prod_{r_i \in \mathcal{R}} m_i^{n(w;r_i)}, & \text{if } w \in L_n, \\ 0, & \text{otherwise} \end{cases}$$

where $n(w;r_i)$ denotes the number of times the word $r_i \in \mathcal{R}$ appears as a subword of the word $w$. An allowed word is said to be repeated if $m(w) > 1$.

Clearly if an allowed word $w$ does not contain any repeated word then $m(w)$ equals 1. Moreover for each $r_i \in \mathcal{R}$, $m(r_i) = m_i$. 
Definition 3.4 (Generalized Language). For $n \geq 1$, define a multiset
\[
\Lambda_n := \{(w, m(w)) : w \in \mathcal{L}_n\},
\]
where we use the notation $(w, m(w))$ to denote the fact that an allowed word $w \in \mathcal{L}_n$ appears $m(w)$ many times in $\Lambda_n$. We define the generalized language corresponding to $\mathcal{F}$ and $\mathcal{R}$ of level $m \geq 1$ as the multiset,
\[
\Lambda^m = \bigcup_{n \geq m} \Lambda_n.
\]

Using the properties of a multiset, the cardinality of $\Lambda_n$, denoted by $|\Lambda_n|$, is given by, $|\Lambda_n| = \sum_{w \in \mathcal{L}_n} m(w)$. In what follows, the statement ‘$w \in \Lambda_n$’ will mean that $w$ is an element of $\mathcal{L}_n$ and is repeated $m(w)$ many times in the multiset $\Lambda_n$.

Example 3.5. Let $\Sigma = \{0,1\}$, $\mathcal{F} = \{01\}$ and $\mathcal{R} = \{00(\alpha),111(\beta)\}$. Then $\Lambda_1 = \{(0,1),(1,1)\}$, $\Lambda_2 = \{(00,\alpha),(10,1),(11,1)\}$, $\Lambda_3 = \{(000,\alpha^2),(100,\alpha),(110,1),(111,\beta)\}$.

Remark 3.6. Let $A = (A_{xy})$ be a non-negative integer matrix and $G_A$ be the associated digraph. Consider the edge shift $\Sigma_A$ associated to $A$ as in Section 3.1 and let $\mathcal{F} = \{xy : A_{xy} = 0\}$ and $\mathcal{R} = \{xy : A_{xy} > 1\}$ with $m(xy) = A_{xy}$. For $n \geq 2$ and an allowed word $w = x_1 \ldots x_n \in \Lambda_n$, $m(w)$ is the multiplicity of $w$ in $G_A$ as given in Definition 3.2 that is, $m(w)$ is the number of paths on $G_A$ with fixed vertices $x_1, x_2, \ldots, x_n$ (in order).

For $n \geq 2$, an allowed word $(x_1x_2)_{i_1}(x_2x_3)_{i_2} \ldots (x_{n-1}x_n)_{i_{n-1}}$ in $\Sigma_A$ corresponds to exactly one of $m(w)$ many repetitions of $w = x_1x_2 \ldots x_n$ giving a bijection between the language of $\Sigma_A$ and the generalized language $\Lambda^2$. Moreover, the topological entropy of $\Sigma_A$ is $h_{\text{top}}(\Sigma_A) = \lim_{n \to \infty} \frac{1}{n} \ln(|\Lambda_n|)$. Hence we can say that $\Lambda^2$ provides an alternate interpretation of $\Sigma_A$. An illustration of this interpretation is given in Figure 1 for the adjacency matrix $A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$. Note that the graph on the left is $G_A$ with edges labelled.

![Figure 1](image-url)

Now we address the converse of Remark 3.6. Let $p$ denote the length of the longest word in $\mathcal{F} \cup \mathcal{R}$. Our aim is to find a matrix $A$ (indexed by $\mathcal{L}_{p-1}$) such that the entropy of the edge shift $\Sigma_A$ is $h_{\text{top}}(\Sigma_A) = \lim_{n \to \infty} \frac{1}{n} \ln(|\Lambda_n|)$. As before, we will need to consider the higher block representation. One possibility for such a matrix $A$ is when the associated edge shift $\Sigma_A$ has language (in bijection with) $\Lambda^p = \bigcup_{n \geq p} \Lambda_n$.

Using sliding block code representation, $\Lambda^p$ is the language of $\Sigma_A$ for some adjacency matrix $A$ indexed by $\mathcal{L}_{p-1}$ if and only if for each $w = x_1 \ldots x_n \in \Lambda_n$ with $n \geq p$, $m(w)$ matches with the number of paths with fixed vertices $X_1, \ldots, X_{n-p+2}$ (in order) where $X_i = x_i \ldots x_{i+p-2}$, for $1 \leq i \leq n - p + 2$. 


Example 3.7. Let \( \Sigma = \{0, 1\} \), \( \mathcal{F} = \{11\} \) and \( \mathcal{R} = \{00(\alpha)\} \). If \( A = \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} \), then by Remark 3.6, \( \Sigma_A \) has language \( \Lambda^2 \). In fact, whenever words from \( \mathcal{F} \) and \( \mathcal{R} \) have length 2, then \( \Sigma_A \) has language \( \Lambda^2 \) where \( A_{xy} = m(xy) \).

Remark 3.8. For given collections \( \mathcal{F}, \mathcal{R} \), a seemingly natural choice for the adjacency matrix \( A \) is the one indexed by words from \( \Lambda_{p-1} \), with \( A_{XY} = m(X \star Y) \), for each \( X, Y \in \Lambda_{p-1} \) (as in Example 3.7). But the following example shows that this is not the right choice always. Let \( \Sigma = \{0, 1\} \), \( \mathcal{F} = \{00\} \) and \( \mathcal{R} = \{110(2), 01(3)\} \). Here \( p = 3 \), and the matrix \( A \), indexed by \( \{01, 10, 11\} \), is given by

\[
\begin{pmatrix}
0 & 3 & 3 \\
3 & 0 & 0 \\
0 & 2 & 1
\end{pmatrix}
\]

and its associated graph \( \mathcal{G}_A \) is as shown in Figure 2.

Here \( m(1010) = 3 \). However, in \( \mathcal{G}_A \), there are 9 distinct paths of length 2 with fixed vertices \( X_1 = 10, X_2 = 01, X_3 = 10 \) (in order). Also, \( h_{\text{top}}(\Sigma_A) \sim \ln(3.9) \), but \( \lim_{n \to \infty} \frac{1}{n} \ln(|\Lambda_n|) \sim \ln(2.6) \) (this limit is calculated using the results from Section 4).

Let \( X = x_1 \ldots x_n, Y = y_1 \ldots y_n \) such that \( X \star Y \) is defined but \( X \star Y \notin \mathcal{R} \). Then,

\[
m(X \star Y) = \frac{m(X)m(Y)}{m(x_2 \ldots x_n)}.
\]

In order to incorporate this term in the denominator, for any word \( v = v_1 \ldots v_n \in \mathcal{L}_n \) we define a number

\[
k(v) := \frac{m(v)}{m(v_2 \ldots v_n)}.
\]

Therefore \( m(X \star Y) = k(X)m(Y) \), provided \( X \star Y \notin \mathcal{R} \). It is easy to see that \( k(v) > 1 \) if and only if \( v \) begins with a word from \( \mathcal{R} \).

For each \( v \in \Lambda_n, n > p \), the multiplicity of \( v \) can be obtained from the multiplicities of words from \( \Lambda_{p-1} \cup \Lambda_p \). Hence the words in \( \Lambda_{p-1} \cup \Lambda_p \) act as building blocks for the words in \( \Lambda_n \).

The following result provides a necessary and sufficient condition on \( \mathcal{F} \) and \( \mathcal{R} \) such that \( \Lambda^p \) is a language of \( \Sigma_A \) for some \( A \).

Theorem 3.9. Let \( p \) be the length of the longest word in \( \mathcal{F} \cup \mathcal{R} \). The generalized language \( \Lambda^p = \bigcup_{n \geq p} \Lambda_n \) is the language of an edge shift if and only if all the words in \( \mathcal{R} \) are of length \( p \).
Conversely, suppose $\Lambda^p$. Then there exists a non-negative integer matrix $A$ indexed by $\mathcal{L}_{p-1}$ as follows,

$$A_{XY} = \begin{cases} m(X * Y), & \text{if } X * Y \in \mathcal{L}_p \\ 0, & \text{otherwise} \end{cases}$$

Consider the associated digraph $G_A$ and the edge shift $\Sigma_A$. We claim that $\Lambda^p$ is the language of $\Sigma_A$. That is, for each word $w = x_1 x_2 \ldots x_n \in \Lambda_n$, $n \geq p$, $m(w)$ is same as the number of paths with fixed vertices $X_1, \ldots, X_{n-p+2}$ (in order), where $X_i = x_i \ldots x_{i+p-2}$, for $1 \leq i \leq n - p + 2$. Note that the number of paths in $G_A$ with fixed vertices $X_1, X_2, \ldots, X_n$ (in order) is

$$m(X_1 * X_2) m(X_2 * X_3) \ldots m(X_{n-1} * X_n),$$

which is the same as $m(w)$ since $\mathcal{R}$ has words of length $p$. Hence the claim.

Conversely, suppose $\Lambda^p$ is the language of an edge shift. Since $\Lambda_{p-1} \cup \Lambda_p$ is the building block for the words in $\Lambda^p$, the edge shift has to be associated with the matrix $A$ as defined in the previous part of the proof (since path of length 1 defines the entries in the matrix). We need to prove that all words in $\mathcal{R}$ are of length $p$. Suppose on the contrary, there exists $X = x_1 \ldots x_{p-1} \in \Lambda_{p-1}$, such that $m(X) > 1$. Choose $x, y \in \Sigma$, such that $xx \ldots x_{p-1}y \in \Lambda_{p+1}$ (use irreducibility of $\Sigma_f$). The number of paths with fixed vertices $xx \ldots x_{p-2}, X, x_2 \ldots x_{p-1}y$ is $m(xx \ldots x_{p-1}y)m(x_1 \ldots x_{p-1}y)$. However,

$$m(xx \ldots x_{p-1}y)m(x_1 \ldots x_{p-1}y) = m(xx \ldots x_{p-1}y)m(X) > m(xx \ldots x_{p-1}y),$$

which is a contradiction. \hfill \Box

By Theorem 3.9 $\Lambda^p$ need not always be a language. The following result describes the structure of $\Lambda^p$ in general. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be languages of two subshifts of finite type. Then we write $\mathcal{L}_1 \subseteq \mathcal{L}_2$ if $\mathcal{L}_1$ is contained in $\mathcal{L}_2$ upon relabeling the symbols of $\mathcal{L}_1$.

**Theorem 3.10.** With the notations as above, let $p$ be the length of the longest word in $\mathcal{F} \cup \mathcal{R}$. Then there exists a non-negative integer matrix $A$ indexed by $\mathcal{L}_{p-1}$ such that the associated edge shift $\Sigma_A$ with language $\mathcal{L}(\Sigma_A)$ satisfies the following.

(i) $\mathcal{L}(\Sigma_A) \subseteq \Lambda^p$,

(ii) The topological entropy $h_{\text{top}}(\Sigma_A) = \lim_{n \to \infty} \frac{1}{n} \ln |\Lambda_n|$.

(iii) $h_{\text{top}}(\Sigma_A) > 0$, if $\mathcal{R} \neq \emptyset$.

**Proof.** Define a matrix $A = (A_{XY})_{X,Y \in \mathcal{L}_{p-1}}$ indexed by $\mathcal{L}_{p-1}$ as follows: for $X, Y \in \mathcal{L}_{p-1}$, let

$$A_{XY} = \begin{cases} k(X * Y), & \text{if } X * Y \in \mathcal{L}_p \\ 0, & \text{otherwise}. \end{cases} \quad (1)$$

(i) Let $G_A$ be the graph associated with the matrix $A$. Note that for a word $w = X * Y \in \Lambda_p$, the number of edges in $G_A$ from vertex labelled $X$ to vertex labelled $Y$ equals $k(X * Y)$, and also $k(X * Y) \leq m(w)$. Let $n > p$. Let $w = x_1 x_2 \ldots x_n \in \Lambda_n$ be an allowed word. Using induction on $n$, for $X_i = x_i \ldots x_{i+p-2}$, we have

$$m(w) = k(X_1 * X_2) \ldots k(X_{n-p} * X_{n-p+1}) m(X_{n-p+1} * X_{n-p+2}).$$

Also, the number of paths with vertices labelled $X_1, \ldots, X_{n-p+2}$ (in order) in $G_A$ is given by $k(X_1 * X_2) \ldots k(X_{n-p+1} * X_{n-p+2})$, which is at most $m(w)$. Hence, clearly $\mathcal{L}(\Sigma_A) \subseteq \Lambda^p$.

(ii) Let $\tau(n)$ denote the number of paths of length $n$ in the graph $G_A$, given by,

$$\tau(n - p + 1) = \sum_{w \in \mathcal{L}_n} k(X_1 * X_2) \ldots k(X_{n-p+1} * X_{n-p+2}).$$
Set $M := \max_{Y \in \mathcal{L}_p} \left\{ \frac{m(Y)}{k(Y)} \right\} > 0$. Then

$$|\Lambda_n| = \sum_{w \in \mathcal{L}_n} m(w) \leq M \sum_{w \in \mathcal{L}_n} k(X_1 \ast X_2) \ldots k(X_{n-p+1} \ast X_{n-p+2}) = M \tau(n - p + 1).$$

Also by part (i), $\tau(n - p + 1) \leq |\Lambda_n|$, for each $n > p$. Hence,

$$h_{top}(\Sigma_A) = \lim_{n \to \infty} \frac{1}{n} \ln \tau(n) = \lim_{n \to \infty} \frac{1}{n} \ln |\Lambda_n|.$$

(iii) Finally, let $r \in \mathcal{R}$ be a repeated word with multiplicity $m \geq 2$. Since $\Sigma_F$ is irreducible, there exist a word $s \in \Lambda_p$ that contains $r$ and a word $sws \in \Lambda^p$. Consider the subsequence $(n_j)_{j \geq 1}$, where $n_j = j(|s| + |w|)$. The word $r$ appears in the word $swsw \ldots sw \in \Lambda_{n_j}$ at least $j$ many times. Thus $|\Lambda_{n_j}| \geq m^j$ and hence

$$h_{top}(\Sigma_A) = \lim_{j \to \infty} \frac{1}{n_j} \ln |\Lambda_{n_j}| \geq \lim_{j \to \infty} \frac{j \ln(m)}{j(|s| + |w|)} = \frac{\ln(m)}{|s| + |w|} > 0.$$

\[\square\]

**Definition 3.11** (Adjacency matrix associated with the collections $\mathcal{F}$ and $\mathcal{R}$). The matrix $A$, defined by Equation (1) is called the adjacency matrix associated with the collections $\mathcal{F}$ and $\mathcal{R}$.

**Example 3.12.** Let $\Sigma = \{0, 1\}$, $\mathcal{F} = \{010, 101, 111\}$, and $\mathcal{R} = \{00(2), 0110(3)\}$. Here $p = 4$ and $\Lambda_3 = \{(000, 4), (001, 2), (011, 1), (100, 2), (110, 1)\}$. The graph corresponding to the adjacency matrix $A$ associated with the collections $\mathcal{F}$ and $\mathcal{R}$ is shown in Figure 3.

![Figure 3. Graph $G_A$ for Example 3.12](image)

**Remark 3.13.** A few observations about the matrix $A$:

1. Since $\mathcal{R}$ is reduced, for $X \ast Y \in \mathcal{L}_p$,

$$k(X \ast Y) = \begin{cases} k(X), & \text{if } X \ast Y \notin \mathcal{R} \\ m(X \ast Y), & \text{otherwise.} \end{cases}$$

2. The non-zero entries of the row of $A$ indexed by a repeated word $X$ (with $m(X) > 1$) are all equal to $k(X)$, using Equation (2).

3. Even though the Perron root of $A$ is given by $\lim_{n \to \infty} |\Lambda_n|^{1/n}$, note that $|\Lambda_n|$ may not be equal to the sum of entries of $A^{n-p+1}$. This equality holds when words in $\mathcal{R}$ are all of equal length $p$, by Theorem (3.9).

4. One can define a new graph by the outsplittings of $A$ at each vertices with multiplicity at least 2 (we refer to [21] for the definition of outsplittings) and then adding a few stranded vertices (i.e., with no outgoing edges) so that the total number of paths on this new graph is exactly equal to as in $\Lambda^p$. The shift on this graph and $\Sigma_A$ are conjugates, since the stranded vertices do not contribute to the infinite sequences. Hence we can prove the existence of a graph where finite paths completely describe $\Lambda^p$. 
Example 3.14. Let us review the example given in Remark 3.8. Consider $\Sigma = \{0, 1\}$, $F = \{00\}$ and $R = \{110(2), 01(3)\}$. Here $p = 3$ and the adjacency matrix $A$ (indexed by $\{01, 10, 11\}$) associated with $F$ and $R$ is given by

\[
\begin{pmatrix}
0 & 3 & 3 \\
1 & 0 & 0 \\
0 & 2 & 1 \\
\end{pmatrix},
\]

and its associated graph is shown in Figure 4. Note that $h_{\text{top}}(\Sigma A) = \lim_{n \to \infty} \frac{1}{n} \ln(|\Lambda_n|) \sim \ln(2.6)$. Here sum of entries of $A$ is 9 but $|\Lambda_3| = |\{(010, 3), (011, 3), (101, 3), (110, 2), (111, 1)\}| = 12$.

Consider the following matrix $\tilde{A}$, indexed by $L_{p-1}$, given by

\[
\tilde{A}_{XY} = \begin{cases} 
m(X \ast Y), & X \ast Y \in L_p \\
0, & \text{otherwise} \end{cases}
\]

Clearly $L(\Sigma_{\tilde{A}}) \supseteq \Lambda^p$, where $L(\Sigma_{\tilde{A}})$ denotes the language of $\Sigma_{\tilde{A}}$. When all the words in $R$ have length $p$, then the adjacency matrix associated with $F$ and $R$ equals $\tilde{A}$. However, as in Example 3.14, the matrices $A$ and $\tilde{A}$ can be different.

Let $L(S)$ denote the language of an SFT $S$ with symbols from $\Sigma$. We will visualize $S$ using its $p$-block presentation. Let $S$ be the collection of all subshifts of finite type $S$ such that $L(S) \subset \Lambda^p$. Clearly $\Sigma_A \in S$ where $A$ is the adjacency matrix associated with the collections $F$ and $R$. Let $S \in S$. If $\tau_S(n)$ denotes the number of allowed words of length $n$ in $S$, then $\tau_S(n) \leq |\Lambda_n|$. Hence $h_{\text{top}}(S) \leq \ln \theta$. Moreover, we have the following result.

Theorem 3.15. There is no subshift of finite type $S \in S$ that satisfies $L(\Sigma_A) \subsetneq L(S) \subsetneq \Lambda^p$.

Proof. Suppose there is such a subshift $S$. Let $t > 1$ be the length of the shortest word in $L(S) \setminus L(\Sigma_A)$. Consider the adjacency matrices $A', A'_S$ corresponding to the $(t - 1)$-step shift of both $\Sigma_A$ and $S$ respectively. Note that since $A' < A'_S$, their respective Perron roots satisfy $\theta < \lambda_S$. Hence $\ln \theta \leq h_{\text{top}}(S)$ which is a contradiction since $L(S) \ Subseteq \Lambda^p$. □

Remark 3.16. If the collection $R$ has words of length strictly smaller than $p$, then we define a new extended collection

\[
\tilde{R} = \{X \ast Y : X, Y \in \Lambda_{p-1}, X \ast Y \in \Lambda_p \text{ and begins with a word from } R\}.
\]

For $X \ast Y \in \tilde{R}$, define its multiplicity to be $k(X \ast Y)$. Clearly, the adjacency matrix associated with $F$ and $\tilde{R}$ is the same as the adjacency matrix $A$ associated with $F$ and $R$. Since all the words in $\tilde{R}$ are of length $p$, by Theorem 3.9, $L(\Sigma_A)$ can be characterized using the generalized language associated with this new collection $F$ and $\tilde{R}$.
For the rest of the paper, let us fix \( q \geq 2 \), a finite symbol set \( \Sigma \) of size \( q \) and reduced collections \( \mathcal{F} = \{a_1, \ldots, a_s\} \) of forbidden words and \( \mathcal{R} = \{r_1(m_1), \ldots, r_\ell(m_\ell)\} \) of repeated words with symbols from \( \Sigma \). Here \( m_j > 1 \) for all \( 1 \leq j \leq \ell \). Fix \( p \) to denote the length of the longest word in \( \mathcal{F} \cup \mathcal{R} \) and \( \Lambda^p = \cup_{n \geq p} \Lambda_n \) be the generalized language associated with \( \mathcal{F} \) and \( \mathcal{R} \).

4. Enumeration of words in the generalized language

In this section, we study the asymptotic behaviour of \( |\Lambda_n| \) (see Definition 3.4) for given \( \mathcal{F} \) and \( \mathcal{R} \), using combinatorial tools. This gives the entropy of the subshift \( \Sigma^A \) obtained in Theorem 3.10. We consider two cases depending on whether \( \mathcal{F} \cup \mathcal{R} \) is reduced or not.

Let \( u, v \) be two words with symbols from \( \Sigma \). Here \( u \) and \( v \) have \( m \) words with symbols from \( \Sigma \). Here \( n \) is the length of the longest word in \( \mathcal{F} \cup \mathcal{R} \).

Let \( \mathcal{L}_n \) be the set of all allowed words of length \( n \) in \( \Sigma_{\mathcal{F}} \) and let \( w = x_1x_2 \ldots x_n \in \mathcal{L}_n \).

Observe that if one of the given repeated words \( r_j \) is a subword of \( w \), then there exists \( t \in (w, r_j) \) such that \( t \geq |r_j| \).

Let \( f(n) = \sum_{w \in \mathcal{L}_n} m(w) = |\Lambda_n| \) (see Definition 3.4). By convention, let \( f(0) = 1 \).

For each \( 1 \leq j \leq \ell \), let \( g_{r_j}(n) = \sum_{w} m(w) \), where summation is over all the words \( w \in \mathcal{L}_n \) which end with a repeated word \( r_j \in \mathcal{R} \). Note that such a word \( w \) can have more repeated words as subwords. By convention, let \( g_{r_j}(0) = 0 \).

Although the multiplicity of any forbidden word is 0 by Definition 3.2, we next redefine the multiplicity of certain kinds of forbidden words \( w \), only limited to this section. For convenience of notation, we denote it also as \( m(w) \).

**Definition 4.1.** [Multiplicity of a word which contains a forbidden word as a subword only at the end – only limited to the discussion in Section 4] Let \( w \) be a forbidden word in \( \Sigma_{\mathcal{F}} \) such that \( w \) ends with \( a_i \in \mathcal{F} \) and this \( a_i \) at the end is the only occurrence of a forbidden word in \( w \). We define multiplicity of \( w \), denoted as \( m(w) \), by

\[
m(w) := \frac{\prod_{r_j \in \mathcal{R}} m_j^{n(a_i, r_j)}}{\prod_{r_j \in \mathcal{R}} m_j^{n(w; r_j)}}.
\]

where \( n(w; v) \) denotes the number of times the word \( v \) appears as a subword of the word \( u \). In particular, for each \( 1 \leq i \leq s \), \( m(a_i) = 1 \).

For \( 1 \leq i \leq s \), let \( f_{a_i}(n) = \sum_{w} m(w) \), where the summation is over all the words \( w \) of length \( n \) with symbols from \( \Sigma \), which end with a forbidden word \( a_i \in \mathcal{F} \), and moreover, this \( a_i \) is the only occurrence of a forbidden word in \( w \). By convention, \( f_{a_i}(0) = 0 \).

The corresponding generating functions for \( f(n) \), \( g_{r_j}(n) \) and \( f_{a_i}(n) \) are defined as follows.

\[
F(z) = \sum_{n=0}^{\infty} f(n) z^{-n}, \quad G_{r_j}(z) = \sum_{n=0}^{\infty} g_{r_j}(n) z^{-n}, \quad 1 \leq j \leq \ell,
\]

\[
F_{a_i}(z) = \sum_{n=0}^{\infty} f_{a_i}(n) z^{-n}, \quad 1 \leq i \leq s.
\]

When \( \mathcal{R} \) is an empty set, the generating functions \( F \) and \( F_{a_i} \) are described in Theorem 2.6. In the general case, the generating functions \( F, F_{a_i}, G_{r_j} \) will be described in
Theorem 4.3 when $\mathcal{F} \cup \mathcal{R}$ is a reduced collection, and in Theorem 4.11 when $\mathcal{F} \cup \mathcal{R}$ is not a reduced collection.

In the following remark, we set up new terminologies which we will use in this section, and also make some related observations.

Remark 4.2. Let $w \in \Lambda_n$.

1. By definition of $f(n)$, we say that $w$ gets counted in $f(n)$ or simply, $w$ is in $f(n)$.
2. We say that $w$ gets counted in $f_a(n + t)$ (or $g_{r_j}(n + s)$) for $t \leq |a_i|$ (or $s \leq |r_j|$) if there exists a word $w'$ in $f_a(n + t)$ (or $g_{r_j}(n + s)$) which begins with $w$. Note that this $w'$ is unique for a given $w$ when considered without the multiplicities.
3. For $w$ and $w'$ as above, we say that $w$ gets counted $m(w')$ times in $f_a(n + t)$ (or $g_{r_j}(n + s)$). Note that $m(w') \geq m(w)$, and moreover if a repeated word appears at some position in $w'$ and not as a subword of $w$, then $m(w') > m(w)$.
4. A repeated word $r_j$ in $w'$ gives $m_j - 1$ many extra counting of $w$ in $f_a(n + t)$ (or $g_{r_j}(n + s)$) if $r_j$ appears in $w'$ other than being a subword of $w$.
5. For any two words $u$ and $v$, let $uv$ denote the word obtained by concatenation of $v$ as a suffix to $u$. We say that a word $w$ appears on the join of $u$ and $v$ if there exists $t > 0$ such that $0 < |w| - t \in (u, w)$ and $t \in (w, v)$. That is, some initial part of $w$ overlaps with the end part of $u$ and the remaining part of $w$ overlaps with the beginning part of $v$.

Suppose now that the collection $\mathcal{R}$ of words with repetitions is non-empty. Consider the collection $\mathcal{F} \cup \mathcal{R}$. In Subsection 4.1, we will discuss the first case where $\mathcal{F} \cup \mathcal{R}$ is reduced. In this case, any forbidden word from $\mathcal{F}$ does not contain a word from $\mathcal{R}$ as its subword. The situation gets complex in general, when $\mathcal{F} \cup \mathcal{R}$ is not reduced, that is, some repeated word sits inside a forbidden word as a subword. This will be discussed in Subsection 4.2.

4.1. When $\mathcal{F} \cup \mathcal{R}$ is reduced. In this case, the collections $\mathcal{F}$ and $\mathcal{R}$ will be automatically reduced and no repeated word in $\mathcal{R}$ appears as a subword of a word in $\mathcal{F}$.

Theorem 4.3. Let $\mathcal{F} = \{a_1, \ldots, a_s\}$ and $\mathcal{R} = \{r_1(m_1), \ldots, r_\ell(m_\ell)\}$ be the collections of forbidden and repeated words respectively, where $m_j > 1$ for all $1 \leq j \leq \ell$. If $\mathcal{F} \cup \mathcal{R}$ is reduced, then the generating functions $F(z), G_{r_j}(z)$ and $F_{a_i}(z)$ satisfy the following system of linear equations.

\begin{align}
(z - q)F(z) - \sum_{j=1}^\ell z \left(1 - \frac{1}{m_j}\right) G_{r_j}(z) + \sum_{i=1}^s z F_{a_i}(z) &= z, \\
F(z) + \sum_{j=1}^\ell \left[ z \left(1 - \frac{1}{m_j}\right) (r_j, r_k)_z - z^{r_j}\delta_{jk}\right] G_{r_j}(z) - \sum_{i=1}^s z (a_i, r_k)_z F_{a_i}(z) &= 0, \\
&= 0, \\
&\text{for } 1 \leq k \leq \ell, \\
&\text{for } 1 \leq k \leq s,
\end{align}

where $\delta_{jk} = 1$, if $j = k$, otherwise $\delta_{jk} = 0$.

Proof. In order to obtain the first equation in the above system, let the word $w$ be counted in $f(n)$ and $x \in \Sigma$ be a symbol. Consider the concatenation of $w$ and $x$ given by $wx$. 


Then the word \(wx\) is counted with multiplicities in either \(f(n + 1)\), or in \(f_{a_i}(n + 1)\), for some \(a_i \in \mathcal{F}\). Therefore
\[
\sum_{x \in \Sigma} \sum_{w \in \mathcal{L}_n} m(wx) = f(n + 1) + \sum_{i=1}^{s} f_{a_i}(n + 1). \tag{6}
\]

To determine the quantity on the left hand side, observe that for each \(x \in \Sigma\), the word \(wx\) is counted at least \(m(w)\) times, which gives total of \(qf(n)\) counting. Moreover, only those \(wx\) which end with some repeated word \(r_j \in \mathcal{R}\), are counted \(m_j - 1\) times extra (one counting is already incorporated in \(qf(n)\)). The word \(wx\) ends with a repeated word if and only if the last symbol of \(r_j\) is \(x\) and a tail of \(w\) matches with the beginning subword of \(r_j\) of length \(|r_j| - 1\), that is, \(|r_j| - 1 \in (w, r_j)\). Hence
\[
m_j \sum_{\{w \in \mathcal{L}_n : |r_j| - 1 \in (w, r_j)\}} m(w) = gr_j(n + 1).
\]

Since each \(w\) satisfying \(|r_j| - 1 \in (w, r_j)\) gets counted extra \(m_j - 1\) times when concatenated with an \(x\) same as the last symbol of \(r_j\) at the end, we have
\[
\sum_{x \in \Sigma} \sum_{w \in \mathcal{L}_n} m(wx) = qf(n) + \sum_{j=1}^{\ell} \frac{g_{r_j}(n + 1)}{m_j} (m_j - 1). \tag{7}
\]

Combining (6) and (7), we get
\[
qf(n) - f(n + 1) + \sum_{j=1}^{\ell} \left(1 - \frac{1}{m_j}\right) g_{r_j}(n + 1) - \sum_{i=1}^{s} f_{a_i}(n + 1) = 0. \tag{8}
\]

On multiplying both the sides of (8) by \(z^{-n}\), taking summation over \(n \geq 0\), making use of the conventions \(f(0) = 1, f_{a_i}(0) = g_{r_j}(0) = 0\), we obtain first equation (3) in the system.

Next, consider an allowed word \(w\) counted in \(f(n)\), \(r_k \in \mathcal{R}\), and the word \(wr_k\) obtained by concatenation. If \(wr_k\) does not contain any forbidden word, then \(w\) is counted in \(g_{r_k}(n + |r_k|)\). Note that it may be counted more than the multiplicity \(m(w)\), due to possible appearance of repeated words from \(\mathcal{R}\) on the join and at the end. We will take care of these extra counting later.

Further if some forbidden word from \(\mathcal{F}\) appears in \(wr_k\), then we look at the first occurrence of a forbidden word in \(wr_k\). Suppose the forbidden word from \(\mathcal{F}\) which occurs first is \(a_i \in \mathcal{F}\). Since \(w\) does not contain a forbidden word and since \(\mathcal{F} \cup \mathcal{R}\) is reduced, the word \(a_i\) appears on the join of \(w\) and \(r_k\). Thus there exists \(0 < t \in (a_i, r_k)\) such that \(w\) is counted in \(f_{a_i}(n + t)\) (counted more than \(m(w)\) times if repeated words appear on the join and before the placement of the word \(a_i\)).

Now, we need to take away the extra counting of the words as mentioned before. Figure (5) illustrates situations like these. Suppose some \(r_j \in \mathcal{R}\) appears on the join of \(w\) and \(r_k\), then there exists \(0 < s \in (r_j, r_k)\). The number of words \(w\) of length \(n\) that give \(r_j\) on the join (and no forbidden word before the placement of \(r_j\)) is precisely \(\frac{g_{r_j}(n + s)}{m_j}\). But, each such \(w\) gets counted \(m_j - 1\) times more in \(g_{r_j}(n + s)\), due to the multiplicity of the occurrence of \(r_j\) at the end. Thus, we subtract such extra counting \(\frac{g_{r_j}(n + s)}{m_j}(m_j - 1)\) to obtain
\[
f(n) = g_{r_k}(n + |r_k|) + \sum_{i=1}^{s} \sum_{0 < t \in (a_i, r_k)} f_{a_i}(n + t) - \sum_{j=1}^{\ell} \sum_{0 < s \in (r_j, r_k)} \left(1 - \frac{1}{m_j}\right) g_{r_j}(n + s). \tag{9}
\]
The condition that $\mathcal{F} \cup \mathcal{R}$ is reduced implies that $t < |a_i|$ and $s < |r_j|$ in (9). Multiply the above equation by $z^{-n}$ and take sum over $n \geq 0$ to obtain (10).

$$f(n) = \sum_{i=1}^{s} \sum_{0 < t \in (a_i, a_k)} f_{a_i}(n + t) - \sum_{j=1}^{t} \sum_{0 < s \in (r_j, a_k)} \left(1 - \frac{1}{m_j}\right) g_{r_j}(n + s). \quad (10)$$

To complete the proof, we multiply (10) by $z^{-n}$ and take sum over $n \geq 0$ to obtain the last equation (5) in the system.

**Example 4.4.** We illustrate Theorem 4.3 using this example. For $q = 2$, let $\mathcal{R} = \{000(2)\}$ and $\mathcal{F} = \{010\}$. Clearly $\mathcal{F} \cup \mathcal{R}$ is a reduced collection. Denote $r_1 = 000$ with $m_1 = 2$ and $a_1 = 010$. Then, $1 \in (a_1, r_1), 1, 2, 3 \in (r_1, r_1), 1, 3 \in (a_1, a_1)$ and $1 \in (r_1, a_1)$. Equations (8), (9) and (10) can be verified using the values given in the table below.

| $n$  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|------|----|----|----|----|----|----|----|----|----|----|
| $f(n)$ | 2  | 4  | 8  | 17 | 37 | 81 | 178| 392| 864|1905|
| $g_{r_1}(n)$ | 0  | 0  | 2  | 6  | 14 | 32 | 72 | 160| 354| 782|
| $f_{a_1}(n)$ | 0  | 0  | 1  | 2  | 4  | 9  | 20 | 44 | 97 | 214|

Matrix representation of the system of linear equations obtained in Theorem 4.3. Consider the matrix function $L(z)$ given as

$$L(z) = \begin{pmatrix} z - q & -z(1 - \frac{1}{m_1}) & \ldots & -z(1 - \frac{1}{m_\ell}) & zI_{\ell+s} \\ 1_{\ell+s} & (z) \\ \end{pmatrix},$$

where $I_m$ denotes the column matrix of size $m$ with all entries 1, $P(z) = [P(z)_{ij}]_{1 \leq i,j \leq \ell+s}$ with

$$P(z)_{ij} = \begin{cases} z(1 - \frac{1}{m_j})(r_j, r_i), & 1 \leq i, j \leq \ell, \\ -z(a_j, r_i), & 1 \leq i \leq \ell, \ell + 1 \leq j \leq \ell + s, \\ z(1 - \frac{1}{m_j})(r_j, a_i), & \ell + 1 \leq i \leq \ell + s, 1 \leq j \leq \ell, \\ -z(a_j, a_i), & \ell + 1 \leq i, j \leq \ell + s, \end{cases} \quad (11)$$

where $\delta_{ij} = 1$ if $i = j$, else $\delta_{ij} = 0$.

Using the matrix function $L(z)$, the linear system of equations in Theorem 4.3 can be

![Figure 5. Words $a_i$ and $r_j$ appear on the join, with $r_j$ placed before $a_i$.](image-url)
We are allowing for existence of a word in \( W \) in the form of a rational function:

\[
L(z) = \begin{pmatrix}
F(z) \\
G_{r_1}(z) \\
\vdots \\
G_{r_s}(z) \\
F_{a_1}(z) \\
\vdots \\
F_{a_t}(z)
\end{pmatrix} = \begin{pmatrix}
z \\
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix}.
\]

With notations as above, the generating function \( F(z) \) takes the following form of a rational function:

\[
F(z) = \frac{z}{z - q + z \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) R_i(z) - z \sum_{j=1}^{s'} R_{\ell+j}(z)},
\]

where \( R_i(z) \) is the \( i \)-th row sum of \( P^{-1}(z) \).

**Proof.** The result follows by computing \( L(z)^{-1} \) using the formula for inverse of a \( 2 \times 2 \) block matrix given in [23].

We will now obtain another form of \( F(z) \) in the following result. We will first define notations required for the result. Consider the diagonal matrix function \( D(z) = \text{diag}\{D_{ii}(z)\}_{1 \leq i \leq \ell+s} \), where

\[
D_{ii}(z) = \begin{cases}
\frac{z - \frac{1}{m_i}}{1 - \frac{1}{m_i}}, & 1 \leq i \leq \ell, \\
-z, & \ell + 1 \leq i \leq \ell + s.
\end{cases}
\]

Define another matrix function

\[
Q = D^{-1}P^TD.
\]

**Theorem 4.6.** With notations as above, \( F(z) \) takes the following form of a rational function:

\[
F(z) = \frac{z}{z - q + z \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) S_i(z) - z \sum_{j=1}^{s'} S_{\ell+j}(z)},
\]

where \( S_i(z) \) denotes the \( i \)-th row sum of inverse of \( Q^{-1}(z) \).

**Proof.** The proof follows by observing the relationship between the matrix functions \( P \) and \( Q \) and Theorem 4.5.

**Remark 4.7.** If the collection \( \mathcal{R} \) is empty, the matrix \( L(z) \) is equal to \( K(z) \) as defined in Theorem 2.6. Consequently, Theorem 2.6 is obtained as a corollary to the Theorem 4.3.

4.2. When \( \mathcal{F} \cup \mathcal{R} \) is not reduced. In this section, we consider the reduced collections \( \mathcal{F} \) and \( \mathcal{R} \) such that \( \mathcal{F} \cup \mathcal{R} \) is not reduced. Since every word in \( \mathcal{R} \) is an allowed word, it cannot contain a forbidden word as its subword. However, since \( \mathcal{F} \cup \mathcal{R} \) is not reduced, some forbidden word in \( \mathcal{F} \) contains a word from \( \mathcal{R} \) as its subword. That is, in this section, we are allowing for existence of \( a_i \in \mathcal{F}, r_j \in \mathcal{R} \) and \( t \geq |r_j| \) such that \( t \in (a_i, r_j) \).

Let \( w \) be in \( f(n) \) and suppose it gets counted in \( f_{a_i}(n+t) \) for some \( 0 < t \leq |a_i| \). Let \( w' \) be the corresponding word in \( f_{a_i}(n+t) \). Suppose \( r_j \) is a subword of \( a_i \) which appears in \( w' \) such that this occurrence of \( r_j \) is not contained in \( w \). Then by considering the
redefined multiplicity of the forbidden $w'$ given in Definition 4.1, this occurrence of $r_j$ does not contribute to any extra counting of $w$ in $f_{a_i}(n + t)$.

We now make another important observation. If $w$ in $f(n)$ gets counted in $f_{a_i}(n + t)$ for some $0 < t \leq |a_i|$ and $a_i$ contains some $r_j$ as a subword except at the end, then clearly $w$ also gets counted in $g_{r_j}(n + s)$ for some $0 < s < t$. However if $r_j$ appears as a subword of $a_i$ at the end, then $w$ can not get counted simultaneously in both $g_{r_j}(n + t)$ and $f_{a_i}(n + t)$ as $g_{r_j}$ counts only the allowed words ending with $r_j$. Therefore, we introduce the following notation.

**Notation.** The total number of distinct positions at which $r_j$ appears in $a_i$ as a subword, except for at the end is denoted as

$$\gamma(a_i, r_j) = \# \{\alpha \in (a_i, r_j) : \alpha > |r_j|\}.$$  \hfill (16)

We also define another concept which will be useful in this section.

**Definition 4.8** ($\alpha$-tail correlation and $\alpha$-tail correlation polynomial). Let $u$ and $v$ be two words with symbols from $\Sigma$, $(u, v) = (c_1, \ldots, c_{|u|})$ be the correlation of $u$ and $v$ and let $1 \leq \alpha \leq |u|$. We define the $\alpha$-tail correlation of $u$ and $v$, denoted by $(u, v)^\alpha$, as the binary string of length $\alpha$ consisting of the last $\alpha$ elements of $(u, v)$, that is,

$$(u, v)^\alpha = (c_{|u| - \alpha + 1}, c_{|u| - \alpha + 2}, \ldots, c_{|u|}).$$

The $\alpha$-tail correlation of $u$ and $v$ can also be interpreted as a polynomial in some variable $z$ as defined below, and is termed as the $\alpha$-tail correlation polynomial of $u$ and $v$. For a given $\alpha \geq 1$, the $\alpha$-tail correlation polynomial of $u$ and $v$, denoted as $(u, v)_z^\alpha$, is defined as

$$(u, v)_z^\alpha = \sum_{i=1}^{|u|} c_i z^{|u| - i} = \sum_{0 < t \in (u, v)^\alpha} z^{t-1}.$$  

**Remark 4.9.** For $\alpha = |u|$, $(u, v)^{|u|} = (u, v)$ and $(u, v)_z^{|u|} = (u, v)_z$.

**Example 4.10.** Let $u = 210210$ and $v = 2102$. Then $(u, v) = (1, 0, 0, 1, 0, 0)$ and $(v, u) = (1, 0, 0, 1)$. The 4-tail correlation of $u$ and $v$ is $(u, v)^4 = (0, 1, 0, 0)$ and the 2-tail correlation of $v$ and $u$ is $(v, u)^2 = (0, 1)$. Moreover the corresponding polynomials are

$$(u, v)_z = z^4 + z^2; \quad (v, u)_z = z^3 + 1; \quad (u, v)_z^4 = z^2; \quad (v, u)_z^2 = 1.$$  

**Theorem 4.11.** Let $\mathcal{F} = \{a_1, \ldots, a_s\}$ be a reduced collection of forbidden words and $\mathcal{R} = \{r_1(m_1), \ldots, r_s(m_s)\}$ be a reduced collection of repeated words with $m_i > 1$, such that the union $\mathcal{F} \cup \mathcal{R}$ is not reduced. Then the generating functions $F(z), G_{r_j}(z)$ and $F_{a_i}(z)$ satisfy the following system of linear equations:

$$(z - q) F(z) - \sum_{j=1}^\ell z \left(1 - \frac{1}{m_j}\right) G_{r_j}(z)$$

$$+ \sum_{i=1}^s z \left[1 + \sum_{j=1}^\ell (m_j - 1) \gamma(a_i, r_j) \right] F_{a_i}(z) = z$$

$$F(z) + \sum_{j=1}^\ell \left[z \left(1 - \frac{1}{m_j}\right) (r_j, r_k)_z z^{|r_j|} \delta_{jk}\right] G_{r_j}(z)$$

$$- \sum_{i=1}^s z(a_i, r_k)^{|r_j|} \left[1 + \sum_{j=1}^\ell (m_j - 1) \gamma(a_i, r_j) \right] F_{a_i}(z) = 0, \quad 1 \leq k \leq \ell,$$  

where $\delta_{jk}$ is the Kronecker delta.
\[ F(z) + \sum_{j=1}^{\ell} z \left( 1 - \frac{1}{m_j} \right) (r_j, a_k)^{r_j-1} G_{r_j}(z) \]

\[ - \sum_{i=1}^{n} \left[ z(a_i, a_k)z + \sum_{0 < t \in (a_i, a_k)} z^t \sum_{j=1}^{\ell} (m_j - 1) \gamma_t(a_i, r_j) \right] F_{a_i}(z) = 0, \quad 1 \leq k \leq s. \]  

The proof of this result is a counting argument as in the proof of Theorem 4.3. The counting is trickier here due to the appearance of repeated words as a subword of a forbidden word since the collection \( \mathcal{F} \cup \mathcal{R} \) is not reduced. Due to heavy technical calculations, the proof is given in the Appendix.

Remark 4.12. When the collection \( \mathcal{F} \cup \mathcal{R} \) is reduced, then \( \gamma(a_i, r_j) = 0 \) for all \( 1 \leq j \leq \ell, 1 \leq i \leq s \). Moreover, the tail-correlation polynomials \( (a_i, r_k)_z^{r_k} \) and \( (r_j, a_k)_z^{r_j-1} \) coincide with the correlation polynomials \( (a_i, r_k)_z \) and \( (r_j, a_k)_z \). Thus Theorem 4.3 is a corollary to Theorem 4.11.

5. Perron Root and Eigenvectors of A

In this section, we obtain a description of the Perron root of the adjacency matrix \( A \) associated with given \( \mathcal{F} \) and \( \mathcal{R} \) using the results from the previous section and also derive expressions for the left and right Perron eigenvectors of \( A \) in terms of correlation polynomials. When \( \Sigma_A \) is irreducible, we apply Theorems 3.10, 4.5 and 4.6 together to obtain the Perron root of \( A \) as the largest real pole of \( F(z) \). The proof follows along the same lines as the proof of [7, Theorem 4.1]. Hence we immediately arrive at the following result.

Theorem 5.1 (The Perron root). With the notations as above, the Perron root of \( A \) is given by the largest real zero of

\[ z - q + z \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) R_i(z) - z \sum_{j=1}^{s} R_{\ell+j}(z), \]

which equals \( z - q + z \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) S_i(z) - z \sum_{j=1}^{s} S_{\ell+j}(z) \) (see (12) and (15) for notations).

For a given non-negative irreducible integer matrix \( A \) of size \( q \) labelled by symbols \( 0, 1, \ldots, q-1 \) (in order), define \( \mathcal{F} = \{ xy : A_{xy} = 0 \} = \{ a_1, \ldots, a_s \} \) and \( \mathcal{R} = \{ xy : A_{xy} > 1 \} = \{ r_1, \ldots, r_r \} \) with multiplicity \( m_i \) of \( r_i \) defined by \( A_{xy} \). Then trivially, \( \mathcal{F} \cup \mathcal{R} \) is reduced with the adjacency matrix \( A \). Hence we can obtain an expression for the Perron root of \( A \) using Theorem 5.1.

Remark 5.2. The Perron root of the adjacency matrix \( A \) associated with \( \mathcal{F} \) and \( \mathcal{R} \) is given by the largest real pole of the generating function \( F(z) \) even when \( \mathcal{F} \cup \mathcal{R} \) is not reduced. Hence, one can solve the system of equations (17)-(19) to calculate the Perron root of \( A \). For instance, consider the example when \( q = 2, \mathcal{F} = \{ 001 \} \) and \( \mathcal{R} = \{ 00(2) \} \), let \( a = 001 \) and \( r = 00 \). Then the system (17)-(19) is given by

\[
\begin{pmatrix}
  z - 2 & \frac{2z}{z - 2} & 2z \\
 1 & \frac{z - 2}{z} & 0 \\
 1 & \frac{z - 2}{z} & -z^2 \\
\end{pmatrix}
\begin{pmatrix}
  F(z) \\
  G_r(z) \\
  F_a(z) \\
\end{pmatrix}
= \begin{pmatrix}
  z \\
  0 \\
  0 \\
\end{pmatrix},
\]

which gives \( F(z) = \frac{z}{z - 2} \). The largest pole of \( F(z) \) is 2, which is the Perron root of the associated adjacency matrix \( A = \begin{pmatrix}
  2 & 0 & 0 & 0 \\
  0 & 0 & 1 & 1 \\
  1 & 1 & 0 & 0 \\
  0 & 0 & 1 & 1 \\
\end{pmatrix} \), indexed by \( \{00, 01, 10, 11\} \) in order.
Now, we will obtain an expression for the Perron eigenvectors of $A$. Let $p$ be the length of the longest word in $\mathcal{F} \cup \mathcal{R}$. Let us assume, from now on in this section, that all the words in $\mathcal{R}$ are of length $p$. Consequently $\mathcal{F} \cup \mathcal{R}$ is reduced. In this case, $A_{XY} = m(X \ast Y)$ if $X \ast Y \in \mathcal{L}_p$, and is 0 otherwise. For a word $X$ of length $p - 1$ with symbols from $\Sigma$, define

$$U_X = 1 - \theta \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) R_i(\theta)(\tilde{r}_i, X)_\theta + \theta \sum_{j=1}^{s} R_{\ell+j}(\theta)(\tilde{a}_j, X)_\theta,$$  

(20)

$$V_X = 1 - \theta \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta)(X, r_i)_\theta + \theta \sum_{j=1}^{s} S_{\ell+j}(\theta)(X, a_j)_\theta,$$  

(21)

where $\tilde{r}_i, \tilde{a}_j$ are the words obtained by removing the first symbol of $r_i, a_j$ respectively (note that $(\tilde{r}_i, X)_\theta = (r_i, X)_\theta$ and $(\tilde{a}_j, X)_\theta = (a_j, X)_\theta$ are the $|r_i| - 1$ and $|a_j| - 1$ tail correlations as described in Section 5.2 respectively) and $(u, w)_\theta$ is the correlation polynomial $(u, w)_z$ evaluated at $z = \theta$. Then we have the following result. As argued in [7] Lemma 5.3 it is easy to show that $R_i(\theta)$ and $S_i(\theta)$ exist.

**Theorem 5.3** (Perron eigenvectors). Let $\theta$ be the Perron eigenvalue of the adjacency matrix $A$ associated with $\mathcal{F}$ and $\mathcal{R}$, $U = (U_X)_{X \in \mathcal{L}_{p-1}}$ and $V = (V_X)_{X \in \mathcal{L}_{p-1}}$ be as defined in (20) and (21). Then $U$ and $V$ are respectively left and right Perron eigenvectors of $A$.

**Proof.** We first prove that $V$ is a right eigenvector of $A$ with respect to the Perron root $\theta$. For a fixed allowed word $X = x_1 x_2 \ldots x_{p-1}$ of length $p - 1$, we need to show that

$$\sum_{b=0}^{q-1} m_{Xb} V_{Xb} = \theta V_X,$$

where $\tilde{X}b = x_2 x_3 \ldots x_{p-1} b$ and $m_{Xb}$ is the multiplicity of $Xb = x_1 x_2 \ldots x_{p-1} b$. By convention, if $Xb$ is forbidden, then $m_{Xb} = 0$. That is, we need to show that,

$$0 = \sum_{b=0}^{q-1} m_{Xb} \left(1 - \theta \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta)(\tilde{X}b, r_i)_\theta + \theta \sum_{j=1}^{s} S_{\ell+j}(\theta)(\tilde{X}b, a_j)_\theta \right)$$

$$- \theta \left(1 - \theta \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta)(X, r_i)_\theta + \theta \sum_{j=1}^{s} S_{\ell+j}(\theta)(X, a_j)_\theta \right)$$

$$= \sum_{b=0}^{q-1} m_{Xb} - \theta \left(\sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta) \left(\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, r_i)_\theta - \theta(X, r_i)_\theta\right)\right)$$

$$+ \theta \left(\sum_{j=1}^{s} S_{\ell+j}(\theta) \left(\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, a_j)_\theta - \theta(X, a_j)_\theta\right)\right).$$  

(22)

First of all, we make a couple of observations. Fix $w \in \mathcal{F} \cup \mathcal{R}$. We consider the following two cases:

i) Let $X$ be such that $Xb$ does not end with $w$ for any $b \in \Sigma$. Then for any $z$,

$$\sum_{b=0}^{q-1} (\tilde{X}b, w)_z = z(X, w)_z + 1.$$  

(23)
ii) Let $X$ be such that $Xb_0$ ends with $w$ for some $b_0 \in \Sigma$. Such $b_0$ is unique and for any $z$,

$$\sum_{b=0}^{q-1} (\tilde{X}b, w)_z + z |w|^{-1} = z(X, w)_z + 1. \quad (24)$$

Let $X$ be such that $Xb_1 = r_1, \ldots, Xb_d = r_d$ and $Xb_{d+1}, \ldots, Xb_{d+n}$ end with $a_1, \ldots, a_n$, respectively, for some $0 \leq d \leq \ell, 0 \leq n \leq s$ so that $m_1 = m_{Xb_1}, \ldots, m_d = m_{Xb_d} > 1$, $m_{Xb_{d+1}} = \ldots = m_{Xb_{d+n}} = 0$ and $m_{Xb_{d+n+1}} = \ldots = m_{Xb_q} = 1$ (rename words from $\mathcal{F}$ or $\mathcal{R}$ if needed). We look at terms in (22) separately.

For $i = 1, \ldots, d$ and $k = 1, \ldots, d$, since $Xb_k = r_k$, we have for $k \neq i, (\tilde{X}b_k, r_i)_\theta = (r_k, r_i)_\theta$, and for $k = i, (\tilde{X}b_i, r_i)_\theta = (r_i, r_i)_\theta - \theta^{p-1}$. Also for $k = 1, \ldots, n$, since $Xb_{d+k}$ ends with $a_k$ and $\mathcal{F} \cup \mathcal{R}$ is reduced, we have $(\tilde{X}b_{d+k}, r_i)_\theta = (a_k, r_i)_\theta$. Hence using (24),

$$\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, r_i)_\theta - \theta(X, r_i)_\theta = \sum_{k=1}^{d} m_k(\tilde{X}b_k, r_i)_\theta + \sum_{k=d+1}^{q} (\tilde{X}b_k, r_i)_\theta - \theta(X, r_i)_\theta$$

$$= \sum_{b=0}^{q-1} (\tilde{X}b, r_i)_\theta - \theta(X, r_i)_\theta + \sum_{k=1}^{d} (m_k - 1)(\tilde{X}b_k, r_i)_\theta - \sum_{k=1}^{n} (a_k, r_i)_\theta \quad (25)$$

Similarly for $i = d + 1, \ldots, \ell$, using (23), we get

$$\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, r_i)_\theta - \theta(X, r_i)_\theta = 1 + \sum_{k=1}^{d} (m_k - 1)(r_k, r_i)_\theta - \sum_{k=1}^{n} (a_k, r_i)_\theta. \quad (26)$$

Combining (25) and (26), we obtain

$$\sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta) \left(\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, r_i)_\theta - \theta(X, r_i)_\theta \right)$$

$$= \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) S_i(\theta) \left(1 + \sum_{k=1}^{d} (m_k - 1)(r_k, r_i)_\theta - \sum_{k=1}^{n} (a_k, r_i)_\theta \right) - \sum_{i=1}^{\ell} (m_i - 1)\theta^{p-1} S_i(\theta). \quad (27)$$

We use similar steps to obtain,

$$\sum_{j=1}^{s} S_{\ell+j}(\theta) \left(\sum_{b=0}^{q-1} m_{Xb}(\tilde{X}b, a_j)_\theta - \theta(X, a_j)_\theta \right)$$

$$= \sum_{j=1}^{s} S_{\ell+j}(\theta) \left(1 - \sum_{k=1}^{d} (m_k - 1)(r_k, a_j)_\theta - \sum_{k=1}^{n} (a_k, a_j)_\theta \right). \quad (28)$$

Also note that $\sum_{b=0}^{q-1} m_{Xb} = q + \sum_{k=1}^{d} (m_k - 1) - n$. Combining (27) and (28) in Equation (22), we obtain the expression
\[ q - \theta + \sum_{k=1}^{d} (m_k - 1) - n - \theta \left( \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) S_i(\theta) \left( 1 + \sum_{k=1}^{d} (m_k - 1)(r_k, r_i) - n(a_k, r_i) \right) \right) \\
+ \theta^p \sum_{i=1}^{d} (m_i - 1) S_i(\theta) + \theta \sum_{j=1}^{s} S_{\ell+j}(\theta) \left( 1 - \sum_{k=1}^{d} (m_k - 1)(r_k, a_j) - n(a_k, a_j) \right). \]

We use that \( q - \theta = \theta \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) S_i(\theta) + \theta \sum_{j=1}^{s} S_{\ell+j}(\theta) = 0. \) Now consider the following terms separately:

\[ \sum_{k=1}^{d} (m_k - 1) \left( 1 - \theta \left( \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) (r_k, r_i) - \theta^p \right) S_i(\theta) + \theta \sum_{j=1}^{s} (r_k, a_j) S_{\ell+j}(\theta) \right), \]

and

\[ \sum_{k=1}^{n} \theta \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) (a_k, r_i) S_i(\theta) - \theta \sum_{j=1}^{s} (a_k, a_j) S_{\ell+j}(\theta) \]

If \( Q_{i,j}(z) \) denotes the \( i,j \)-th entry of \( Q(z) \), then \( \sum_{j=1}^{\ell+s} Q_{i,j}(z) S_j(z) = 1 \), for all \( i = 1, \ldots, \ell+s \) and for all \( z \). Using this, we get that both the terms (29) and (30) are zero.

To prove that \( U \) is a left eigenvector of \( A \) with respect to \( \theta \), we define another adjacency matrix \( B \) corresponding to the collections \( \tilde{F} = \{ \tilde{a}_1, \ldots, \tilde{a}_s \} \) and \( \tilde{R} = \{ \tilde{r}_1(m_1), \ldots, \tilde{r}_\ell(m_\ell) \} \) where \( \tilde{w} \) is the reverse of \( w \). Unlike \( A \), here the rows and columns are indexed by the reverse of the allowed words of length \( p - 1 \). For \( X = x_1 \ldots x_{p-1} \) and \( Y = y_1 \ldots y_{p-1} \), the \( \tilde{X} \tilde{Y} \)-th entry of \( B \) is given by multiplicity of \( \tilde{X} \tilde{Y} \) (with respect to \( \tilde{F} \) and \( \tilde{R} \)) where \( \tilde{X} \tilde{Y} \) is defined if \( x_{p-2} \ldots x_1 = y_{p-1} \ldots y_2 \) and \( \tilde{X} \tilde{Y} = x_{p-1}y_{p-1} \ldots y_1 \). Note that this is same as the \( YX \)-th entry of \( A \). Hence \( B \) is the transpose of \( A \). Since \( \mathcal{F} \cup \mathcal{R} \) is reduced, for any \( u, w \in \mathcal{F} \cup \mathcal{R} \), note that \( (\tilde{u}, \tilde{w})_z = (w, u)_z \) for all \( z \). Hence by the first part of this proof, \( B \) has a right Perron eigenvector \( (A \) has a left Perron eigenvector) with the \( \tilde{X} \)-th entry (\( X \)-th entry) given by

\[ 1 - \theta \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) R_i(\theta)(\tilde{X}, \tilde{r}_i) + \theta \sum_{j=1}^{s} R_{\ell+j}(\theta)(\tilde{X}, \tilde{a}_j). \]

To obtain the final expression, observe that \( (\tilde{X}, \tilde{r}_i)_z = (\tilde{r}_i, X)_z \) and \( (\tilde{X}, \tilde{a}_j)_z = (\tilde{a}_j, X)_z \).

**Example 5.4.** Let \( \Sigma = \{0, 1\} \), \( \mathcal{F} = \{010\} \), \( \mathcal{R} = \{100(3)\} \). Then the adjacency matrix is indexed by \( \mathcal{L}_2 = \{00, 01, 10, 11\} \) and is given by

\[ A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \]

Here \( P(z) = \begin{pmatrix} -z^3/3 & -z^2 \\ 2z^3/3 & -z(z^2 + 1) \end{pmatrix} \), \( Q(z) = \begin{pmatrix} -z^3/3 & -z \\ 2z^2/3 & -z(z^2 + 1) \end{pmatrix} \), \( R_1(z) = -3(1 - z + z^2)/z^2(2 + z + z^2) \), \( R_2(z) = -(2 + z^2)/z^2(2 + z + z^3) \); \( S_1(z) = -3/(2 + z + z^3) \); and \( S_2(z) = -(2 + z)/(2z + z^2 + z^4) \). This gives the
Perron root of $A$ to be the largest zero of \( z - 2 + \frac{2}{3}zR_1(z) - zR_2(z) = z - 2 + \frac{2}{3}zS_1(z) - zS_2(z) \), which is \( \theta = 2 \). Here

\[
U = \begin{pmatrix} 3/2 \\ 1/2 \\ 1 \end{pmatrix}, \quad \text{and} \quad V = \begin{pmatrix} 2/3 \\ 2/3 \\ 4/3 \end{pmatrix}
\]

**Remark 5.5.** Consider the case when $\mathcal{R}$ has a word of length strictly less than $p$. Here we define a new collection of repeated words $\tilde{\mathcal{R}}$ as given in Remark 3.16 that consists of all words of length $p$ that start with words from $\mathcal{R}$. In this case Theorem 5.3 gives Perron eigenvector of $A$ in terms of words from $\mathcal{F}$ and $\tilde{\mathcal{R}}$. For example, if $\Sigma = \{0,1\}$, $\mathcal{F} = \{0000\}$ and $\mathcal{R} = \{01(2)\}$. Then we look at the following collection $\tilde{\mathcal{R}} = \{0100(2), 0101(2), 0110(2), 0111(2)\}$ and calculate the eigenvectors of $A$ indexed by $\mathcal{L}_3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$ where $A_{XY} = k(X \ast Y)$ for $X, Y \in \mathcal{L}_3$.

**Perron eigenvectors for a non-negative matrix.** Let $A = [A_{xy}]_{0 \leq x,y \leq q-1}$ be a non-negative irreducible integer matrix of size $q$. Let $\mathcal{F} = \{xy : A_{xy} = 0\} = \{a_1, \ldots, a_s\}$ and $\mathcal{R} = \{xy : A_{xy} > 1\} = \{r_1, \ldots, r_t\}$ with multiplicity $m_i$ of $r_i = xy$ given by $A_{xy}$. If $P(z), Q(z)$ denote the matrix functions as given in (11) and (12) for these collections $\mathcal{F}$ and $\mathcal{R}$. Let $R_i(z), S_i(z)$ denote the $i$th row sum of $P^{-1}(z), Q^{-1}(z)$, respectively. Since all the forbidden words and repeated words are of length 2 here, the expressions for Perron eigenvectors take a simplified form as given in the following result.

**Corollary 5.5.1** (Perron eigenvectors for a non-negative matrix). Let $U = (U_x)_{0 \leq x \leq q-1}$ and $V = (V_x)_{0 \leq x \leq q-1}$ be the vectors defined as

\[
U_x = 1 - \theta \left( \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) R_i(\theta) + \sum_{j=1}^{s} \frac{R_{\ell+j}(\theta)}{a_j \text{ ends with } x} \right),
\]

\[
V_x = 1 - \theta \left( \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) S_i(\theta) + \sum_{j=1}^{s} \frac{S_{\ell+j}(\theta)}{a_j \text{ begins with } x} \right),
\]

where \( \theta \) is the Perron root of $A$. Then $U$ and $V$ are left and right Perron eigenvectors, respectively, of $A$.

**Remark 5.6.** Let $A$ be the adjacency matrix associated with the collections $\mathcal{F}$ and $\mathcal{R}$. Then $A$ is indexed by $\mathcal{L}_{p-1}$. Let $\mathcal{G}_A = (\mathcal{L}_{p-1}, \mathcal{E}_A)$ denote the graph where $\mathcal{E}_A = \{(X \ast Y)_i : X, Y \in \mathcal{L}_{p-1}, 1 \leq i \leq A_{XY}\}$ is the edge set of $\mathcal{G}_A$. Let $\Sigma_A$ denote the edge shift associated to the matrix $A$. Let us now consider a binary matrix $E$, indexed by $\mathcal{E}_A$. For $(X \ast Y)_i, (W \ast Z)_j \in \mathcal{E}_A$, $E \left( (X \ast Y)_i, (W \ast Z)_j \right) = 1$ if and only if $Y = W$. If $\Sigma^e_E$ denotes the vertex shift associated with $E$, then $\Sigma^e_E = \Sigma_A$. Note that $A$ and $E$ have the same Perron root, say \( \theta \). If $U, V$ denote the left and right Perron (column) eigenvectors of $A$ and $\hat{U}, \hat{V}$ denote the left and right Perron (column) eigenvectors of $E$, respectively, such that $U^T V = 1 = \hat{U}^T \hat{V}$, then we have for $X \in \mathcal{L}_{p-1}$,

\[
V_X = \sum_{Y \in \mathcal{L}_{p-1}} \sum_{1 \leq i \leq k(X \ast Y)} \hat{V}_{(X \ast Y)_i},
\]

\[
\theta U_X = \sum_{Z \in \mathcal{L}_{p-1}} \sum_{1 \leq i \leq k(Z \ast X)} \hat{U}_{(Z \ast X)_i}.
\]
Hence the Perron root and eigenvectors of $E$ (equivalently, the Perron root and eigenvectors of $A$) can be calculated using the techniques discussed in [2], since $E$ is irreducible whenever $A$ is irreducible. Here, we get a combinatorial expression for $\theta$ and $U$ and $V$ in terms of the correlation between the words corresponding to the zero entries in $E$. Note that the zero entries in $E$ correspond to all the pairs of non-adjacent edges in $G_A$. It is usually a much bigger collection, especially when the size of $L_{p-1}$ is large. Hence the results presented in this paper are computationally less expensive in most such cases. Other advantages of looking at $A$ instead of $E$ are discussed in [33].

6. Normalized Perron Eigenvectors of $A$

Let $A$ be the adjacency matrix associated with given $F$ and $R$ and $p$ denote the length of the longest word in $F \cup R$. In this section, we assume that all words in $R$ are of length $p$. Then $U$ and $V$ as obtained in Theorem 5.3 are left and right Perron (column) eigenvectors of $A$. We will obtain a simple expression for the normalizing factor $U^T V$. This will then be used to obtain a combinatorial expression for the Parry measure on the edge shift associated with $A$.

In [7], subshifts with no repeated words (multiple edges) were considered and the concept of local escape rate was used to compute $U^T V$. We will use similar techniques here. The results in this section require $\Sigma_A$ to be an irreducible subshift with positive topological entropy, that is, the Perron root $\theta$ is strictly bigger than one. We will now define the concept of escape rate in the setting of an edge shift $\Sigma_A$. Let $\mu$ be the Parry measure on $\Sigma_A$.

**Definition 6.1** (Escape rate). Let $W$ be an allowed word in $\Sigma_A$. Consider the cylinder $C_W$ in $\Sigma_A$ based at $W$, which is the collection of all sequences which begin with $W$. The escape rate into the hole $C_W$ measures the rate at which the orbits escape into the hole $C_W$ and is defined as

$$\rho(C_W) := - \lim_{n \to \infty} \frac{1}{n} \ln \mu(W_n(W)),$$

if the limit exists, where $W_n(W)$ denotes the collection of all sequences in $\Sigma_A$ which do not include $W$ as a subword in their first $n$ positions.

In general, the escape rate is defined for any hole of positive measure in $\Sigma_A$, but we only consider that the hole is a cylinder. In the given setting, the limit exists and is given by the following result. Let $h_W(n)$ be the number of allowed words of length $n$ in $\Sigma_A$ that do not contain $W$ as a subword, let $\ln(\lambda_W) = \lim_{n \to \infty} \frac{1}{n} \ln(h_W(n))$, and let $\theta$ be the Perron root of $A$. Then we have the following result which is a direct generalization of [16] Theorem 3.1.

**Theorem 6.2.** The escape rate into the hole $C_W$ satisfies $\rho(C_W) = \ln(\theta/\lambda_W) > 0$.

For each word $W = (X_1 \ast X_2)_{i_1}(X_2 \ast X_3)_{i_2} \ldots (X_{n-p+1} \ast X_{n-p+2})_{i_{n-p+1}}$, of length $n-p+1$ in $\Sigma_A$, there is a unique word $w = x_1 \ldots x_n$ of length $n$ in $\Sigma_F$ where $X_i = x_i \ldots x_{i+p-2}$. Also, if $m(w) = 1$, then $h_W(n-p+1) = \tau_w(n)$, where $\tau_w(n)$ is the number of allowed words of length $n$ in $\Sigma_{F \cup \{w\}}$ counted with multiplicity with respect to $\mathcal{R}$. Hence, $\lambda_W = \theta_w$, where $\ln \theta_w = \lim_{n \to \infty} \frac{1}{n} \ln \tau_w(n)$ can be calculated using the correlation between the words from $F \cup \{w\}$ and $\mathcal{R}$ (as in Theorem 5.1). Note that $m(w) = 1$ implies $F \cup \{w\} \cup \mathcal{R}$ is reduced.

**Remark 6.3.** If $w$ contains a repeated word, then $\lambda_W$ and $\theta_w$ may not be equal. For example, let $A = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ and $G_A = \{0, 1\}, \mathcal{E}_A$ where $\mathcal{E}_A = \{(01)_1, (01)_2, (10)_1, (11)_1\}$. If $W = (01)_2(11)_1$, then $w = 011$, and $\tau_w(3) = \{(011, 2), (101, 2), (110, 1), (111, 1)\} = 6$. 

By Theorem 5.1, the Perron root of \(A\) if it exists, where \(D\) and \(\chi\) then the word \(X, Y\) with \(X, Y\) starts with \(Z\) and ends with \(W\). Moreover, \(\alpha = \frac{1}{\theta}\).

**Definition 6.4** (Local escape rate). Let \(\alpha = \alpha_1 \alpha_2 \cdots \in \Sigma_A\). The local escape rate around \(\alpha\) is defined as
\[
\rho(\alpha) = \lim_{n \to \infty} \frac{\rho(C_{W_n})}{\mu(C_{W_n})},
\]
if it exists, where \(W_n = \alpha_1 \alpha_2 \cdots \alpha_n\).

For fixed \(X, Y \in \mathcal{L}_{p-1}\), choose a point \(\alpha \in \Sigma_A\) such that there exists a subsequence \((n_k)\) where for each \(k\), \(w_{n_k}\) (associated with \(W_{n_{k}-p+1}\)) is a word that begins with \(X\) and ends with \(Y\) (in other words \(W_{n_{k}-p+1}\) is a path from vertex \(X\) to vertex \(Y\) in the graph \(G_A\)). Such an \(\alpha\) exists as \(A\) is irreducible. We assume that there exist \(X, Y \in \mathcal{L}_{p-1}\) such that the word \(w_{n_k}\) has multiplicity 1 for all \(k\). Using Theorem 6.2 and the expression for the Parry measure, we get that
\[
\rho(\alpha) = \lim_{k \to \infty} \frac{U^TV \theta^{n_k-p+1} \ln(\theta/\lambda_{W_{n_k}-p+1})}{U_X V_Y} = \frac{U^TV}{\theta^{p-1}U_X V_Y} \lim_{k \to \infty} \theta^{n_k} \ln(\theta/\theta_{w_{n_k}}),
\]
where \(U\) and \(V\) are the Perron eigenvectors of \(A\) as given in Theorem 5.3.

**Remark 6.5.** (1) If \(\alpha = (X_1 * X_2)_1(X_2 * X_3)_2 \cdots \in \Sigma_A\) is periodic with period \(t\), then the word \(\chi = X_1 * X_2 * X_3 * \cdots \in \Sigma_F\) is periodic with period \(t\). Using this and the local escape rate formula given by Ferguson and Pollicott in [12, Corollary 5.4.], \(\lim_{k \to \infty} \theta^{-k+1}(w_k, \theta) = \frac{1}{\theta(\alpha)}\).

**Definition 6.6** (Property (P)). We say that the subshift \(\Sigma_A\) satisfies property (P) if there exist \(X, Y \in \mathcal{L}_{p-1}\), a point \(\alpha = \alpha_1 \alpha_2 \cdots \in \Sigma_A\), and a strictly increasing sequence \((n_k)_{k \geq 1}\) of natural numbers such that for each \(k \geq 1\), \(W_{n_k-p+1} = \alpha_1 \cdots \alpha_{n_k-p+1}\) begins with \(X\) and ends with \(Y\) and \(m(w_{n_k}) = 1\), where \(w_{n_k}\) is the word in \(\Sigma_F\) associated to \(W_{n_k-p+1}\).

**Remark 6.7.** For property (P) to be satisfied, it is enough to show the existence of two words \(X, Y \in \mathcal{L}_{p-1}\) for which there are two words \(Z, W \in \mathcal{L}\), each with multiplicity 1, such that the word \(Z\) starts with \(X\) and ends with \(Y\), and the word \(W\) both starts and ends with \(Y\). A large family of subshifts satisfy property (P). For instance, if at least one vertex in \(G_A\) has a single loop (in other words, the adjacency matrix \(A\) has at least one diagonal entry equal to 1), then \(\Sigma_A\) satisfies property (P).

Let \(P(z)\) be the matrix function corresponding to the collections \(\mathcal{F}\) and \(\mathcal{R}\) as defined in (11) and let \(R_i(z)\) be the \(i^{th}\) row sum of \(P^{-1}(z)\). Define rational function \(R(z)\) as follows:
\[
R(z) = z \sum_{i=1}^{\ell} \left(1 - \frac{1}{m_i}\right) R_i(z) - z \sum_{j=1}^{s} R_{\ell+i}(z), \tag{31}
\]
By Theorem 5.1, the Perron root of \(A\) is given by the largest real zero of \(z - q + R(z)\). Note that \(R(z) = \sum DP^{-1}\), where \(\sum\) denotes the sum of all the entries of the matrix and \(D\) is the diagonal matrix as defined in Equation (13).

Now assume that \(\Sigma_A\) satisfies property (P). With notations as in the definition of property (P), \(w_{n_k}\) starts with \(X\) and ends with \(Y\), for all \(k \geq 1\). Denote \(\mathcal{F}_k = \mathcal{F} \cup \{w_{n_k}\}\), then \(\mathcal{F}_k \cup \mathcal{R}\) is reduced for all \(k \geq 1\). Suppose \(P_k(z)\) be the matrix function (see (11)) and \(R_k(z)\) be the rational function (see (31)) associated with the collections \(\mathcal{F}_k\) and \(\mathcal{R}\).
Let $D_k(z)$ and $D(z)$ denote the determinant of $P_k(z)$ and $P(z)$, respectively. Then

$$P_k(z) = \begin{pmatrix} P(z) & -zY^T(z) \\ X(z)D & Z(z) \end{pmatrix},$$

where $Z(z) = (w_{n_k}, w_{n_k})_z$, $X^T(z) = \begin{pmatrix} (r_1, w_{n_k})_z \\ (\hat{r}_1, X)_z \\ \vdots \\ (\hat{a}_s, X)_z \end{pmatrix}$ and $Y^T(z) = \begin{pmatrix} (w_{n_k}, r_1)_z \\ (w_{n_k}, \hat{r}_1)_z \\ \vdots \\ (w_{n_k}, a_s)_z \end{pmatrix}$.

where for a word $u$, $\hat{u}$ is the word obtained by removing the first symbol of $u$. Observe that if $U_X(z) = 1 - \sum XDP^{-1}(z)$ and $V_Y(z) = 1 - \sum Y(DP^{-1})^T(z)$, then $U_X(\theta)$ and $V_Y(\theta)$ are the Perron eigenvectors obtained as in Theorem 5.3. Using the inverse formula for a $2 \times 2$ block matrix, we get

$$R_k(z) - R(z) = -\frac{zD(z)}{D_k(z)} U_X(z)V_Y(z).$$

Then we have the following result, proof of which is similar to the proof of [7, Theorem 7.6], using (32) and Remark 6.5(1).

**Theorem 6.8.** With the notations as above, if $\Sigma_A$ satisfies property (P), then

$$U^TV = \theta^{p-1}(1 + R'(\theta)),$$

where $R'(\theta)$ is the derivative of the function $R(z)$ (defined in 31) evaluated at $z = \theta$.

**Example 6.9.** Let $\Sigma = \{0, 1\}, \mathcal{F} = \{010\}, \mathcal{R} = \{100(3)\}$. Here $R(z) = \frac{2-z}{z^2+1+z^2}$. The Perron root of the associated adjacency matrix $A$ is $\theta = 2$. The left and right eigenvectors of $A$ is calculated in Example 5.3 and is given as $U = \begin{pmatrix} 3/2 \\ 1/2 \\ 1 \end{pmatrix}$, and $V = \begin{pmatrix} 2/3 \\ 2/3 \\ 4/3 \\ 4/3 \end{pmatrix}$. Note that $U^TV = \theta^2(1 + R'(\theta)) = 11/3$.  

**Example 6.10.** Let $\Sigma = \{0, 1, \ldots, q-1\}, \mathcal{F} = \{a_1, \ldots, a_s\} \mathcal{R} = \{r_1(m_1), \ldots, r_t(m_t)\}$ be such that $|a_i| = |r_i| = 2$ satifying $(r_i, r_i)_z = z, (a_j, a_j)_z = z$ and $(r_i, a_j)_z = (a_j, r_i)_z = 0$ for all $r_i \in \mathcal{R}, a_j \in \mathcal{F}$. Further, $(r_i, r_j)_z = (a_i, a_j)_z = 0$ for all $r_i \neq r_j \in \mathcal{R}$ and $a_i \neq a_j \in \mathcal{F}$. Then the Perron root $\theta$ is the largest real zero of

$$z - q - \left( \frac{\alpha + \ell - s}{z} \right)$$

where $\alpha = m_1 + \cdots + m_t$. Therefore $\theta = \frac{\alpha^2 + \sqrt{\alpha^2 + 4(\alpha - \ell + s)}}{2}$. For instance, consider $q = 4, \mathcal{R} = \{10(a), 20(b), 30(c)\}, \mathcal{F} = \emptyset$. Set $\alpha = a + b + c$. Then the Perron root $\theta$ is the largest real zero of

$$z - 4 - \left( \frac{\alpha - 3}{z} \right).$$
Therefore \( \theta = 2 + \sqrt{1 + \alpha} \). Also in this case,
\[
U = \frac{1}{2 + \sqrt{1 + \alpha}} \begin{pmatrix} \alpha - 1 + \sqrt{1 + \alpha} \\ 2 + \sqrt{1 + \alpha} \\ 2 + \sqrt{1 + \alpha} \\ 2 + \sqrt{1 + \alpha} \end{pmatrix}, \quad V = \frac{1}{2 + \sqrt{1 + \alpha}} \begin{pmatrix} 2 + \sqrt{1 + \alpha} \\ a + 1 + \sqrt{1 + \alpha} \\ b + 1 + \sqrt{1 + \alpha} \\ c + 1 + \sqrt{1 + \alpha} \end{pmatrix}.
\]

Since property (P) is satisfied by the adjacency matrix for the corresponding shift space, 
\( U^T V = \theta(1 + R'(\theta)) \), where \( R(z) = \frac{\alpha - 3}{z} \). Hence
\[
U^T V = \frac{5 - \alpha + \sqrt{1 + \alpha}}{2 + \sqrt{1 + \alpha}}.
\]

This normalization result immediately gives an alternate definition for the Parry measure on \( \Sigma_A \) and is stated below as a corollary.

**Corollary 6.10.1** (A combinatorial expression for the Parry measure). **With the notations as above, assume that \( \Sigma_A \) satisfies property (P). If \( W = (X_1 * X_2)_1 \ldots (X_n * X_{n+1})_i \) is an allowed word in \( \Sigma_A \), then the Parry measure of the cylinder based at \( W \) is given by,
\[
\mu(C_W) = \frac{U_{X_1} V_{X_{n+1}}}{\theta^{n+p-1}(1 + R'(\theta))},
\]
where \( U \) and \( V \) are as given in equations (20) and (21).

**Remark 6.11.** As discussed in Remark 6.5.2, the above mentioned techniques cannot be used if \( \Sigma_A \) does not satisfy property (P). However, we conjecture that Corollary 6.10.1 holds true provided \( \mathcal{F} \cup \mathcal{R} \) is reduced and all words from \( \mathcal{R} \) have length \( p \). That is, if \( U \) and \( V \) denote the Perron eigenvectors of \( A \) as given in Theorem 5.3 then \( U^T V = \theta^{p-1}(1 + R'(\theta)) \). We are tempted to give an example where the conjecture is true. Consider the shift \( \Sigma_A \) with \( A = \begin{pmatrix} 0 & 2 \\ 3 & 2 \end{pmatrix} \). Here \( \Sigma = \{0, 1\} \), \( \mathcal{F} = \{00\} \) and \( \mathcal{R} = \{01(2), 10(3), 11(2)\} \).

Clearly property (P) is not satisfied as, there does not exist a point in \( \Sigma_\mathcal{F} \) that contains no repeated words. Here we get \( F(z) = \frac{z}{z - q + R(z)} \), where \( R(z) = \frac{-3(z + 2)}{z^2 - 3} \). Hence the Perron root of \( A \) is given by \( \theta = 1 + \sqrt{7} \). Also, by Theorem 5.3 the left and right eigenvalues of \( A \) is given by \( U = \begin{pmatrix} \sqrt{7} - 1 \\ 2 \end{pmatrix} \) and \( V = \begin{pmatrix} 2(\sqrt{7} - 2) \\ 5 - \sqrt{7} \end{pmatrix} \), respectively. Note that \( U^T V = \theta(1 + R'(\theta)) = 28 - 8\sqrt{7} \).

7. \( \Sigma_\mathcal{F} \) AS A FACTOR OF \( \Sigma_A \)

In this section, we study the edge shift \( \Sigma_\mathcal{F} \) as a factor of the edge shift \( \Sigma_A \) and discuss some properties of Markov measures on these spaces. Let \( \Sigma_\mathcal{F}, \mathcal{R} \) and \( p \) as before and \( A \) be the adjacency adjacency matrix associated with \( \mathcal{F} \) and \( \mathcal{R} \). Let \( \hat{A} \) be the binary matrix indexed by \( \mathcal{L}_{p-1} \), compatible with \( A \), that is, \( \hat{A}_{XY} = 1 \) if and only if \( A_{XY} > 0 \). Let \( \Sigma_A \) and \( \Sigma_\hat{A} \) be the edge shifts associated with the matrices \( A \) and \( \hat{A} \), respectively. Note that \( \Sigma_\hat{A} = \Sigma_{\mathcal{F}^{[p]}} \). Observe that if \( \mathcal{R} \) is an empty collection, then \( A \) and \( \hat{A} \) are the same. However, it was observed earlier, that \( \Sigma_A \) and \( \Sigma_\hat{A} \) are not conjugate to each other if the collection \( \mathcal{R} \) is non-empty.

Let \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \) and \( \hat{\mathcal{G}} = (\hat{\mathcal{V}}, \hat{\mathcal{E}}) \) be the digraphs associated to the matrices \( A \) and \( \hat{A} \), respectively, both having the same vertex sets, \( \mathcal{V} = \hat{\mathcal{V}} = \mathcal{L}_{p-1} \), but (possibly) different edge sets. For \( X, Y \in \mathcal{L}_{p-1} \), let \( \mathcal{E}_{XY} \) and \( \hat{\mathcal{E}}_{XY} \) denote the set of edges from \( X \).
to $Y$ in $\mathcal{G}$ and $\hat{\mathcal{G}}$, respectively. If $X \ast Y \in \mathcal{L}_p$, then $\hat{\mathcal{E}}_{XY} = \{X \ast Y\}$ and $\mathcal{E}_{XY} = \{(X \ast Y)_j : 1 \leq j \leq k(X \ast Y)\}$, else, $\hat{\mathcal{E}}_{XY} = \mathcal{E}_{XY} = \emptyset$.

**Definition 7.1** (Projection map for $\mathcal{G}$ and $\hat{\mathcal{G}}$). With notations as above, a projection map $\pi : \mathcal{G} \to \hat{\mathcal{G}}$ is defined as follows. On vertex sets, $\pi(\mathcal{V}) = \hat{\mathcal{V}} = \mathcal{V}$ is an identity map and on edge sets, $\pi(\mathcal{E}) = \hat{\mathcal{E}}$, with $\pi(\mathcal{E}_{XY}) = \hat{\mathcal{E}}_{XY}$ for all $X, Y \in \mathcal{V}$.

The projection map $\pi$ identifies all the edges in $\mathcal{G}$ from vertex $X$ to $Y$ as a single edge. Whenever $X \ast Y$ is allowed, all the edges from $X$ to $Y$ in $\mathcal{G}$ are projected onto the only edge $X$ to $Y$ in $\hat{\mathcal{G}}$.

**Example 7.2.** We look at a simple example to understand the map $\pi$. Let $q = 2$, $\Sigma = \{0, 1\}$. Let $\mathcal{F} = \{11\}$ and a matrix $A = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}$. Then $\hat{A} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Consider the digraphs $\mathcal{G}$ and $\hat{\mathcal{G}}$ defined by the adjacency matrices $A$ and $\hat{A}$ as depicted in Figure 6.

The vertex set for both the graphs is $\mathcal{V} = \{0, 1\}$. Let us name the edges in $\mathcal{G}$ as $a, b, \ldots, f$ as shown in the figure. Then $\pi(\{a, b, c\}) = \{00\}, \pi(\{d\}) = \{01\}, \pi(\{e, f\}) = \{10\}$.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{graph.png}
\caption{Graph $\mathcal{G}$ on the left and graph $\hat{\mathcal{G}}$ on the right.}
\end{figure}

The definition of the projection map $\pi$ can be extended to finite paths as well. A word $W = (X_1 \ast X_2)_{i_1} (X_2 \ast X_3)_{i_2} \ldots (X_n \ast X_{n+1})_{i_n}$ denotes a path from $X_1$ to $X_{n+1}$ of length $n$ in $\mathcal{G}$. As $i_1, i_2, \ldots, i_n$ vary over all the possible choices, we get all the paths from $X_1$ to $X_{n+1}$ with the fixed intermediate vertices (in order) $X_2, \ldots, X_n$. We define $\pi$ to map all these different paths in $\mathcal{G}$ onto a fixed path in $\hat{\mathcal{G}}$ as,

$$\pi ((X_1 \ast X_2)_{i_1} (X_2 \ast X_3)_{i_2} \ldots (X_n \ast X_{n+1})_{i_n}) := (X_1 \ast X_2) (X_2 \ast X_3) \ldots (X_n \ast X_{n+1}),$$

for all $1 \leq i_j \leq k(X_j \ast X_{j+1}), 1 \leq j \leq n$.

The shifts $\Sigma_A$ and $\Sigma_{F}^{[p]}$ represent the spaces of all one-sided infinite paths on the graphs $\mathcal{G}$ and $\hat{\mathcal{G}}$, respectively. The map $\pi$ can be further extended to $\pi : \Sigma_A \to \Sigma_{F}^{[p]}$ as follows.

$$\pi ((X_1 \ast X_2)_{i_1} (X_2 \ast X_3)_{i_2} (X_3 \ast X_4)_{i_3} \ldots) := (X_1 \ast X_2)(X_2 \ast X_3)(X_3 \ast X_4)\ldots, \quad (33)$$

for all $1 \leq i_j \leq k(X_j \ast X_{j+1}), j \geq 1$.

**Proposition 7.3.** The shift $\Sigma_F$ is a factor of the edge shift $\Sigma_A$ with factor map $\pi$.

**Proof.**

$$\Sigma_A \xrightarrow{\sigma_A} \Sigma_A \quad \downarrow \pi \quad \downarrow \pi \quad \Sigma_{F}^{[p]} \xrightarrow{\sigma} \Sigma_{F}^{[p]}$$

We first recall that $\pi : \Sigma_A \to \Sigma_F$ is called a *factor map* or *projection* if $\pi$ is onto and the above diagram commutes. The proof now follows directly from the definitions of $\pi$ as given in (33) and of the left shift maps $\sigma_A$ and $\sigma$. \qed
We now recall some basic concepts related to a Markov measure on a shift space. A square matrix is called a stochastic matrix if it is non-negative, with each row sum being 1. Let $\hat{P}$ be a stochastic matrix compatible with $\hat{A}$. Then 1 is the Perron root of $\hat{P}$ with right Perron eigenvector given by $(1, 1, \ldots, 1)^T$. Let $\hat{\rho}$ be the normalized left Perron eigenvector of $\hat{P}$. A Markov chain corresponding to $\hat{P}$ on the graph $\hat{G}$ is the assignment of probability to the vertices given by $\nu(X) := \hat{\rho}_X$, and to the edges given by $\nu(X * Y | X) := \hat{P}_{XY}$ which is the conditional probability of the edge $X * Y$. The probability of the edge $X * Y$ is then given by $\nu(X * Y) = \hat{\rho}_X \hat{P}_{XY}$. The probability of a path is expressed as the product of probability of the initial vertex and conditional probabilities of the edges in the path.

Let $\hat{W} = (X_1 * X_2)(X_2 * X_3) \ldots (X_n * X_{n+1})$ be an allowed word of length $n$ in $\Sigma_\hat{F}^{[p]}$ and let $C_W$ be the cylinder in $\Sigma_\hat{F}^{[p]}$ based at $\hat{W}$. The stochastic matrix $\hat{P}$ defines a Markov measure $\nu$ on $\Sigma_\hat{F}^{[p]}$ which is described for the cylinder sets as follows,

$$\nu(C_W) = \hat{\rho}_{X_1} \hat{P}_{X_1X_2} \hat{P}_{X_2X_3} \ldots \hat{P}_{X_nX_{n+1}}.$$

The Parry measure on $\Sigma_\hat{F}^{[p]}$ defined in Section 2 is a special kind of Markov measure. The stochastic matrix associated with it is called the Parry matrix, and is given as follows.

**Definition 7.4.** The Parry matrix $\hat{P}$ associated to a binary matrix $\hat{A}$ is a stochastic matrix compatible with $\hat{A}$ and is defined as

$$\hat{P}_{XY} := \frac{\hat{A}_{XY} \hat{V}_Y}{\hat{\theta}_X \hat{V}_X}, \quad \text{for } X, Y \in \mathcal{L}_{p-1},$$

where $\hat{\theta}$ is the Perron root and $\hat{U}$ and $\hat{V}$ are the left and right Perron (column) eigenvectors of $\hat{A}$ such that $\hat{U}^T \hat{V} = 1$. The normalized left Perron eigenvector $\hat{\rho}$ of $\hat{P}$ is given by

$$\hat{\rho}_X = \hat{U}_X \hat{V}_X.$$

Let $A$ be the adjacency matrix associated with $\mathcal{F}$ and $\mathcal{R}$. A Markov measure on $\Sigma_A$ is defined in a similar way. Let $\mathcal{P}$ be a stochastic matrix compatible with $A$ and $\rho$ be its normalized left Perron eigenvector. A Markov chain $\mu$ associated to $\mathcal{P}$ on the graph $\mathcal{G}$ (associated to matrix $A$) is given by $\mu(X) = \rho_X$ and $\sum_{1 \leq i \leq k(X * Y)} \mu((X * Y)_i | X) = \mathcal{P}_{XY}$.

Here the sum of the conditional probabilities of all the edges from $X$ to $Y$ is the $XY^{th}$ entry of $\mathcal{P}$. Let $C_W$ be the cylinder in $\Sigma_A$ based at a word $W = (X_1 * X_2)_{i_1}(X_2 * X_3)_{i_2} \ldots (X_n * X_{n+1})_{i_n}$ in $\Sigma_A$. The stochastic matrix $\mathcal{P}$ gives rise to a Markov measure on $\Sigma_A$, which is described for cylinder sets as follows: for a cylinder $C_W$,

$$\mu(C_W) = \mu(X_1) \mu((X_1 * X_2)_{i_1} | X_1) \mu((X_2 * X_3)_{i_2} | X_2) \ldots \mu((X_n * X_{n+1})_{i_n} | X_n).$$

Consider the preimage of the word $\hat{W} = (X_1 * X_2) \ldots (X_n * X_{n+1})$ in $\Sigma_\hat{F}^{[p]}$ under $\pi$, which is the set of all paths from $X_1$ to $X_{n+1}$ with the fixed intermediate vertices $X_2, \ldots, X_n$, in order, in the graph $\mathcal{G}$. That is,

$$\pi^{-1}(\hat{W}) = \{(X_1 * X_2)_{i_1} \ldots (X_n * X_{n+1})_{i_n} | 1 \leq i_j \leq k(X_j * X_{j+1}), 1 \leq j \leq n \}.$$ 

The Markov measure of the set of all cylinder sets based at words $W \in \pi^{-1}(\hat{W})$ is given as

$$\mu \left( \bigcup_{W \in \pi^{-1}(\hat{W})} C_W \right) = \sum_{W \in \pi^{-1}(\hat{W})} \mu(C_W).$$
Definition 7.5. The Shannon-Parry matrix $U$ and Shannon-Parry measure $\rho$ Perron eigenvector $\nu$ is such as a push forward measure under the projection $\pi$.

Remark 7.6. Let $W = (X_1 \ast X_2)_{i_1} (X_2 \ast X_3)_{i_2} \ldots (X_n \ast X_{n+1})_{i_n}$ be a word of length $n$ in $\Sigma_A$. The Shannon-Parry measure $\mu$ of $C_W$ is given in terms of the Shannon-Parry matrix $\mathcal{P}$ as

$$\mu(C_W) = \rho_{X_1} \frac{\mathcal{P}_{X_1 X_2}}{\theta X_2} \frac{\mathcal{P}_{X_2 X_3}}{\theta X_3} \ldots \frac{\mathcal{P}_{X_n X_{n+1}}}{\theta X_{n+1}}.$$ 

Note that if $W' = (X_1 \ast X_2)_{j_1} (X_2 \ast X_3)_{j_2} \ldots (X_n \ast X_{n+1})_{j_n}$ is another path of length $n$ in $\mathcal{G}$, then $\mu(C_W) = \mu(C_{W'})$. In this case, all the edges from two fixed vertices are given equal conditional probabilities which need not be the case for a general Markov measure on $\Sigma_A$.

Theorem 7.7. Any Markov measure on $\Sigma_A$ induces a Markov measure on $\Sigma_{\mathcal{F}}^{[p]}$ as a push forward measure under $\pi$. That is, for every Markov measure $\mu$ on $\Sigma_A$, there exists a Markov measure $\nu$ on $\Sigma_{\mathcal{F}}^{[p]}$ such that for any cylinder $C_W$ in $\Sigma_{\mathcal{F}}^{[p]}$ based at a word $\tilde{W} \in \mathcal{L}$,

$$\nu(C_W) = \mu \left( \pi^{-1}(C_W) \right).$$

Proof. Let $\tilde{W} = (X_1 \ast X_2)(X_2 \ast X_3) \ldots (X_n \ast X_{n+1})$ be a word of length $n$ in $\Sigma_{\mathcal{F}}^{[p]}$ and let $C_W$ be the cylinder in $\Sigma_{\mathcal{F}}^{[p]}$ based at $\tilde{W}$. Let $\mathcal{P}$ be any stochastic matrix compatible with $A$ and let $\mu$ be the corresponding Markov measure on $\Sigma_A$. Observe that $\mathcal{P}$ is a stochastic matrix compatible with $\hat{A}$ as well, and thus it defines a Markov measure $\nu$ on $\Sigma_{\mathcal{F}}^{[p]}$ as

$$\nu(C_W) = \rho_{X_1} \mathcal{P}_{X_1 X_2} \mathcal{P}_{X_2 X_3} \ldots \mathcal{P}_{X_n X_{n+1}}.$$ 

Note that $\mu \left( \pi^{-1}(C_W) \right) = \mu \left( \bigcup_{W \in \pi^{-1}(\tilde{W})} C_W \right) = \rho_{X_1} \mathcal{P}_{X_1 X_2} \mathcal{P}_{X_2 X_3} \ldots \mathcal{P}_{X_n X_{n+1}}$, which is same as $\nu(C_{W'})$. Hence the result.

Since the Shannon-Parry measure on $\Sigma_A$ is a Markov measure, we obtain the following corollary.

Corollary 7.7.1. The Shannon-Parry measure on $\Sigma_A$ induces a Markov measure on $\Sigma_{\mathcal{F}}$ as a push forward measure under the projection $\pi$. 


As evident from the above proof, a sufficient condition for a Markov measure on $\Sigma_f^{[p]}$ to be a push forward of the Shannon-Parry measure on some $\Sigma_A$ under $\pi$, is that the stochastic matrices corresponding the two measures should be the same. An intriguing question now arises is that, given any Markov measure on $\Sigma_f^{[p]}$, can it be obtained as a push forward of a Shannon-Parry measure on $\Sigma_A$, for some non-negative irreducible integer matrix $A$. We show that this is not true in general. We further obtain certain conditions on the Markov measure under which this holds true.

**Theorem 7.8.** Suppose $\nu$ is a Markov measure on $\Sigma_f^{[p]}$ such that the associated stochastic matrix $P_\nu$ of $\nu$ has all rational entries. Then there exists a non-negative irreducible integer matrix $A$ such that $\nu$ is the push-forward of the Shannon-Parry measure on $\Sigma_A$.

**Proof.** Given that $P_\nu$ is a rational matrix, any nonzero $XY$th entry of $P_\nu$ is $\frac{d_{XY}}{d_{XY}}$ (reduced form). Set $L = \text{lcm}\{d_{XY} : (P_\nu)_{XY} > 0\}$. It can be verified with simple calculations that the matrix $A := L P_\nu$ is the required matrix. Here $P_\nu$ itself is the Shannon-Parry matrix for $A$. We use the fact that 1 is the Perron root of a stochastic matrix with corresponding right eigenvector $(1, 1, \ldots, 1)$.

\[\square\]

**Example 7.9.** We give an example to show that the hypothesis given in Theorem 7.8 is not necessary. Let $\Sigma = \{0, 1\}$, and $F = \{11\}$. We give two examples of Markov measures $\nu_1$ and $\nu_2$ on $\Sigma_f^{[2]}$ with associated stochastic matrices $P_1$ and $P_2$, respectively, having irrational entries. We prove that $\nu_1$ is the push-forward for the Shannon-Parry measure on $\Sigma_A$ for some $A$ whereas $\nu_2$ is not.

1. Let $P_1 = \left(\begin{array}{cc} 2(\sqrt{2} - 1) & 3 - 2\sqrt{2} \\ 1 & 0 \end{array}\right)$. Note that $\nu_1$ is a push-forward of the Shannon-Parry measure on $\Sigma_A$ where $A = \left(\begin{array}{cc} 2 & 1 \\ 1 & 0 \end{array}\right)$.

2. Let $P_2 = \left(\begin{array}{cc} 1/\pi & 1 - 1/\pi \\ 1 & 0 \end{array}\right)$. If there exists a non-negative integer matrix $A = \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right)$ such that $P_{XY} = \frac{A_{XY}v_Y}{v_X}$ where $\theta$ is the Perron root of $A$ and $V$ is the right Perron eigenvector of $A$. We have, in particular, $\frac{1}{\pi} = P_{11} = \frac{A_{11}}{\theta}$ which gives $\theta = A_{11}\pi$. Note that $\theta$ is an algebraic number as it satisfies the characteristic equation of $A$ which gives a contradiction.

**Theorem 7.10.** Let $F$ be a given collection of forbidden words and $p$ be the length of the longest word in $F$. Set $R = L_p$, such that all words in $R$ are assigned equal multiplicities, say $M$. Let $A$ be the adjacency matrix associated with $F$ and $R$, and $\Sigma_A$ be the corresponding edge shift. Then the Parry measure on $\Sigma_f^{[p]}$ is the push forward of the Shannon-Parry measure on $\Sigma_A$.

**Proof.** Let $\hat{A}$ be the binary matrix compatible with $A$ and $\hat{P}$ be the associated Parry matrix. It is given that $A = MA\hat{A}$. If $P$ denotes the Shannon-Parry matrix associated to $A$, then simple calculations show that $\hat{P} = P$. That is, the same stochastic matrix induces the Parry measure on $\Sigma_f^{[p]}$ and the Shannon-Parry measure on $\Sigma_A$. Hence the result follows. \[\square\]

**Example 7.11.** We give an example to show that the hypothesis given in Theorem 7.10 is not necessary. Let $\Sigma = \{0, 1\}$ and $F = \{11\}$. The adjacency matrix of $\Sigma_f^{[2]}$ is given
by $\hat{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ which has Perron root $\hat{\theta} = (1 + \sqrt{5})/2$ and right Perron eigenvector $\hat{V} = \begin{pmatrix} (1 + \sqrt{5})/2 \\ 1 \end{pmatrix}$. Let $\mathcal{R} = \{00(2), 10(4)\}$. Note that $\mathcal{R}$ does not satisfy the conditions as in Theorem 7.10. The adjacency matrix associated with $\mathcal{F}$ and $\mathcal{R}$ are $A = \begin{pmatrix} 2 & 1 \\ 4 & 0 \end{pmatrix}$ which has Perron root $\theta = 1 + \sqrt{5}$ and right Perron eigenvector $V = \begin{pmatrix} (1 + \sqrt{5})/4 \\ 1 \end{pmatrix}$. Note that for $X, Y \in \{0, 1\}$, $A_{XY}V_Y = \hat{A}_{XY}\hat{V}_Y\hat{V}_X$.

8. Concluding remarks and Potential questions

In this section, we discuss some immediate consequences of our work and also potential questions that emerge from this study.

The methods proposed in this paper for computing the Perron root and eigenvectors (Theorems 5.1, 5.3) are particularly useful when the matrices $P, Q$ are sparse. That is, a lot of pairs of distinct words in $\mathcal{F} \cup \mathcal{R}$ have zero cross-correlations (i.e., $(u, w)_z = 0$ for all $u \neq w \in \mathcal{F} \cup \mathcal{R}$). For instance, if all the cross-correlations are zero, then the matrices $P, Q$ take the form of a diagonal matrix. Hence it is much easier to use our methods than the traditional methods. Here the rational function $F(z)$ of which the Perron root $\theta$ of the adjacency matrix $A$ is the largest real pole and also the expressions of the Perron eigenvectors take a simple form since

$$R_i(z) = Q_i(z) = \begin{cases} 
(z \left(1 - \frac{1}{m_i}\right)(r_i, r_i)_z - z^{r_i})^{-1}, & 1 \leq i \leq \ell, \\
(-z(a_i, a_i)_z)^{-1}, & \ell + 1 \leq i \leq \ell + s.
\end{cases}$$

For an illustration, revisit Example 6.10. Moreover, this method is extremely useful when the matrix $A_{n \times n}$ differs from $I_{n \times n}$ (the matrix where all entries are 1) only on a few places, in which case, the matrices $P, Q$ are small and the calculations are simpler in terms on computational complexity. Also, these algorithms do not depend on the size of $A$ hence may have applications in network theory where usually matrices of large order are studied. Another interesting observation about our work is that the expressions of the Perron root and eigenvectors only depend on the correlation between the words corresponding to zero or greater than 1 entries in the matrix, which implies that when the correlations are the same, no matter where they are located in the matrix, these quantities are the same.

For a matrix $A$ with non-negative rational entries, choose the smallest $L > 0$ so that $LA$ is an integer matrix. Then if $\theta, U, V$ are the Perron root and left and right Perron eigenvectors of $A$, respectively, then $L\theta, U, V$ are the Perron root and left and right Perron eigenvectors of $LA$, respectively. Thus our results are easily extendable to matrices with non-negative rational entries.

Further, since the recurrence relations obtained in [13] (when the collection $\mathcal{R}$ is empty) are useful in non-transitive games, period prefix-synchronized coding, information theory or network theory, we believe that the generalized recurrence relations obtained in Section 4 may have several other independent applications.

There are some immediate potential questions that emerge from this study such as obtaining an expression for the Perron eigenvectors when $\mathcal{R}$ has words of different lengths or when $\mathcal{F} \cup \mathcal{R}$ is not reduced or proving the normalization result, as discussed in Section 6 when property (P) is not satisfied. Moreover, this work leads to the question of characterizing collections $\mathcal{F}$ and $\mathcal{R}$ that maximize the entropy of the shift $\Sigma_A$. 
References

[1] A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns: enumerative aspects. In Language and automata theory and applications, volume 10792 of Lecture Notes in Comput. Sci., pages 195–206. Springer, Cham, 2018.

[2] A. Asinowski, C. Banderier, and V. Roitner. Generating functions for lattice paths with several forbidden patterns. Sém. Lothar. Combin., 84B:Art. 95, 12, 2020.

[3] R. B. Bapat and T. E. S. Raghavan. Nonnegative matrices and applications, volume 64 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1997.

[4] C. Bonanno, G. Cristadoro, and M. Lenci. Maximal escape rate for shifts. Discrete Contin. Dyn. Syst., 42(12):6007–6029, 2022.

[5] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. Comput. Networks ISDN Syst., 30(1-7):107–117, 1998.

[6] R. Cerf and J. Dalmau. A Markov chain representation of the normalized Perron-Frobenius eigenvector. Electron. Commun. Probab., 22:Paper No. 52, 6, 2017.

[7] H. Cheriyath and N. Agarwal. On the Perron root and eigenvectors associated with a subshift of finite type. Linear Algebra Appl., 633:42–70, 2022.

[8] R. Dabrowski, G. Timoszkuk, and K. Stencel. One graph to rule them all: software measurement and management. Fund. Inform., 125(1-2):47–63, 2013.

[9] D. Dentchev. A method for computing the Perron root for primitive matrices. Numer. Linear Algebra Appl., 28(1):Paper No. e2340, 14, 2021.

[10] O. Elishco, T. Meyerovitch, and M. Schwartz. Semiconstrained systems. IEEE Trans. Inform. Theory, 62(4):1688–1702, 2016.

[11] O. Elishco, T. Meyerovitch, and M. Schwartz. On encoding semiconstrained systems. IEEE Trans. Inform. Theory, 64(4, part 1):2474–2484, 2018.

[12] A. Ferguson and M. Pollicott. Escape rates for Gibbs measures. Ergodic Theory Dynam. Systems, 32(3):961–988, 2012.

[13] L. J. Guibas and A. M. Odlyzko. Maximal prefix-synchronized codes. SIAM J. Appl. Math., 35(2):401–418, 1978.

[14] L. J. Guibas and A. M. Odlyzko. String overlaps, pattern matching, and nontransitive games. J. Combin. Theory Ser. A, 30(2):183–208, 1981.

[15] J. Hadamard. Les surfaces à courbures opposées et leurs lignes géodésique. J. Math. pures appl., 4.

[16] C. Haritha and N. Agarwal. Product of expansive Markov maps with hole. Discrete Contin. Dyn. Syst., 39(10):5743–5774, 2019.

[17] J. Hawkins. Ergodic dynamics-from basic theory to applications, volume 289 of Graduate Texts in Mathematics. Springer, Cham, [2021] © 2021.

[18] K. A. S. Immink. Codes for mass data storage systems. Shannon Foundation Publisher, 2004.

[19] N. Kashyap. Maximizing the Shannon capacity of constrained systems with two constraints. SIAM J. Discrete Math., 17(2):298–319, 2003.

[20] N. Keyfitz and H. Caswell. Applied mathematical demography, volume 47. Springer, 2005.

[21] D. Lind and B. Marcus. An introduction to symbolic dynamics and coding. Cambridge University Press, Cambridge, 1995.

[22] D. A. Lind. Perturbations of shifts of finite type. SIAM J. Discrete Math., 2(3):350–365, 1989.

[23] T.-T. Lu and S.-H. Shiou. Inverses of $2 \times 2$ block matrices. Comput. Math. Appl., 43(1-2):119–129, 2002.

[24] J. A. Marrero, J. Núñez Valdés, and M. T. Villar. Associating hub-directed multigraphs to arrowhead matrices. Math. Methods Appl. Sci., 41(6):2360–2369, 2018.

[25] M. Morse and G. A. Hedlund. Symbolic Dynamics. Amer. J. Math., 60(4):815–866, 1938.

[26] G. F. Newell and E. W. Montroll. On the theory of the ising model of ferromagnetism. Reviews of Modern Physics, 25(2):353, 1953.

[27] W. Parry. Intrinsic Markov chains. Trans. Amer. Math. Soc., 112:55–66, 1964.

[28] R. C. Penner. A construction of pseudo-Anosov homeomorphisms. Trans. Amer. Math. Soc., 310(1):179–197, 1988.

[29] D. Ruelle. Thermodynamic formalism. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2004.

[30] C. E. Shannon. A mathematical theory of communication. Bell System Tech. J., 27:379–423, 1948.

[31] S. Smale. Differentiable dynamical systems. Bull. Amer. Math. Soc., 73:747–817, 1967.
Proof of Theorem 4.14. To obtain the first equation, consider the word $wx$ obtained by adjoining a word $w$ in $f(n)$, and a symbol $x \in \Sigma$. Then, by similar arguments as in the proof of Theorem 4.3, we get

$$\sum_{w \in \mathcal{L}_n} \sum_{x \in \Sigma} m(wx) = qf(n) + \sum_{i=1}^{\ell} \frac{g_i(n+1)}{m_i} (m_i - 1).$$

A word $wx$ gets counted either in $f(n+1)$ or in $f_a(n+1)$ with extra counting and thus

$$qf(n) + \sum_{i=1}^{\ell} \frac{g_i(n+1)}{m_i} (m_i - 1) = f(n+1) + \sum_{i=1}^{s} f_{a_i}(n+1) + T(n). \tag{34}$$

Here the term $T(n)$ appears only when $wx$ ends with a forbidden word $a_i$, such that $a_i$ contains a repeated word $r_j$, except at the end. In this case, $\alpha \in (a_i, r_j)$ with $\alpha > |r_j|$. This implies $r_j$ is a subword of $w$ and thus $w$ is counted $m_j$ times in $qf(n)$ in the LHS of (34). Further, in the RHS of (34), $w$ is counted only once in $f_{a_i}(n+1)$, since $r_j$ being a subword of $a_i$ does not contribute to any extra counting. Thus we need to add $(m_j - 1)f_{a_i}(n+1)$ in all such cases. The term $T(n)$ refers to all such extra counting. It is given by

$$T(n) = \sum_{i=1}^{s} \sum_{j=1}^{\ell} \sum_{0 < \alpha \in (a_i, r_j)} \sum_{\alpha > |r_j|} (m_j - 1) f_{a_i}(n+1).$$

Using the definition of $\gamma(a_i, r_j)$ as given in (16), we obtain the following

$$qf(n) - f(n+1) = \sum_{i=1}^{s} f_{a_i}(n+1) - \sum_{i=1}^{\ell} \left( 1 - \frac{1}{m_i} \right) g_i(n+1)$$

$$+ \sum_{i=1}^{s} \left[ \sum_{j=1}^{\ell} \gamma(a_i, r_j)(m_j - 1) \right] f_{a_i}(n+1). \tag{35}$$

Multiplying the above equation by $z^{-n}$ and taking the sum over $n \geq 0$, we get (17) as required.

In order to obtain the second equation, we adjoin a repeated word $r_k$ to an allowed word $w$ counted in $f(n)$ as a suffix. Our claim is that $f(n)$ satisfies

$$f(n) = g_{r_k}(n+|r_k|) - \sum_{j=1}^{\ell} \sum_{0 < s \in (r_j, r_k)} \left( 1 - \frac{1}{m_j} \right) g_{r_j}(n+s) + \sum_{i=1}^{s} \sum_{0 < t \in (a_i, r_k)^{r_k}} f_{a_i}(n+t)$$

$$+ \sum_{i=1}^{s} \sum_{0 < t \in (a_i, r_k)} \gamma(a_i, r_j)(m_j - 1)f_{a_i}(n+t). \tag{36}$$
Two possible situations arise, as described below:

1) If no forbidden word appears on the join of \(w\) and \(r_k\), then \(w\) is counted in \(g_{r_k}(n+|r_k|)\), the first term on the RHS of (36). Further, if a repeated word \(r_j\) appears on the join, then there exists \(0 < s \in (r_j, r_k)\) such that \(w\) is counted extra \(m_j - 1\) times in \(g_{r_k}(n+|r_k|)\). The total number of \(w\) in \(f(n)\) which give \(r_j\) on the join is \(\frac{g_{r_j}(n+s)}{m_j}\), for some \(0 < s \in (r_j, r_k)\). We subtract these extra counting to obtain the second term on the RHS of (36).

2) Now, suppose there exists a forbidden word on the join of \(w\) and \(r_k\), and let \(a_i\) be the first such occurrence. That means, there exists \(0 < t \in (a_i, r_k)\) such that \(w\) is counted in \(f_{a_i}(n+t)\). These terms constitute the third term on the RHS of (36). Further, since the forbidden word \(a_i\) on the join may contain some repeated word as its subword, we need to carefully handle the multiple counting of the words in such situations. As explained earlier, it is enough to look at the situations when \(r_j\) is a subword of \(a_i\), except at the end. There are only two possible positions where \(r_j\) (as a subword of \(a_i\)) can appear in \(w r_k\), which contributes extra to the counting. These are depicted in Figure 7. The figure on the left represents the situation when \(r_j\) appears before the join, as a subword of \(w\), and the figure on the right represents the situation when \(r_j\) appears on the join.

i) In the first one (figure on the left), \(r_j\) appears before the join as a subword of \(a_i\), except at the end, and also as a subword of \(w\) if and only if there exists \(\alpha \in (a_i, r_j)\) such that \(\alpha > |r_j|\) and \(\alpha > t + |r_j|\), where \(0 < t \in (a_i, r_k)\) is fixed. In these cases, \(r_j\) contributes \(m_j\) to the multiplicity of \(w\) in \(f(n)\). Further, \(w\) is counted in the third term on the RHS of (36), but \(r_j\) as a subword of \(a_i\) does not contribute to the counting of \(w\) in \(f_{a_i}(n+t)\). Therefore, to balance the counting, we need to add the term \((m_j - 1)f_{a_i}(n+t)\) for appropriate choices of \(m_j\), determined using \(\alpha\). These terms are included in the last term on the RHS of (36).

ii) In the second one (figure on the right), \(r_j\) occurs on the join. This holds if and only if there exists \(\alpha \in (a_i, r_j)\) such that \(\alpha > |r_j|\) and \(\alpha > t\). Note that \(\alpha > t\) is implicit here, for if \(\alpha \leq t\), then \(r_j\) is a subword of \(r_k\), which is not allowed since \(R\) is a reduced collection. Let us suppose \(w\) gets counted only once in \(f(n)\). Moreover, \(w\) gets counted \(m_j - 1\) times in \(\frac{g_{r_j}(n+s)}{m_j}\) for a suitable choice of \(0 < s \in (r_j, r_k)\) and once in \(f_{a_i}(n+t)\). In this case as well, we need to add \((m_j - 1)f_{a_i}(n+t)\), which are included in the last term on the RHS of (36) for appropriate choices of \(\alpha\).

These two ((i) and ii)) together give us: \(\#\{\alpha \in (a_i, r_j) : \alpha > |r_j|\) and \(\alpha > t + |r_j|\} = \#\{\alpha \in (a_i, r_j) : \alpha > |r_j|\} = \gamma(a_i, r_j)\). Therefore the terms obtained from these two cases constitute the last term in the RHS of (36).

Now, (36) is obtained by multiplying (35) by \(z^{-n}\) and taking sum over \(n \geq 0\).

Finally, consider \(w a_k\), where \(w\) is counted in \(f(n)\), and \(a_k \in F\). Using the arguments developed so far in this proof and the proof of Theorem 1.3 we can obtain an expression
for \( f(n) \) given by

\[
\begin{align*}
f(n) &= \sum_{i=1}^{s} \sum_{0 < t \in (a_i,a_k)} f_{a_i}(n + t) - \sum_{j=1}^{\ell} \sum_{0 < s \in (r_j,a_k) r_j|^{-1}} \left( 1 - \frac{1}{m_j} \right) g_{r_j}(n + s) \\
&\quad + \sum_{i=1}^{s} \sum_{0 < t \in (a_i,a_k)} \left[ \sum_{j=1}^{\ell} (m_j - 1) \gamma_t(a_i, r_j) \right] f_{a_i}(n + t),
\end{align*}
\]

where \( \gamma_t(a_i, r_j) = \# \{ \alpha \in (a_i, r_j) : \alpha > \max\{|r_j|, t\} \} \). Here the condition on \( \alpha \) is obtained as \( \alpha > \max\{|r_j|, t\} \), for each fixed \( 0 < t \in (a_i, a_k) \). Here \( \alpha > t \) needs to be imposed explicitly unlike the previous case since, for \( \alpha \leq t \), \( a_k \) can contain \( r_j \) as a subword, however it does not contribute to the counting. To take care of this extra condition, we define \( \gamma_t(a_i, r_j) \) as above and obtain Equation (37). Finally to complete the proof, we multiply the above equation by \( z^{-n} \) and take sum over \( n \geq 0 \) to obtain (19).