

Distribution of integers with digit restrictions via Markov chains

VICENTE SAAVEDRA-ARAYA 

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
(e-mail: vicente.saavedra-araya@warwick.ac.uk)

(Received 28 January 2025 and accepted in revised form 19 October 2025)

Abstract. In this paper, we introduce a new technique to study the distribution in residue classes of sets of integers with digit and sum-of-digits restrictions. From our main theorem, we derive a necessary and sufficient condition for integers with missing digits to be uniformly distributed in arithmetic progressions, extending previous results going back to the work of Erdős, Mauduit and Sárközy. Our approach uses Markov chains and does not rely on Fourier analysis as many results of this nature do. Our results apply more generally to the class of multiplicatively invariant sets of integers. This class, defined by Glasscock, Moreira and Richter using symbolic dynamics, is an integer analogue to fractal sets and includes all missing digits sets. We address uniform distribution in this setting, partially answering an open question posed by the same authors.

Key words: symbolic dynamics, number theory

2020 Mathematics Subject Classification: 11N69 (Primary); 60J10, 37B10 (Secondary)

Contents

1	Introduction	1
2	Preliminaries	8
3	Construction of Markov chains	13
4	From Markov chains to uniform distribution	19
5	Applications to the distribution of multiplicatively invariant sets of integers	27
6	Transversality between multiplicatively invariant sets and arithmetic progressions	38
	Acknowledgements	46
	References	46

1. Introduction

In this paper, we investigate the distribution of sets of integers formed by imposing restrictions on the digits in a given base. Throughout this paper, we denote by \mathbb{N}_0

the set of natural numbers including zero. For an integer $g \geq 2$ and a set of digits $\mathcal{D} \subseteq \{0, \dots, g-1\}$, let

$$C_{g,\mathcal{D}} := \left\{ \sum_{i=0}^m w_i g^i : m \in \mathbb{N}_0, w_i \in \mathcal{D} \right\} \quad (1.1)$$

be the set of non-negative integers that can be written in base g only using digits of \mathcal{D} . Such sets have been widely studied (see, for example, [18, 29, 38, 41]), and we will refer to them as *missing digits sets*. A broader class of integers was introduced by Glasscock, Moreira and Richter in [27].

Definition 1.1. Let $g \geq 2$ be an integer. The set $A \subseteq \mathbb{N}_0$ is said to be $\times g$ -invariant if for every $n \in A$ with base- g expansion given by $n = w_m g^m + \dots + w_1 g + w_0$, where $m \in \mathbb{N}_0$ and $w_m \neq 0$, it holds

$$w_{m-1} g^{m-1} + \dots + w_1 g + w_0 \in A \quad \text{and} \quad w_m g^{m-1} + \dots + w_2 g + w_1 \in A.$$

We say a set of integers is *multiplicatively invariant* if it is $\times g$ -invariant for some $g \geq 2$.

Notice that the first condition in Definition 1.1 indicates that the set A is invariant under the operation of deleting the most significant digit in base g , while the second condition ensures that it is invariant under the operation of deleting the least significant digit. Clearly, integers with missing digits belong to this class, but so do many other sets described by broader digit restrictions (for example, by forbidding specific combinations of digits). Unlike missing digits sets, much less is known about multiplicatively invariant sets.

Motivated by the study of the distribution of integers with missing digits in arithmetic progressions by Erdős, Mauduit and Sárközy in [18], and by an open question posed by Glasscock, Moreira and Richter in [27, Question 5.6] concerning the interaction between multiplicatively invariant sets and arithmetic progressions, this paper investigates distribution properties of multiplicatively invariant sets. A central idea is to understand when a set of integers visits different congruence classes with the same frequency. In this sense, $A \subseteq \mathbb{N}_0$ is said to be *uniformly distributed mod a* (or in arithmetic progressions) if

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A : n \equiv b \pmod{a}\} \cap [0, N)|}{|A \cap [0, N)|} = \frac{1}{a}$$

for every $b \in \mathbb{Z}_a := \mathbb{Z}/a\mathbb{Z}$. To explore this idea in more generality, we use the notion of g -additive functions (first considered by Bellman and Shapiro in [7]). We say $f : \mathbb{N}_0 \rightarrow \mathbb{Z}$ is g -additive if

$$f\left(\sum_{i=0}^m w_i g^i\right) = \sum_{i=0}^m f(w_i g^i)$$

for every $w_0, \dots, w_m \in \{0, \dots, g-1\}$. Let $A \subseteq \mathbb{N}_0$ be $\times g$ -invariant and $r \in \mathbb{N}$. For all $i \in \{1, \dots, r\}$, let f_i be a g -additive function, $a_i \in \mathbb{N}$ and $b_i \in \mathbb{Z}_{a_i}$. The main objective in this work is to study the limiting behaviour of

$$\frac{|\{n \in A : f_i(n) \equiv b_i \pmod{a_i} \text{ for } i = 1, \dots, r\} \cap [0, N)|}{|A \cap [0, N)|} \quad (1.2)$$

when $N \rightarrow \infty$, and determine conditions to guarantee the limit equals to $1/(a_1 \cdot a_2 \cdot \dots \cdot a_r)$ for all $b_i \in \mathbb{Z}_{a_i}$. When this happens, we will say that A is *uniformly distributed* $\mathbf{f} \bmod \mathbf{a}$, where $\mathbf{f} = (f_1, \dots, f_r)$ and $\mathbf{a} = (a_1, \dots, a_r)$. Our main idea lies in the construction of suitable Markov chains that capture the behaviour of (1.2), and then study it using classical convergence results for Markov chains. Before stating our main results, we will briefly review some previous results and introduce the context of the problem to be investigated.

1.1. History and context. The study of the distribution of sets of integers with digit restrictions dates back to the works of Fine [20] and Gelfond [25]. For an integer $g \geq 2$, they considered the function sum-of-digits in base g , denoted by S_g and defined as follows. If $w_0, \dots, w_m \in \{0, \dots, g-1\}$,

$$S_g\left(\sum_{i=0}^m w_i g^i\right) = w_0 + \dots + w_m. \quad (1.3)$$

In this context, Gelfond studied the interplay between n and $S_g(n)$ for positive integers in terms of their distribution in residue classes, providing sufficient conditions for uniform distribution. A straightforward consequence of [25, Theorem 1] is presented as the following corollary.

COROLLARY 1.2. *Let $a, a' \in \mathbb{N}$ such that $\gcd(g-1, a') = 1$. Then,*

$$\lim_{N \rightarrow \infty} \frac{|\{n \leq N : n \equiv b \bmod a, S_g(n) \equiv b' \bmod a'\}|}{N} = \frac{1}{aa'} \quad (1.4)$$

for any $b \in \mathbb{Z}_a$ and $b' \in \mathbb{Z}_{a'}$.

It is worth mentioning that Gelfond's result also gives quantitative information about the convergence. Corollary 1.2 shows that the sequences $n \bmod a$ and $S_g(n) \bmod a'$ are independent; in other words, integers with a fixed sum of digits $\bmod a'$ are uniformly distributed in arithmetic progressions (a related result can be found in [40]). As part of our work, we provide an equivalent condition to the uniform distribution exhibited in (1.4), thereby extending previous works. Results concerning the independence between sum-of-digits functions in different bases can be found in [8, 30]. For an overview of results of this type, see [46, §4.3].

In a related direction, the distribution of missing digits sets in residue classes was initially studied by Erdős, Mauduit and Sárközy. Let $\mathcal{D} \subset \{0, \dots, g-1\}$ be a subset of digits (with at least two elements). Regarding the set $C_{g,\mathcal{D}}$ defined in (1.1), they showed the following result in [18, Theorem 1].

THEOREM 1.3. *Assume $\gcd(g(g-1), a) = 1$, $0 \in \mathcal{D}$ and $\gcd\{d \in \mathcal{D} : d \neq 0\} = 1$. Then, there exists c_1, c_2 and c_3 (all dependent only on g and $|\mathcal{D}|$) such that*

$$\left| \frac{|\{n \in C_{g,\mathcal{D}} : n \equiv b \bmod a\} \cap [0, N]|}{|C_{g,\mathcal{D}} \cap [0, N]|} - \frac{1}{a} \right| \leq \frac{c_2}{m} \exp\left(-c_3 \frac{\log(N)}{\log(a)}\right)$$

for any $a < \exp(c_1(\log(N))^{1/2})$ and $b \in \mathbb{Z}_a$.

Further generalization of Theorem 1.3 was provided by Konyagin [31, Corollary 1] and Aloui [2, Corollary 2.2], with the latter also generalizing Corollary 1.2 to the case of missing digits sets. A qualitative version of Theorem 1.3 can be presented as the following corollary.

COROLLARY 1.4. *Assume $\gcd(g(g-1), a) = 1$, $0 \in \mathcal{D}$ and $\gcd\{d \in \mathcal{D} : d \neq 0\} = 1$. Then, $C_{g,\mathcal{D}}$ is uniformly distributed mod a .*

As a consequence of our work, we strengthen Corollary 1.4, providing an equivalent condition to the uniform distribution when $\gcd(g, a) = 1$.

The study of arithmetic properties for sets with digit restrictions has been an active area in number theory and is not limited to the problems covered in this paper. Different directions that have been explored include, for instance, studies on the number of prime factors for elements of sets with digit restrictions [12, 13, 19, 32], as well as properties of divisibility [6, 11]. It is also worth mentioning the breakthrough work of Mauduit and Rivat [39] about the distribution of the function sum-of-digits for primes in residue classes and the works of Maynard [41, 42] about primes with missing digits, showing the existence of infinitely many primes with one missing digit. Additional references include [3, 5, 14–17, 43, 49–51, 54].

However, a famous conjecture in fractal geometry and dynamical systems was posed by Furstenberg in [23], which stated that $\times p \bmod 1$ and $\times q \bmod 1$ invariant sets of $[0, 1]$ are transverse when $\log(p)/\log(q) \notin \mathbb{Q}$. (For a given notion of dimension, the sets $A, B \subseteq [0, 1]$ are said to be transverse if $\dim(A \cap B) \leq \max\{0, \dim(A) + \dim(B) - 1\}$.) This conjecture was independently proven by Shmerkin [47] and Wu [53] (see [4] for an alternative proof). Motivated by an integer version of this conjecture, Glasscock, Moreira and Richter introduced multiplicatively invariant sets of integers (Definition 1.1), proving in [27, Theorem B] an analogous result to Furstenberg's conjecture in the setting of integer fractals. Seeking a better understanding of invariant sets and their interaction with other sets from a fractal point of view, they raised the following question [27, Question 5.6].

Open question 1.5. Let $A \subseteq \mathbb{N}_0$ be a multiplicatively invariant set and P be an arithmetic progression. Is it true that $\dim(A \cap P)$ is either 0 or $\dim(A)$?

Particularly, they used the notion of mass dimension (see Definition 2.11). The answer to this question is only known for certain missing digits sets (those covered by previous results about uniform distribution), but has not been explored in generality. In this paper, we provide a negative answer to this question in the general case and describe a class of sets where the answer is affirmative. Finally, we propose a conjecture identifying the precise class of multiplicatively invariant sets that satisfy this property.

1.2. Overview of main results. The main contribution of this paper lies in the introduction of a new approach to study the distribution of certain sets of integers via Markov chains. From this technique, we can retrieve several results, some of which are presented below.

In the direction of the works of Fine [20], Gelfond [25], and Erdős, Mauduit and Sárközy [18], we study the distribution in residue classes of set of integers with digits restrictions. A first consequence of Theorem 4.9, our main technical result, establishes conditions for a set with missing digits to be uniformly distributed in residue classes. In contrast to Corollary 1.4 and Aloui's result [2, Theorem 2.9], which provide only sufficient conditions, our result offers conditions that are both necessary and sufficient for uniform distribution.

THEOREM A. *Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$ such that $\gcd(g, a) = \gcd(a, a') = 1$. Let $\mathcal{D} = \{d_1, \dots, d_t\} \subseteq \{0, \dots, g-1\}$ be a subset of digits (where $d_1 < \dots < d_t$) and $C_{g, \mathcal{D}}$ be the missing digits set defined in (1.1). Then,*

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, N]|}{|C_{g, \mathcal{D}} \cap [0, N]|} = \frac{1}{aa'}$$

for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$ if and only if $\gcd(aa', d_2 - d_1, \dots, d_t - d_1) = 1$.

Indeed, we provide a comprehensive description of the limiting distribution even when $\gcd(aa', d_2 - d_1, \dots, d_t - d_1) \neq 1$ (Theorem 5.4) and offer insight into the rate of convergence. In Proposition 5.7, we also present an equivalent condition without the assumption $\gcd(a, a') = 1$.

By setting $a' = 1$, we generalize Corollary 1.4 by providing an equivalent condition. Furthermore, we are able to recover the same convergence speed as that obtained by Erdős, Mauduit and Sárközy, albeit with a difference in the constants involved.

COROLLARY A. *Let $g \geq 2$ be an integer and $a \in \mathbb{N}$ such that $\gcd(g, a) = 1$. Let $\mathcal{D} = \{d_1, \dots, d_t\} \subseteq \{0, \dots, g-1\}$ be a subset of digits, where $d_1 < \dots < d_t$. Then, $C_{g, \mathcal{D}}$ is uniformly distributed mod a if and only if $\gcd(a, d_2 - d_1, \dots, d_t - d_1) = 1$.*

For a full description of the distribution mod a , see Corollary 5.5. Our results are not limited to the distribution of integers with missing digits, but more general digit restrictions are allowed. In §§5.3 and 5.4, we address more general restrictions and provide examples of sets with missing combinations of digits that exhibit uniform distribution. Additionally, our method allows us to revisit the result of Gelfond, yielding an equivalent condition to the uniform distribution exhibited in (1.4). In particular, this shows that the sufficient condition $\gcd(g-1, a') = 1$ provided in Corollary 1.2 is not necessary.

THEOREM B. *Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$. Then,*

$$\lim_{N \rightarrow \infty} \frac{|\{n \leq N : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\}|}{N} = \frac{1}{aa'}$$

for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$ if and only if there exists $n \in \mathbb{N}$ such that $n \equiv 0 \pmod{a}$ and $S_g(n) \equiv 1 \pmod{a'}$.

Notice that the existence of $n \in \mathbb{N}$ that satisfies $n \equiv 0 \pmod{a}$ and $S_g(n) \equiv 1 \pmod{a'}$ can be established from an elementary construction under the condition $\gcd(g-1, a') = 1$. (The author is grateful for the useful construction provided in <https://mathoverflow.net/q/482398>.)

Finally, we provide a negative answer in the general case to the open question posed by Glasscock, Moreira and Richter [27, Question 5.6], and we provide a description using symbolic dynamics of sets of integers satisfying affirmatively the statement. An important idea is that every multiplicatively invariant set of integers can be characterized by a dynamical structure known as a subshift (see Proposition 2.13). Two important classes of subshifts are sofic subshifts and transitive subshifts (see §2.2 for details). Using these notions, we establish the following partial answer to the Open question 1.5.

THEOREM C. *Let $A \subseteq \mathbb{N}_0$ be a $\times g$ -invariant set that can be represented by a transitive sofic subshift. Then, for every infinite arithmetic progression P , $\dim_{\mathbb{M}}(A \cap P)$ is either 0 or $\dim_{\mathbb{M}}(A)$. Moreover, if the subshift is only transitive, or only sofic, the result is not always true.*

Refer to Theorem 6.7 for a more general version of this result. Furthermore, we show that the sofic and transitive setting is not optimal, as the affirmative answer can be extended to more general classes of subshifts. Thus, using the notion of topological entropy (see Definitions 2.3 and 6.10), we propose the following conjecture.

Conjecture 1.6. *Let $A \subseteq \mathbb{N}_0$ be a $\times g$ -invariant set. There exists an entropy minimal subshift that represents A if and only if, for every arithmetic progression P , $\dim_{\mathbb{M}}(A \cap P)$ is either 0 or $\dim_{\mathbb{M}}(A)$.*

1.3. Outline of the proof. We present an outline of the proof of Corollary A. Despite being a consequence of more general results, the proof contains the key ideas of this paper while avoiding some obstructions that arise when dealing with additional digit restrictions.

Consider an integer $g \geq 2$, an integer $a \in \mathbb{N}$ coprime with g and a set of digits $\mathcal{D} = \{d_1, \dots, d_t\} \subseteq \{0, \dots, g-1\}$. We aim to construct Markov chains that describe the behaviour of

$$\frac{|\{n \in C_{g,\mathcal{D}} : n \equiv b \pmod{a}\} \cap [0, N)|}{|C_{g,\mathcal{D}} \cap [0, N)|} \quad (1.5)$$

when $N \rightarrow \infty$. For a finite sequence of digits $w = w_0 w_1 \dots w_m$ (we call it a word of length $m+1$), we denote

$$(w)_g := w_m g^m + \dots + w_1 g + w_0.$$

We identify \mathcal{D}^i with the set of words of length i and we say $w \in \mathcal{D}^i$ represents $n \in \mathbb{N}_0$ if $(w)_g = n$. Clearly,

$$C_{g,\mathcal{D}} = \{(w)_g : w \in \mathcal{D}^n \text{ for } n \in \mathbb{N}_0\},$$

so equivalently, we can study the behaviour of (1.5) in terms of words. For every i , we define the measure μ_i on \mathbb{Z}_a as

$$\mu_i(\{b\}) := \frac{|\{w \in \mathcal{D}^i : (w)_g \equiv b \pmod{a}\}|}{|\mathcal{D}|^i}.$$

For simplicity, we will denote $\mu_i(b) = \mu_i(\{b\})$. This measure represents the distribution in \mathbb{Z}_a of elements in $C_{g,\mathcal{D}}$ that can be expressed using i digits of \mathcal{D} in base g . Our goal is

to understand the evolution of μ_i , and to do this, we will study how a word of length i is extended. (We say $x' = x'_0 \cdots x'_k$ is an extension of $x = x_0 \cdots x_m$ if $k > m$ and $x'_j = x_j$ for all $j \leq m$.) For this purpose, we consider $p \in \mathbb{N}$ such that $g^p \equiv 1 \pmod{a(g-1)}$ (whose existence is guaranteed by the condition $\gcd(g, a) = 1$). Let us define the sets E_i for $i \in \mathbb{N}$ as

$$E_i(b) := \{w \in \mathcal{D}^p : g^i(w)_g \equiv b \pmod{a}\}, \quad (1.6)$$

and we also define the matrix on $\mathbb{Z}_a \times \mathbb{Z}_a$ given by

$$M_i(b, b') := \frac{|E_i(b' - b)|}{|\mathcal{D}|^p}.$$

It is easy to see that M_i is a double stochastic matrix. Let $n \in \mathbb{N}$ and let $x \in \mathcal{D}^{i+np}$ such that $(x)_g \equiv b \pmod{a}$. We can ask the following question. For $b' \in \mathbb{Z}_a$, how many extensions $x' \in \mathcal{D}^{i+(n+1)p}$ of x satisfy $(x')_g \equiv b' \pmod{a}$? The answer is the cardinality of $E_{i+np}(b' - b)$ and we can observe that $E_i(b - b') = E_{i+np}(b - b')$ since $g^{np} \equiv 1 \pmod{a}$. From this, we can obtain the relation

$$\mu_{i+np} = \mu_i M_i^n.$$

If X^i is a Markov chain with initial distribution μ_i and transition matrix M_i , it holds that

$$\mathbb{P}(X_n^i = b) = \mu_{i+np}(b).$$

Moreover, if X^i is irreducible and aperiodic, the Markov chain converges to its invariant distribution, and hence,

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n^i = b) = \lim_{n \rightarrow \infty} \frac{|\{w \in \mathcal{D}^{i+np} : (w)_g \equiv b \pmod{a}\}|}{|\mathcal{D}|^{i+np}} = \frac{1}{a}. \quad (1.7)$$

If $0 \notin \mathcal{D}$, there is a one-to-one correspondence between elements of $C_{g, \mathcal{D}}$ and words. However, since $(w0 \cdots 0)_g = (w)_g$, every element of $C_{g, \mathcal{D}}$ can be represented by infinitely many words when $0 \in \mathcal{D}$. By making this distinction and under the assumption that X^1, \dots, X^p are irreducible and aperiodic, we can elucidate (1.7) in terms of the set of integers with missing digits.

- If $0 \in \mathcal{D}$,

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}\} \cap [0, g^N)|}{|C_{g, \mathcal{D}} \cap [0, g^N)|} = \frac{1}{a}.$$

- If $0 \notin \mathcal{D}$,

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}\} \cap [g^{N-1}, g^N)|}{|C_{g, \mathcal{D}} \cap [g^{N-1}, g^N)|} = \frac{1}{a}.$$

For an arbitrary set of integers, this does not guarantee the existence of the limit when considering intervals of the form $[0, N)$. However, as we will prove in §4, in this case, it happens. Therefore, we can show that

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}\} \cap [0, N]|}{|C_{g, \mathcal{D}} \cap [0, N]|} = \frac{1}{a} \quad (1.8)$$

for every $b \in \mathbb{Z}_a$. Certainly, we must still ensure that each Markov chain X^i converges to the equilibrium. This property will hold if we choose p as above and $\gcd(a, d_2 - d_1, \dots, d_t - d_1) = 1$, where d_1 is the smallest digit of \mathcal{D} .

Alternatively, if $\gcd(a, d_2 - d_1, \dots, d_t - d_1) \neq 1$, it is possible to show that there exist $b \in \mathbb{Z}_a$ and $j \in \{1, \dots, p\}$ such that

$$\mathbb{P}(X_n^j = b) = 0$$

for every $n \in \mathbb{N}$. From this fact, we will conclude that (1.8) does not hold for such a b .

Some difficulties arise when dealing with additional constraints on the digits, but we will extend the construction presented above to more general cases by leveraging ideas from symbolic dynamics. We are going to discuss this in detail in the following sections.

2. Preliminaries

In this section, we introduce some concepts of Markov chains, symbolic dynamics and integer fractals that will be used throughout this paper.

2.1. Markov chains. Let \mathcal{S} be a finite set and let M be a square matrix indexed by $\mathcal{S} \times \mathcal{S}$ with entries in $[0, 1]$. The matrix M is called *stochastic* if the sum of each row is 1, that is, for every $s \in \mathcal{S}$,

$$\sum_{s' \in \mathcal{S}} M(s, s') = 1.$$

Given a stochastic matrix M and a probability measure μ on \mathcal{S} , a *discrete-time Markov chain* with state space \mathcal{S} , transition matrix M and initial distribution μ is a sequence of random variables $(X_n)_{n \in \mathbb{N}_0}$ on a probability space (with probability \mathbb{P}) such that $\mathbb{P}(X_0 = s) = \mu(\{s\}) =: \mu(s)$, and

$$\mathbb{P}(X_{n+1} = s_{n+1} | X_n = s_n, \dots, X_0 = s_0) = \mathbb{P}(X_1 = s_{n+1} | X_0 = s_n) = M(s_n, s_{n+1})$$

for every $n \in \mathbb{N}_0$ and $s_0, \dots, s_{n+1} \in \mathcal{S}$. In this case,

$$\mathbb{P}(X_n = s') = \mu \cdot M^n(\cdot, s') = \sum_{s \in \mathcal{S}} \mu(s) M^n(s, s').$$

For $s, s' \in \mathcal{S}$, we say that s' is *accessible from* s if $M^n(s, s') > 0$ for some $n \in \mathbb{N}$. The matrix M is called *irreducible* if for every $s, s' \in \mathcal{S}$, there exists $n \in \mathbb{N}$ such that $M^n(s, s') > 0$ (that is, all states are accessible from each other). We define the *period of the state* s as

$$\text{Per}(s) := \gcd\{n \in \mathbb{N} : M^n(s, s) > 0\}.$$

If every state has period 1, we say M is *aperiodic*. When M is irreducible, all states share the same period, so aperiodicity can be checked by finding one state with period 1.

The following classical result on the convergence of Markov chains will be a key tool in this paper.

THEOREM 2.1. [35, Theorem 4.9] *Let X be a Markov chain with state space \mathcal{S} (not necessarily finite), transition matrix M and initial distribution μ . Suppose M is irreducible and aperiodic, and there exists an invariant distribution λ for M (that is, $\lambda = \lambda M$). Then, there exist $\rho \in (0, 1)$ and $C > 0$ such that*

$$|\mathbb{P}(X_n = s) - \lambda(s)| \leq C\rho^n$$

for every $n \in \mathbb{N}$ and $s \in \mathcal{S}$.

Notice that the invariant distribution for an irreducible and aperiodic transition matrix with a finite state space always exists (and it is unique) as a consequence of the Perron–Frobenius theorem. In addition, Theorem 2.1 is only dependent on the transition matrix, so that λ , C and ρ are independent of the initial distribution. A particularly interesting case occurs when the sum of each column of M is also 1, in which case, M is called *doubly stochastic*. In this case, the distribution $\lambda(s) = 1/|\mathcal{S}|$ is invariant for M . For further discussion on the convergence of Markov chains, see [44, §1.8] and [45, §3].

2.2. Elements of symbolic dynamics. Let \mathcal{D} be a finite set (in this paper, we assume $\mathcal{D} \subseteq \{0, \dots, g-1\}$ for some $g \in \mathbb{N}$) and consider the *left-shift* operator σ over $\mathcal{D}^{\mathbb{N}_0}$, defined by

$$\sigma : (w_n)_{n \in \mathbb{N}_0} \mapsto (w_{n+1})_{n \in \mathbb{N}_0}.$$

Endowing \mathcal{D} with the discrete topology and considering the product topology on $\mathcal{D}^{\mathbb{N}_0}$, σ is a continuous function, and the topological system $(\mathcal{D}^{\mathbb{N}_0}, \sigma)$ is called a *full shift*. If $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ is closed and $\sigma(\Sigma) \subseteq \Sigma$, the dynamical system (Σ, σ) is called a *subshift*. Although a subshift is a set of infinite sequences of symbols, it can be fully characterized by the finite sequences that appear in Σ , referred to as *words*.

We define the *language* of Σ , denoted by $\mathcal{L}(\Sigma)$, as the set of all the words in Σ ,

$$\mathcal{L}(\Sigma) := \{w_0 \cdots w_{n-1} : n \in \mathbb{N}_0, \text{ there exists } x \in \Sigma \text{ such that } x_0 = w_0, x_1 = w_1, \dots, x_{n-1} = w_{n-1}\}.$$

When there is no confusion, we will simply denote the language by \mathcal{L} . For any word $w = w_0 \cdots w_{n-1} \in \mathcal{L}(\Sigma)$, we say w has *length* n (and we denote the length by $|w|$). The set $\mathcal{L}^n(\Sigma)$ denotes all words in the language of length n . Also, we will commonly use the notions of concatenation of words and subwords. For $u = u_0 \cdots u_{m-1}$ and $v = v_0 \cdots v_{n-1}$, the *concatenation* of u with v is written as $uv = u_0 \cdots u_{m-1}v_0 \cdots v_{n-1}$ and we say that uv is an *extension* of u . For a given word u and $k \in \mathbb{N}$, we also denote by $u^k = u \cdots u$ the concatenation k -times of u . We say w is a *subword* of x if the sequence of x contains w .

A subshift is a particular case of a topological dynamical system and some classical notions of topological dynamics have a special characterization in the context of symbolic systems.

Definition 2.2. The topological dynamical system (Σ, σ) is said to be *transitive* (or *irreducible*) if for every $u, v \in \mathcal{L}(\Sigma)$, there exists $w \in \mathcal{L}(\Sigma)$ such that $uwv \in \mathcal{L}(\Sigma)$. Also, the system is said to be *mixing* if for every u, v in the language, there exists $n \in \mathbb{N}$ such that for every $m \geq n$, there exists $w \in \mathcal{L}^m(\Sigma)$ such that $uwv \in \mathcal{L}(\Sigma)$.

However, the notion of *topological entropy* is a measure of the complexity of a dynamical system. In symbolic systems, it quantifies the complexity of the language.

Definition 2.3. The (topological) entropy of the subshift Σ is defined as

$$h(\Sigma) := \lim_{N \rightarrow \infty} \frac{\log |\mathcal{L}^N(\Sigma)|}{N}.$$

See [36, Ch. 4] for further discussion. Regarding the classification of subshifts, we will consider two important classes, *shifts of finite type* and *sofic shifts*.

Definition 2.4. Consider the full shift $X = \mathcal{D}^{\mathbb{N}_0}$ and let $\mathcal{F} \subseteq \mathcal{L}(X)$ be a finite set of *forbidden words*. We define the subshift

$$X_{\mathcal{F}} := \{x \in X : x \text{ does not contain any } w \in \mathcal{F} \text{ as a subword}\}.$$

We say that $\Sigma \subseteq X$ is a *shift of finite type* (SFT for short) if there exists a finite set \mathcal{F} such that $\Sigma = X_{\mathcal{F}}$.

If the longest forbidden word has length $M + 1$, we say that $X_{\mathcal{F}}$ is an *M-step* shift of finite type. In a 1-step SFT, all forbidden words have length 2, so it can be described by a matrix indexed by $\mathcal{D} \times \mathcal{D}$ with values in $\{0, 1\}$. In this sense, Σ is a 1-step SFT if and only if there exists a matrix T such that

$$\Sigma = \{x \in \mathcal{D}^{\mathbb{N}_0} : T(x_i, x_{i+1}) = 1 \text{ for all } i\}.$$

A more general way to construct subshifts is by reading infinite sequences in a *labelled directed multi-graph*. A labelled multi-graph is a triple $G = (V, E, \lambda)$, where V is a finite set of nodes, E is a set of edges and $\lambda : E \rightarrow \mathcal{A}$ is a labelling function for some finite alphabet \mathcal{A} . Each edge $e \in E$ has an associated starting node $s(e) \in V$ and a terminal node $t(e) \in V$. A sequence of edges $\pi = e_0 \cdots e_n$ is called a *path* in G if $t(e_i) = s(e_{i+1})$ for every $i \in \{0, \dots, n-1\}$, and we say that π has length $|\pi| := n + 1$. Note that multiple edges may share the same starting and terminal nodes while having different labels.

Definition 2.5. Let $G = (V, E, \lambda)$ be a labelled directed multi-graph. If \mathcal{D} is the finite alphabet associated with the labelling function, we define the subshift $\Sigma_G \subset \mathcal{D}^{\mathbb{N}_0}$ as

$$\Sigma_G = \{(\lambda(e_j))_{j \in \mathbb{N}_0} \in \mathcal{D}^{\mathbb{N}_0} : s(e_{j+1}) = t(e_j) \text{ for all } j \in \mathbb{N}_0\}.$$

Thus, Σ_G are all infinite sequences that we can read in some infinite path of G . It is well known that Σ_G is a subshift.

Definition 2.6. The subshift $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ is said to be *sofic* if there exists a labelled directed multi-graph such that $\Sigma = \Sigma_G$. The graph G is called a *cover* of Σ .

Note that the cover of a sofic shift is not unique. Additionally, every shift of finite type is sofic.

Remark 2.7. Let $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ be a 1-step shift of finite type, and consider a matrix T that represents Σ , that is, $T(d, d') = 1$ if and only if $dd' \in \mathcal{L}(\Sigma)$. We create a node $V(d)$ for every $d \in \mathcal{D}$, and place an edge from $V(d)$ to $V(d')$ labelled by d' if and only if $T(d, d') = 1$. The resulting labelled graph is a cover for Σ and we denote it by \mathcal{G}_T .

To construct a cover for a sofic shift, we can rely on the notion of *follower sets* [36, Ch. 3].

Definition 2.8. Let Σ be a subshift of $\mathcal{D}^{\mathbb{N}_0}$. For any $w \in \mathcal{L}(\Sigma)$, the *follower set* of w , denoted $F_\Sigma(w)$, is the set of all words that can extend w , that is,

$$F_\Sigma(w) := \{v \in \mathcal{L}(\Sigma) : wv \in \mathcal{L}(\Sigma)\}.$$

We denote by $\mathcal{F}_\Sigma := \{F_\Sigma(w) : w \in \mathcal{L}(\Sigma)\}$ the collection of all follower sets. When the underlying subshift Σ is clear from the context, we will simply write $F(w)$ and \mathcal{F} .

While each word has its own follower set, a folklore result states that Σ is sofic if and only if there are finitely many distinct follower sets [36, Theorem 3.2.10], that is, $|\mathcal{F}_\Sigma| < \infty$. If $d \in F(u) = F(v)$, it follows that $F(ud) = F(vd)$. Using these ideas, a cover can be constructed for Σ .

Definition 2.9. Let $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ be a sofic shift. Consider the finite set of nodes $V = \mathcal{F}_\Sigma$, and place an edge labelled by $d \in \mathcal{D}$ from $F \in \mathcal{F}_\Sigma$ to $F' \in \mathcal{F}_\Sigma$ if $d \in F$, and $F(wd) = F'$ for any word w such that $F = F(w)$. The labelled multi-graph generated is called the *follower set graph* of Σ .

The follower set graph is a cover of Σ [36, Proposition 3.2.9]. Notice that the follower set graph is *right-resolving*, meaning that for any node, all the edges going out have different labels. When the sofic shift is transitive, we would expect the cover to be an irreducible graph (that is, there is a path between each pair of vertices), but this is not always true for the follower set graph. To address this problem, a subgraph can be extracted, which remains a cover and is irreducible.

Definition 2.10. A word $w \in \mathcal{L}(\Sigma)$ is said to be *synchronizing* if for every $u, v \in \mathcal{L}(\Sigma)$, the condition that both uv and wv are in the language implies that $uwxv \in \mathcal{L}(\Sigma)$.

While not every subshift has synchronizing words, it is well known that transitive sofic shifts do. Moreover, every word can be extended to be synchronizing and any extension of a synchronizing word is also synchronizing. In the case of an M -step SFT, every word with length at least $M + 1$ is synchronizing.

An important result due to Fischer [21, 22] establishes the existence and uniqueness (up to isomorphism) of a minimal (in terms of the number of nodes), right-resolving and irreducible cover for a transitive sofic shift. This unique representation is known as the *Fischer cover* and plays an important role in this paper. The following outlines the construction of this cover.

2.3. Construction of the Fischer cover. Let $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ be a transitive sofic shift. For each follower set $F \in \mathcal{F}_\Sigma$, we create a node associated to F if there exists a synchronizing word w such that $F(w) = F$; in other words, we consider the set of nodes

$$\mathcal{V}_\Sigma := \{F(w) : w \in \mathcal{L}(\Sigma) \text{ is synchronizing}\} \subseteq \mathcal{F}_\Sigma.$$

Let $F, F' \in \mathcal{V}_\Sigma$ and $d \in \mathcal{D}$. Then, there exists $w \in \mathcal{L}(\Sigma)$ such that $F = F(w)$. We create an edge with starting node F , terminal node F' and labelled by d if and only if $F' = F(wd)$. The resulting graph, denoted by $\mathcal{G}_\Sigma = (\mathcal{V}_\Sigma, \mathcal{E}_\Sigma, \lambda_\Sigma)$, is the Fischer cover for Σ (see [36, Ch. 3] and [48, Lemma 6.11] for more details).

Thus, the Fischer cover is a subgraph of the follower set graph, constructed by considering only the nodes corresponding to follower sets of synchronizing words. Notice that for any synchronizing word $w \in \mathcal{L}(\Sigma)$, every path in \mathcal{G}_Σ labelled by w must end at the node associated with $F(w)$. In the case where Σ is a mixing sofic subshift, it is possible to find $n \in \mathbb{N}$ such that for every pair of nodes $F, F' \in \mathcal{V}$ and $m \geq n$, there exists a path of length m from F to F' . In such a case, for every pair of words u and v in the language, there exists $w \in \mathcal{L}^m(\Sigma)$ such that $uwv \in \mathcal{L}(\Sigma)$.

2.4. Fractal sets of integers. In [27], Glasscock, Moreira and Richter introduced the notion of multiplicatively invariant sets of integers (recall Definition 1.1), and studied them from a fractal perspective. Here, we focus on the notion of *mass dimension*, which is analogous to the box dimension in fractal geometry, but adapted to the setting of integers.

Definition 2.11. Let $A \subseteq \mathbb{N}_0$ be non-empty. We define the *lower mass dimension* and the *upper mass dimension* of A as

$$\begin{aligned} \underline{\dim}_M(A) &:= \liminf_{N \rightarrow \infty} \frac{\log |A \cap [0, N)|}{\log N} = \sup \left\{ \gamma \geq 0 : \liminf_{N \rightarrow \infty} \frac{|A \cap [0, N)|}{N^\gamma} > 0 \right\}, \\ \overline{\dim}_M(A) &:= \limsup_{N \rightarrow \infty} \frac{\log |A \cap [0, N)|}{\log N} = \sup \left\{ \gamma \geq 0 : \limsup_{N \rightarrow \infty} \frac{|A \cap [0, N)|}{N^\gamma} > 0 \right\}. \end{aligned}$$

If $\underline{\dim}_M(A) = \overline{\dim}_M(A)$, we denote this value by $\dim_M(A)$ and say that the *mass dimension* of A exists.

From [27, Proposition 3.6], the mass dimension exists for any multiplicatively invariant set of integers. A natural way to generate $\times g$ -invariant sets is by considering subshifts of $X = \{0, \dots, g-1\}^{\mathbb{N}_0}$.

Definition 2.12. The *g-language set* associated with a subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ is the set $A_\Sigma \subseteq \mathbb{N}_0$ defined by

$$A_\Sigma := \{(w)_g := w_0 + w_1g + \dots + w_{n-1}g^{n-1} : w = w_0 \dots w_{n-1} \in \mathcal{L}(\Sigma)\}. \quad (2.1)$$

In general, the expansion of a number in a certain base is read from left to right, where the leftmost digit represents the most significant digit. For a given word w , in this paper, we will consider the rightmost digit to be the most significant digit of the integer $(w)_g$, this to be consistent with the theory of subshifts. However, when it comes to applications, this does not make a major distinction.

From the definition of $\times g$ -invariant sets and the shift-invariance of Σ , it is evident that A_Σ is $\times g$ -invariant. It is also possible to see that every multiplicatively invariant set can be written as A_Σ for some subshift Σ and the mass dimension is related to the entropy of the subshift. This is summarized in the following proposition (see [27, Proposition 3.10]).

PROPOSITION 2.13. *The g -language set $A_\Sigma \subseteq \mathbb{N}_0$ corresponding to any non-empty subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ is a $\times g$ -invariant set and the mass dimension of A_Σ is equal to the normalized topological entropy of the subshift (Σ, σ) , that is,*

$$\dim_{\mathbf{M}}(A_\Sigma) = h(\Sigma) / \log(g). \quad (2.2)$$

Moreover, for any $\times g$ -invariant set $A \subseteq \mathbb{N}_0$, there exists a subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ such that $A = A_\Sigma$. In such a case, we say that A is represented by Σ .

COROLLARY 2.14. *Let $g \geq 2$ be an integer and let $\mathcal{D} \subseteq \{0, \dots, g-1\}$. If we consider the full shift $\Sigma = \mathcal{D}^{\mathbb{N}_0}$, $C_{g,\mathcal{D}} = A_\Sigma$. In addition,*

$$\dim_{\mathbf{M}}(C_{g,\mathcal{D}}) = \log(|\mathcal{D}|) / \log(g).$$

Finally, the following result allows us to simplify the calculation of the dimension to simpler intervals.

LEMMA 2.15. [27, Lemma 3.2. (IV)]. *For $r \in \mathbb{N}$, $r \geq 2$,*

$$\overline{\dim}_{\mathbf{M}}(A) = \limsup_{N \rightarrow \infty} \frac{\log |A \cap [0, r^N]|}{N \log r}. \quad (2.3)$$

Analogously, we can compute $\underline{\dim}_{\mathbf{M}}(A)$ by using \liminf instead of \limsup in (2.3).

3. Construction of Markov chains

In §1.3, we provided an overview of the construction of Markov chains in the context of integers with missing digits and their distribution in residue classes, but now we extend this approach to a more general framework. Let f_1, \dots, f_r be g -additive functions, $a_1, \dots, a_r \in \mathbb{N}$ and $A \subseteq \mathbb{N}_0$ a $\times g$ -invariant set. We aim to find conditions to ensure

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A : f_i(n) \equiv b_i \pmod{a_i} \text{ for } i = 1, \dots, r\} \cap [0, N]|}{|A \cap [0, N]|} = \frac{1}{\prod_{i=1}^r a_i} \quad (3.1)$$

for any $(b_1, \dots, b_r) \in \mathbb{Z}_{a_1} \times \dots \times \mathbb{Z}_{a_r}$.

Notation. Throughout this paper, we denote $\mathbf{a} := (a_1, \dots, a_r)$, $|\mathbf{a}| := |\mathbb{Z}_{a_1} \times \dots \times \mathbb{Z}_{a_r}| = \prod_{i=1}^r a_i$ and $\mathbf{f} := (f_1, \dots, f_r)$. For any $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}} := \mathbb{Z}_{a_1} \times \dots \times \mathbb{Z}_{a_r}$ and $n \in \mathbb{N}_0$, we will write $\mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}}$ if $f_j(n) \equiv b_j \pmod{a_j}$ for every $j \in \{1, \dots, r\}$.

From Proposition 2.13, there exists a subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ such that

$$A = A_\Sigma = \{(w)_g : w \in \mathcal{L}(\Sigma)\}.$$

Therefore, we will focus on studying the distribution of words of a certain length and then extend the result to the set of integers A . Similarly to §1.3, we need to find an analogue

to the value p for periodicity and properly understand the concatenation of words in the language.

Definition 3.1. Let f_1, \dots, f_r be g -additive functions and $\mathbf{a} \in \mathbb{N}^r$. We say \mathbf{f} is *eventually periodic with respect to \mathbf{a}* if there exists $\ell \in \mathbb{N}_0$ and $p \in \mathbb{N}$ such that

$$\mathbf{f}(dg^{i+np}) \equiv \mathbf{f}(dg^i) \pmod{\mathbf{a}}$$

for every $d \in \{0, \dots, g-1\}$, $n \in \mathbb{N}$ and $i \geq \ell$. Any pair (p, ℓ) satisfying this condition is called an *eventual period*.

When $\ell' \geq \ell$ and $p \mid p'$, the pair (p', ℓ') is also an eventual period. Notice that \mathbf{f} is eventually periodic if each f_i is eventually periodic. Clearly, the function sum-of-digits in base g satisfies this condition with $p = 1$ and $\ell = 0$. Also, the identity is always g -invariant and eventually periodic.

PROPOSITION 3.2. For any $a, g \in \mathbb{N}$, there exist $p \in \mathbb{N}$ and $\ell \leq a-1$ such that

$$dg^i \equiv dg^{i+np} \pmod{a}$$

for every $d \in \{0, \dots, g-1\}$, $n \in \mathbb{N}$ and $i \geq \ell$.

Proof. At least two elements of $\{g^j \pmod{a} : j \in \{0, \dots, a\}\}$ agree. Then, there exist $j_0 < j_1 \leq a$ such that $g^{j_0} \equiv g^{j_1} \pmod{a}$. We conclude by considering $\ell = j_0$ and $p = j_1 - j_0$. \square

If A is a missing digits set, then $A = A_\Sigma$ for $\Sigma = \mathcal{D}^{\mathbb{N}_0}$, where \mathcal{D} is the respective set of digits. In this context, the concatenation of two words is in $\mathcal{L}(\Sigma)$ if and only if both words are in $\mathcal{L}(\Sigma)$. Unlike full shifts, this property generally does not hold. Nevertheless, it is possible to adapt the construction when concatenation is governed by a suitable rule, as occurs in transitive sofic shifts through the notions of follower sets and the Fischer cover.

As discussed in §2.2, we can identify the nodes of the Fischer cover $\mathcal{G}_\Sigma = (\mathcal{V}_\Sigma, \mathcal{E}_\Sigma, \lambda_\Sigma)$ with the collection of follower sets associated to synchronizing words of the language. Thus, we can identify

$$\mathcal{V}_\Sigma = \{F_1, F_2, \dots, F_q\} = \{F(w) : w \in \mathcal{L}(\Sigma) \text{ is a synchronizing word}\}.$$

The concatenation of words can be understood from \mathcal{G}_Σ . For $x, y \in \mathcal{L}(\Sigma)$,

$$xy \in \mathcal{L}(\Sigma) \iff y \in F(x) \iff \text{there exists a path of } \mathcal{G}_\Sigma \text{ labelled by } xy.$$

When $F(x) \in \mathcal{V}_\Sigma$, it is also equivalent to the existence of a path π in the Fischer cover starting at the node associated to $F(x)$ and it is labelled by y .

Notation. For a word $w \in \mathcal{L}(\Sigma)$ and $F, F' \in \mathcal{V}_\Sigma$, we write $F \xrightarrow{w} F'$ if there is a path π in \mathcal{G}_Σ from F to F' labelled by w . Since \mathcal{G}_Σ is right-resolving, any other path from F to F' has a different label.

Notice that $F \xrightarrow{w} F'$ if and only if $w \in F$, and for every word u such that $F = F(u)$, $F(uw) = F'$.

Remark 3.3. For the rest of this paper, we will consider a base $g \geq 2$, $r \in \mathbb{N}$, $\mathbf{a} \in \mathbb{N}^r$, a collection of g -additive functions $\mathbf{f} = (f_1, \dots, f_r)$ eventually periodic with respect to \mathbf{a} , and an eventual period (p, ℓ) . In addition, we consider a transitive sofic subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$, and $\mathcal{G}_\Sigma = (\mathcal{V}_\Sigma, \mathcal{E}_\Sigma, \lambda_\Sigma)$ its Fischer cover (if there is no confusion, we omit the subscript). For technical reasons, we will assume that ℓ is large enough such that there exists some synchronizing word of length equal to ℓ .

3.1. Transition matrix. To define the transition matrix, we will group the elements of $\mathcal{L}(\Sigma)$ according to how they extend other words. Recalling that $(w_0 \cdots w_n)_g := w_0 + w_1g + \cdots + w_n g^n$, we introduce the following subsets of $\mathcal{L}^p(\Sigma)$.

Definition 3.4. Let $i \in \mathbb{N}$. For any $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}} = \mathbb{Z}_{a_1} \times \cdots \times \mathbb{Z}_{a_r}$ and $F, F' \in \mathcal{V}$, we define the set

$$E_i(\mathbf{b}, F, F') := \{w \in \mathcal{L}^p(\Sigma) : \mathbf{f}(g^i(w)_g) \equiv \mathbf{b} \pmod{\mathbf{a}}, F \xrightarrow{w} F'\}.$$

Notice that E_i depends on $g, \mathbf{f}, \mathbf{a}, \Sigma$ and p , but as in Remark 3.3, we set these elements, but we will not make the dependency explicit to lighten the notation. The following lemma clarifies the motivation for defining these sets.

LEMMA 3.5. Let $x \in \mathcal{L}^i(\Sigma)$ and $w \in E_i(\mathbf{b}, F(x), F')$. Then, $xw \in \mathcal{L}(\Sigma)$, $F(xw) = F'$ and $\mathbf{f}((xw)_g) \equiv \mathbf{f}((x)_g) + \mathbf{b} \pmod{\mathbf{a}}$. If $i \geq \ell$, it also holds that $E_i = E_{i+np}$ for all $n \in \mathbb{N}$.

Proof. Since $F(x) \xrightarrow{w} F'$, $xw \in \mathcal{L}(\Sigma)$ and $F(xw) = F'$. Also,

$$\mathbf{f}((xw)_g) = \mathbf{f}\left(\sum_{j=0}^{i-1} x_j g^j + \sum_{j=i}^{i+p-1} w_j g^j\right) = \mathbf{f}((x)_g) + \mathbf{f}(g^i(w)_g) \equiv \mathbf{f}((x)_g) + \mathbf{b} \pmod{\mathbf{a}}.$$

If $i \geq \ell$, $\mathbf{f}(g^i(y)_g) = \mathbf{f}(g^{i+np}(y)_g)$ for every $n \in \mathbb{N}$ and $y \in \mathcal{L}^p(\Sigma)$, concluding that $E_i = E_{i+np}$. \square

Despite that the collection of sets E_i is not a partition of \mathcal{L}^p (since a word might be read along different paths in \mathcal{G}), we still can understand it as a partition of the paths of length p in \mathcal{G} .

PROPOSITION 3.6. Let $i \in \mathbb{N}$. For every $F \in \mathcal{V}$,

$$\sum_{(\mathbf{b}, F') \in \mathbb{Z}_{\mathbf{a}} \times \mathcal{V}} |E_i(\mathbf{b}, F, F')| = |\{\pi = (\pi_0 \cdots \pi_{p-1}) \text{ is a path of } \mathcal{G} \text{ and } s(\pi) = F\}|. \quad (3.2)$$

Similarly, for every $F' \in \mathcal{V}$,

$$\sum_{(\mathbf{b}, F) \in \mathbb{Z}_{\mathbf{a}} \times \mathcal{V}} |E_i(\mathbf{b}, F, F')| = |\{\pi = (\pi_0 \cdots \pi_{p-1}) \text{ is a path of } \mathcal{G} \text{ and } t(\pi) = F'\}|. \quad (3.3)$$

Proof. Let $F, F' \in \mathcal{V}$. Notice that the sets $E_i(\mathbf{b}, F, F')$ and $E_i(\mathbf{b}', F, F')$ are disjoint when $\mathbf{b} \neq \mathbf{b}' \in \mathbb{Z}_{\mathbf{a}}$. For any $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$, if $w \in E_i(\mathbf{b}, F, F')$, there exists a path π in \mathcal{G} of length p such that $\lambda(\pi) = w$ from F to F' . However, any path π of length p from F to F' represents a word $x := \lambda(\pi)$ of $E_i(\mathbf{b}, F, F')$, where $\mathbf{b} := \mathbf{f}(g^i(x)_g)$. Since F and F' are fixed, each of these paths represents a different word. Then,

$$\begin{aligned}
|\{\pi = (\pi_0 \cdots \pi_{p-1}) \text{ is a path in } \mathcal{G} \text{ from } F \text{ to } F'\}| &= \left| \bigcup_{\mathbf{b} \in \mathbb{Z}_a} E_i(\mathbf{b}, F, F') \right| \\
&= \sum_{\mathbf{b} \in \mathbb{Z}_a} |E_i(\mathbf{b}, F, F')|.
\end{aligned}$$

By summing over F' , the equality (3.2) is obtained and similarly, (3.3) follows by summing over F . \square

Usually, (3.2) and (3.3) do not agree. For this work, we require an extra regularity of \mathcal{G} to ensure that the number of paths of length p starting at some $F \in \mathcal{V}$ is equal to the number of paths of length p with terminal node $F' \in \mathcal{V}$. From graph theory, a directed multi-graph is said to be k -regular if each vertex has k edges going in and k edges going out (two edges are different if they differ in the starting node, terminal node or label). We can pose this definition in subshifts.

Definition 3.7. We say a transitive sofic shift is k -regular if the graph representation given by the Fischer cover is a k -regular graph. We say that a transitive sofic shift is regular if it is k -regular for some $k \geq 2$.

If a graph is 1-regular, it must represent a full shift in one symbol. Thus, we are interested in the case $k \geq 2$, so we include it directly in the definition. Trivially, any full shift satisfies this condition, with k being the number of symbols. When Σ is k -regular and $F \in \mathcal{V}$, there are k^p paths of length p starting from F and k^p paths of length p ending at F . From this idea, we can define a transition matrix for the Markov chains.

Definition 3.8. In the framework of Remark 3.3 and assuming Σ is k -regular, we define the state space $\mathcal{S} := \mathbb{Z}_a \times \mathcal{V}$ and the matrix M_i indexed by $\mathcal{S} \times \mathcal{S}$ as

$$M_i((\mathbf{b}, F), (\mathbf{b}', F')) := k^{-p} \cdot |E_i(\mathbf{b}' - \mathbf{b}, F, F')|$$

for any $(\mathbf{b}, F), (\mathbf{b}', F') \in \mathcal{S}$.

Clearly, the matrix M_i depends on several elements, including the choice of p , but we will not make the dependence explicit. The regularity condition allows us to show that this matrix M_i is suitable for defining a Markov chain.

PROPOSITION 3.9. For any $i \in \mathbb{N}$, the matrix M_i is a doubly stochastic matrix.

Proof. Let $(\mathbf{b}, F) \in \mathcal{S}$. From (3.2),

$$\begin{aligned}
\sum_{s \in \mathcal{S}} M_i((\mathbf{b}, F), s) &= k^{-p} \sum_{(\mathbf{b}', F') \in \mathcal{S}} |E_i(\mathbf{b}' - \mathbf{b}, F, F')| \\
&= k^{-p} \sum_{(\hat{\mathbf{b}}, F') \in \mathcal{S}} |E_i(\hat{\mathbf{b}}, F, F')| \\
&= k^{-p} |\{\pi = (\pi_0 \pi_1 \cdots \pi_{p-1}) \text{ is a path in } \mathcal{G} \text{ and } s(\pi) = F\}|.
\end{aligned}$$

We conclude that $\sum_{s \in \mathcal{S}} M_i((\mathbf{b}, F), s) = 1$ from k -regularity. Analogously, from (3.3), we obtain $\sum_{s \in \mathcal{S}} M_i(s, (\mathbf{b}, F)) = 1$. \square

Note that the regularity is important to ensure that the normalization factor to make the matrix stochastic is the same for all rows. Without this, while a stochastic matrix could still be defined, the link between M_i and the distribution of integers becomes unclear, and our method no longer seems to work.

3.2. Initial distribution. In our attempt to study A_Σ , the idea is to extend all the ℓ -long words that have a follower set in \mathcal{V} , as we can fully describe those extensions using the Fischer cover. However, for some $w \in \mathcal{L}^\ell$, it might be that $F(w) \notin \mathcal{V}$. This leads to the following definition.

Definition 3.10. Let Σ be a transitive sofic shift. We define the ℓ -restricted language of Σ as all words of $\mathcal{L}(\Sigma)$ that are extensions of a word of length ℓ with follower set in \mathcal{V} ,

$$\mathcal{L}_{\mathcal{V},\ell}(\Sigma) := \{w \in \mathcal{L}(\Sigma) : |w| \geq \ell \text{ and } F(w_0 \cdots w_{\ell-1}) \in \mathcal{V}\}.$$

For every $i \geq \ell$, we also define $\mathcal{L}_{\mathcal{V},\ell}^i(\Sigma) := \{w \in \mathcal{L}_{\mathcal{V},\ell}(\Sigma) : |w| = i\}$.

The reason for considering ℓ large enough in Remark 3.3 is to ensure that this set is not empty. If Σ is an M -step shift of finite type, any word w such that $|w| \geq M + 1$ is synchronizing; therefore, $\mathcal{L}_{\mathcal{V},\ell}^i(\Sigma) = \mathcal{L}^i(\Sigma)$ for all $i \geq M + 1$. Under the condition of regularity, it is easy to describe the cardinality of $\mathcal{L}_{\mathcal{V},\ell}^i$. For the upcoming results, if $w = w_0 \cdots w_n$ is a word, we will identify it with a pair of words (x, y) such that $w = xy$.

PROPOSITION 3.11. Let Σ be a transitive sofic shift and k -regular. Then,

$$|\mathcal{L}_{\mathcal{V},\ell}^i(\Sigma)| = |\mathcal{L}_{\mathcal{V},\ell}^\ell(\Sigma)| \cdot k^{i-\ell}$$

for every $i \geq \ell$.

Proof. Notice $x \in \mathcal{L}_{\mathcal{V},\ell}^i$ if and only if $F := F(x_0 \cdots x_{\ell-1}) \in \mathcal{V}$ and there is a path π in the Fischer cover of length $i - \ell$ from the vertex associated to F , and π is labelled by $x_\ell \cdots x_{i-1}$. Using the identification,

$$\mathcal{L}_{\mathcal{V},\ell}^i = \bigcup_{F \in \mathcal{V}} \{w \in \mathcal{L}_{\mathcal{V},\ell}^\ell : F(w) = F\} \times \{\lambda(\pi) : \pi = \pi_0 \cdots \pi_{i-\ell-1} \in \mathcal{G}, s(\pi) = F\}. \quad (3.4)$$

Since the Fischer cover is right-resolving, all paths in $\{\pi = \pi_0 \cdots \pi_{i-\ell-1} \in \mathcal{G}, s(\pi) = F\}$ have different labels. Then, the conclusion follows by using the k -regularity in (3.4). \square

We introduce a probability measure on \mathcal{S} .

Definition 3.12. For every $i \geq \ell$, the distribution μ_i on $\mathcal{S} = \mathbb{Z}_a \times \mathcal{V}$ is defined as

$$\mu_i(\{(\mathbf{b}, F)\}) := \frac{|\{w \in \mathcal{L}_{\mathcal{V},\ell}^i(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \pmod{a}, F(w) = F\}|}{|\mathcal{L}_{\mathcal{V},\ell}^i(\Sigma)|}$$

for any $(\mathbf{b}, F) \in \mathcal{S}$.

We will denote $\mu_i(\mathbf{b}, F) := \mu_i(\{(\mathbf{b}, F)\})$. By understanding the powers of the matrix M_i , we can describe the evolution of the measure μ_{i+np} with respect to $n \in \mathbb{N}$.

PROPOSITION 3.13. *For any $i, n \in \mathbb{N}$ and $s = (\mathbf{b}, F), s' = (\mathbf{b}', F') \in \mathcal{S}$,*

$$M_i^n(s, s') = k^{-pn} |\{w \in \mathcal{L}^{pn}(\Sigma) : \mathbf{f}(g^i(w)_g) \equiv \mathbf{b}' - \mathbf{b} \bmod \mathbf{a}, F \xrightarrow{w} F'\}|. \quad (3.5)$$

Moreover, for every $i \geq \ell$ and $n \in \mathbb{N}$,

$$\mu_{i+np} = \mu_i M_i^n. \quad (3.6)$$

Proof. We proceed by induction to show (3.5). Let $s = (\mathbf{b}, F), s' = (\mathbf{b}', F') \in \mathcal{S}$. The case $n = 1$ holds by definition. Suppose that the result holds for some $n \in \mathbb{N}$, so we have to prove the result for $n + 1$. Since

$$M_i^{n+1}(s, s') = \sum_{\tilde{s} \in \mathcal{S}} M_i^n(s, \tilde{s}) M_i(\tilde{s}, s'),$$

from the induction hypotheses, it will be sufficient to prove that

$$\begin{aligned} & \{w \in \mathcal{L}^{p(n+1)}(\Sigma) : \mathbf{f}(g^i(w)_g) \equiv \mathbf{b}' - \mathbf{b} \bmod \mathbf{a}, F \xrightarrow{w} F'\} \\ &= \bigcup_{(\tilde{\mathbf{b}}, \tilde{F}) \in \mathcal{S}} (\{x \in \mathcal{L}^{pn}(\Sigma) : \mathbf{f}(g^i(x)_g) \equiv \tilde{\mathbf{b}} - \mathbf{b} \bmod \mathbf{a}, F \xrightarrow{x} \tilde{F}\} \\ & \quad \times \{y \in \mathcal{L}^p(\Sigma) : \mathbf{f}(g^i(y)_g) \equiv \mathbf{b}' - \tilde{\mathbf{b}} \bmod \mathbf{a}, \tilde{F} \xrightarrow{y} F'\}). \end{aligned} \quad (3.7)$$

Let $w \in \mathcal{L}^{p(n+1)}$ in the left-hand side of (3.7) and consider the decomposition $w = (x, y)$, where $x = w_0 \cdots w_{pn-1}$ and $y = w_{pn} \cdots w_{n(p+1)-1}$. Since $\mathbf{f}(g^i(w)_g) = \mathbf{f}(g^i(x)_g) + \mathbf{f}(g^{i+pn}(y)_g)$, the eventual periodicity leads to

$$\mathbf{f}(g^i(w)_g) \equiv_a \mathbf{b}' - \mathbf{b} \text{ if and only if } \mathbf{f}(g^i(x)_g) \equiv_a \tilde{\mathbf{b}} - \mathbf{b} \text{ and } \mathbf{f}(g^i(y)_g) \equiv_a \mathbf{b}' - \tilde{\mathbf{b}},$$

where $\tilde{\mathbf{b}} := \mathbf{f}(g^i(x)_g) + \mathbf{b} \bmod \mathbf{a}$. In addition, $F \xrightarrow{w} F'$ is equivalent to the existence of a path π in \mathcal{G} of length $(n+1)p$ from F to F' labelled by w . If we denote by \tilde{F} the terminal node for the subpath $\pi_0 \cdots \pi_{np-1}$, $F \xrightarrow{w} F'$ if and only if $F \xrightarrow{x} \tilde{F}$ and $\tilde{F} \xrightarrow{y} F'$. Thus, we conclude (3.7). However,

$$(\mu_i M_i^n)(s) = \sum_{s' \in \mathcal{S}} \mu_i(s') M_i^n(s', s).$$

By Proposition 3.11, $k^{pn} |\mathcal{L}_{\mathcal{V}, \ell}^i(\Sigma)| = |\mathcal{L}_{\mathcal{V}, \ell}^{i+np}(\Sigma)|$. Hence, using (3.5), we can conclude (3.6) by showing that

$$\begin{aligned} & \{w \in \mathcal{L}_{\mathcal{V}, \ell}^{i+(n+1)p}(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, F(w) = F\} \\ &= \bigcup_{s' \in \mathcal{S}} (\{x \in \mathcal{L}_{\mathcal{V}, \ell}^i(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b}' \bmod \mathbf{a}, F(w) = F'\} \\ & \quad \times \{y \in \mathcal{L}^{pn}(\Sigma) : \mathbf{f}(g^i(y)_g) \equiv \mathbf{b} - \mathbf{b}' \bmod \mathbf{a}, F' \xrightarrow{y} F\}). \end{aligned} \quad (3.8)$$

If $w \in \mathcal{L}_{\mathcal{V}, \ell}^{i+(n+1)p}(\Sigma)$, it is straightforward that $w_0 \cdots w_{i-1} \in \mathcal{L}_{\mathcal{V}, \ell}^i(\Sigma)$. The rest of the proof follows by replicating the argument used to prove (3.7). \square

From the previous results, we can characterize the distribution using Markov chains.

Definition 3.14. Let $g, \mathbf{a}, \mathbf{f}, p, \ell$ and Σ be as defined in Remark 3.3, and assume that the subshift Σ is k -regular (where $k \geq 2$). For $i \geq \ell$, we denote by $X^i = (X_n^i)_{n \in \mathbb{N}_0}$ a Markov chain on the state space \mathcal{S} with transition matrix M_i (Definition 3.8) and initial distribution μ_i (Definition 3.12). We denote by \mathbb{P} the probability measure associated to the process. Notice that $X^i = X^i(g, \mathbf{a}, \mathbf{f}, p, \ell, \Sigma)$, but we omit this dependency in the notation for simplicity.

Remark 3.15. Let $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ be a 1-step shift of finite type. For any $w = w_0 \cdots w_n \in \mathcal{L}(\Sigma)$, $F(w) = F(w_n)$. Thus, $\mathcal{V}_\Sigma = \{F(d) : d \in \mathcal{D}\}$. Note that the cover presented in Remark 2.7 is the Fischer cover for Σ , except if there exist $d \neq d' \in \mathcal{D}$ such that $F(d) = F(d')$. In such a case, the symbols d and d' are associated to the same node in the Fischer cover.

If the graph of Remark 2.7 is irreducible and regular, we can use this cover to make the construction of the Markov chains instead of the Fischer cover. The only distinction is that different symbols are forced to be associated with different nodes, but the rest of the construction follows immediately. It is easy to see that the results that we will present in §4 also hold when using this cover.

4. From Markov chains to uniform distribution

Throughout this section, we will work within the framework outlined in Remark 3.3 and assume that Σ is a k -regular subshift for some $k \geq 2$. The connection between the Markov chains introduced in Definition 3.14 and the distribution of words in the ℓ -restricted language $\mathcal{L}_{\mathcal{V}, \ell}(\Sigma)$ is straightforward from Proposition 3.13.

COROLLARY 4.1. For any $i \geq \ell$, $(\mathbf{b}, F) \in \mathcal{S}$ and $n \in \mathbb{N}$,

$$\mathbb{P}(X_n^i = (\mathbf{b}, F)) = \frac{|\{w \in \mathcal{L}_{\mathcal{V}, \ell}^{i+np}(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, F(w) = F\}|}{|\mathcal{L}_{\mathcal{V}, \ell}^{i+np}(\Sigma)|}.$$

Proof. Given the Markov chain X^i with initial distribution μ_i and transition matrix M_i , for every $(\mathbf{b}, F) \in \mathcal{S}$ and $n \in \mathbb{N}$,

$$\mathbb{P}(X_n^i = (\mathbf{b}, F)) = \sum_{s \in \mathcal{S}} \mu_i(s) M_i(s, (\mathbf{b}, F)).$$

The conclusion follows from Definition 3.12 and (3.6) of Proposition 3.13. \square

Since we are restricted to words in $\mathcal{L}_{\mathcal{V}, \ell}(\Sigma) \subseteq \mathcal{L}(\Sigma)$, it is natural to consider the subset of A_Σ defined by

$$A_{\Sigma, \ell} := \left\{ \sum_{i=0}^n w_i g^i : n \geq \ell - 1, w_0 \cdots w_n \in \mathcal{L}_{\mathcal{V}, \ell}(\Sigma) \right\}.$$

Through X_n^i , the distribution of integers represented by words in $\mathcal{L}_{\mathcal{V}, \ell}^{i+np}$ can be understood. By running different Markov chains for $i = \ell, \ell + 1, \dots, \ell + (p - 1)$, we can recover the distribution in residue classes for $A_{\Sigma, \ell}$. Furthermore, as ℓ increases, the set difference

$A_\Sigma \setminus A_{\Sigma, \ell}$ becomes negligible in terms of relative density, allowing us to recover the distribution for A_Σ .

To ensure the convergence of the Markov chains, we introduce the following definition.

Definition 4.2. Let $g, \mathbf{a}, \mathbf{f}, p, \ell$ and Σ be as defined in Remark 3.3, and assume that the subshift Σ is regular. We say that the *Markov condition* holds if M_i (Definition 3.8) is irreducible and aperiodic for every $i \geq \ell$.

If $i \geq \ell$, $M_i = M_{i+np}$ for all $n \in \mathbb{N}$, so the Markov condition it is equivalent to shows that M_i is irreducible and aperiodic for $i \in \{\ell, \dots, \ell + (p-1)\}$. Notice that the Markov condition implicitly assumes that Σ is mixing (while the construction only requires transitivity). Assuming the convergence of the Markov chains and using Corollary 4.1, we devote the rest of the section to obtain the uniform distribution $\mathbf{f} \bmod \mathbf{a}$ for the $\times g$ -invariant set A_Σ .

LEMMA 4.3. *If the Markov condition holds, there exist $C > 0$ and $\rho \in (0, 1)$ such that*

$$\left| \frac{|\{w \in \mathcal{L}_{V, \ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, F(w) = F\}|}{|\mathcal{L}_{V, \ell}^n(\Sigma)|} - \frac{1}{|\mathcal{S}|} \right| \leq C\rho^n \quad (4.1)$$

for every $n \geq \ell$ and $(\mathbf{b}, F) \in \mathcal{S}$. In addition,

$$\left| \frac{|\{w \in \mathcal{L}_{V, \ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, F(w) = F\}|}{|\{w \in \mathcal{L}_{V, \ell}^n(\Sigma) : F(w) = F\}|} - \frac{1}{|\mathbf{a}|} \right| \leq C\rho^n. \quad (4.2)$$

Proof. Notice that the uniform measure $1/|\mathcal{S}|$ is invariant for M_i by Proposition 3.9. From Theorem 2.1, for every $i \in \{\ell, \dots, \ell + p - 1\}$, there exist $C_i > 0$ and $\rho_i \in (0, 1)$ such that

$$\left| \mathbb{P}(X_n^i = (\mathbf{b}, F)) - \frac{1}{|\mathcal{S}|} \right| \leq C_i \rho_i^n \quad (4.3)$$

for every $(\mathbf{b}, F) \in \mathcal{S}$ and $n \in \mathbb{N}$. By choosing $C' := \max_i \{C_i \rho_i^{-i/p}\}$ and $\rho := \max_i \{\rho_i^{1/p}\}$,

$$\left| \mathbb{P}(X_n^i = (\mathbf{b}, F)) - \frac{1}{|\mathcal{S}|} \right| \leq C' \rho^{np+i} \quad (4.4)$$

for every $i \in \{\ell, \dots, \ell + (p-1)\}$. From Corollary 4.1, we can write

$$\left| \frac{|\{w \in \mathcal{L}_{V, \ell}^{np+i}(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}; F(w) = F\}|}{|\mathcal{L}_{V, \ell}^{np+i}(\Sigma)|} - \frac{1}{|\mathcal{S}|} \right| \leq C' \rho^{np+i}.$$

Since it holds for every $i \in \{\ell, \dots, \ell + (p-1)\}$ and $n \in \mathbb{N}$, (4.1) is obtained. Equivalently, for every $n \in \mathbb{N}$,

$$\begin{aligned} |\mathcal{L}_{V, \ell}^n(\Sigma)| \left(\frac{1}{|\mathcal{S}|} - C' \rho^n \right) &\leq |\{w \in \mathcal{L}_{V, \ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}; F(w) = F\}| \\ &\leq |\mathcal{L}_{V, \ell}^n(\Sigma)| \left(\frac{1}{|\mathcal{S}|} + C' \rho^n \right). \end{aligned} \quad (4.5)$$

By summing (4.5) over all $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$, we obtain

$$|\mathcal{L}_{\mathcal{V},\ell}^n(\Sigma)| \left(\frac{1}{|\mathcal{V}|} - C'|\mathbf{a}|\rho^n \right) \leq |\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : F(w) = F\}|. \quad (4.6)$$

Let $n_0 \in \mathbb{N}$ be large enough such that $1 - C'|\mathcal{S}|\rho^{n_0} > 1/2$. Combining (4.5) and (4.6),

$$\begin{aligned} & |\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}; F(w) = F\}| \\ & \leq \left(\frac{1}{|\mathcal{S}|} + C'\rho^n \right) \left(\frac{1}{|\mathcal{V}|} - C'|\mathbf{a}|\rho^n \right)^{-1} |\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : F(w) = F\}| \end{aligned} \quad (4.7)$$

for every $n \geq n_0$. A quick computation leads to

$$\left(\frac{1}{|\mathcal{S}|} + C'\rho^n \right) \left(\frac{1}{|\mathcal{V}|} - C'|\mathbf{a}|\rho^n \right)^{-1} = \frac{1}{|\mathbf{a}|} + 2C'|\mathcal{V}| \frac{\rho^n}{1 - C'|\mathcal{S}|\rho^n} \leq \frac{1}{|\mathbf{a}|} + 4C'|\mathcal{V}|\rho^n.$$

Using the previous estimation in (4.7),

$$\frac{|\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv_{\mathbf{a}} \mathbf{b}; F(w) = F\}|}{|\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : F(w) = F\}|} \leq \frac{1}{|\mathbf{a}|} + 4C'|\mathcal{V}|\rho^n$$

for every $n \geq n_0$. With the obvious modifications, we can obtain in an analogous way that

$$\frac{|\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : \mathbf{f}((w)_g) \equiv_{\mathbf{a}} \mathbf{b}; F(w) = F\}|}{|\{w \in \mathcal{L}_{\mathcal{V},\ell}^n(\Sigma) : F(w) = F\}|} \geq \frac{1}{|\mathbf{a}|} - \frac{2C'|\mathcal{V}|\rho^n}{1 + C'|\mathcal{S}|\rho^n} \geq \frac{1}{|\mathbf{a}|} - 4C'|\mathcal{V}|\rho^n.$$

The conclusion follows by choosing $C > 0$ large enough. \square

When $0 \notin \mathcal{L}(\Sigma)$, each element of $A_{\Sigma,\ell}$ has a unique representation in the language, which is no longer true if $0 \in \mathcal{L}(\Sigma)$. This lack of uniqueness can introduce complications when analysing the distribution of integers, as multiple representations may correspond to the same integer. We deal with this situation in the following result by using follower sets and the estimation (4.2).

LEMMA 4.4. *If the Markov condition holds, there exist $C > 0$ and $\rho \in (0, 1)$ such that*

$$\left| \frac{|\{n \in A_{\Sigma,\ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, g^m]|}{|A_{\Sigma,\ell} \cap [0, g^m]|} - \frac{1}{|\mathbf{a}|} \right| \leq C\rho^m$$

for every $m \geq \ell$ and $(\mathbf{b}, F) \in \mathcal{S}$.

Proof. For $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$, $F \in \mathcal{V}$ and $t \geq \ell$, we denote

$$\begin{aligned} \mathcal{L}_{\mathcal{V},\ell}^t(\mathbf{b}, F) &:= \{w \in \mathcal{L}_{\mathcal{V},\ell}^t(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, F(w) = F\}, \\ \mathcal{L}_{\mathcal{V},\ell}^t(\mathbf{b}) &:= \{w \in \mathcal{L}_{\mathcal{V},\ell}^t(\Sigma) : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}\} \text{ and} \\ \mathcal{L}_{\mathcal{V},\ell}^t(F) &:= \{w \in \mathcal{L}_{\mathcal{V},\ell}^t(\Sigma) : F(w) = F\}. \end{aligned}$$

If two different words represent the same integer, they must have different lengths. Additionally, any word of length t ending with the digit 0 represents an integer that can also be represented by a word of length $t - 1$. Therefore,

$$\begin{aligned}
& |\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, g^m)| \\
&= |\mathcal{L}_{\mathcal{V}, \ell}^{\ell}(\mathbf{b})| + \sum_{t=\ell+1}^m |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}, w_{t-1} \neq 0\}|. \quad (4.8)
\end{aligned}$$

Any word of length t that ends at 0 must be an extension by 0 of a word w of length $t-1$. Recalling that $\mathbf{f}(0) = 0$, and since 0 can extend w if and only if $0 \in F(w)$, it follows that

$$|\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}; w_{t-1} \neq 0\}| = |\mathcal{L}_{\mathcal{V}, \ell}^t(\mathbf{b})| - \sum_{F \in \mathcal{V}} |\mathcal{L}_{\mathcal{V}, \ell}^{t-1}(\mathbf{b}, F)| \cdot \mathbb{1}_F(0). \quad (4.9)$$

Let $C > 0$ and $\rho \in (0, 1)$ be given by Lemma 4.3. From (4.1),

$$|\mathcal{L}_{\mathcal{V}, \ell}^t(\mathbf{b})| \leq |\mathcal{L}_{\mathcal{V}, \ell}^t| \left(\frac{1}{|\mathbf{a}|} + C|\mathcal{V}|\rho^n \right). \quad (4.10)$$

From (4.9), (4.10) and the estimation provided by (4.2) of Lemma 4.3 for $|\mathcal{L}_{\mathcal{V}, \ell}^{t-1}(\mathbf{b}, F)|$,

$$\begin{aligned}
& |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t : \mathbf{f}((w)_g) \equiv \mathbf{b} \bmod \mathbf{a}; w_{t-1} \neq 0\}| \\
&\leq \frac{1}{|\mathbf{a}|} \left(|\mathcal{L}_{\mathcal{V}, \ell}^t| - \sum_{F \in \mathcal{V}} |\mathcal{L}_{\mathcal{V}, \ell}^{t-1}(F)| \mathbb{1}_F(0) \right) + C|\mathcal{V}|\rho^t |\mathcal{L}_{\mathcal{V}, \ell}^t| + C \sum_{F \in \mathcal{V}} \rho^{t-1} |\mathcal{L}_{\mathcal{V}, \ell}^{t-1}(F)| \mathbb{1}_F(0) \\
&\leq \frac{1}{|\mathbf{a}|} |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t : w_{t-1} \neq 0\}| + C'\rho^t |\mathcal{L}_{\mathcal{V}, \ell}^t|
\end{aligned}$$

for some constant $C' > C$. Using the previous estimation and (4.10) in (4.8),

$$\begin{aligned}
& |\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, g^m)| \\
&\leq \frac{1}{|\mathbf{a}|} \left(|\mathcal{L}_{\mathcal{V}, \ell}^{\ell}| + \sum_{t=\ell+1}^m |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t : w_{t-1} \neq 0\}| \right) + C\rho^{\ell} |\mathcal{L}_{\mathcal{V}, \ell}^{\ell}| + C' \sum_{t=\ell+1}^m \rho^t |\mathcal{L}_{\mathcal{V}, \ell}^t| \\
&= \frac{1}{|\mathbf{a}|} |A_{\Sigma, \ell} \cap [0, g^m)| + C' \sum_{t=\ell}^m \rho^t |\mathcal{L}_{\mathcal{V}, \ell}^t|. \quad (4.11)
\end{aligned}$$

Since $|A_{\Sigma, \ell} \cap [0, g^m)| \geq |\mathcal{L}_{\mathcal{V}, \ell}^m|$, Proposition 3.11 leads to

$$|A_{\Sigma, \ell} \cap [0, g^m)|^{-1} \sum_{t=\ell}^m \rho^t |\mathcal{L}_{\mathcal{V}, \ell}^t| \leq \sum_{t=\ell}^m \rho^t \frac{|\mathcal{L}_{\mathcal{V}, \ell}^t|}{|\mathcal{L}_{\mathcal{V}, \ell}^m|} \leq \sum_{t=\ell}^m \rho^t k^{-(m-t)} \leq m\rho^m,$$

where in the last inequality, we assumed that $k^{-1} \leq \rho$ (otherwise, we change ρ by k^{-1}). Hence, it follows from (4.11) that

$$\frac{|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, g^m)|}{|A_{\Sigma, \ell} \cap [0, g^m)|} - \frac{1}{|\mathbf{a}|} \leq C'm\rho^m.$$

Following the same approach, the analogue lower inequality can be obtained. If $r = \rho^{1/2}$ and $C'' > 0$ is large enough, it holds that $C'm\rho^m \leq C''r^m$ for all $m \in \mathbb{N}$, concluding the result. \square

Let $N \in \mathbb{N}$ and consider its expansion in base g , $N = d_0 + d_1g + \cdots + w_mg^m$, where $w_m \neq 0$. We can write $[0, N) = [0, g^m) \cup [g^m, N)$ and using Lemma 4.4, the distribution can be recovered when restricted to the interval $[0, g^m)$. The next lemma focuses on addressing the distribution in the interval $[g^m, N)$.

LEMMA 4.5. *If the Markov condition holds, there exist $C > 0$ and $\rho \in (0, 1)$ such that for all $t, m \in \mathbb{N}$ (with $t < m$) and for any sequence of digits $d_t, d_{t+1}, \dots, d_m \in \{0, \dots, g-1\}$ where $d_m \neq 0$, defining $M := d_mg^m + \cdots + d_{t+1}g^{t+1}$ and assuming $A_{\Sigma, \ell} \cap [M, M + d_tg^t] \neq \emptyset$, the following estimation holds for every $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$:*

$$\left| \frac{|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [M, M + d_tg^t]\}|}{|A_{\Sigma, \ell} \cap [M, M + d_tg^t]|} - \frac{1}{|\mathbf{a}|} \right| \leq C\rho^t.$$

Proof. Let C and ρ be as given by Lemma 4.3, and let $d \in \{0, \dots, g-1\}$. Let

$$x \in \{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [M + dg^t, M + (d+1)g^t]\}. \quad (4.12)$$

Since $x \in [M + dg^t, M + (d+1)g^t]$, the representation of x can be written as $x = (x_0 \cdots x_m)_g$, where $x_t = d$ and $x_{t+1} \cdots x_m = d_{t+1} \cdots d_m$. Therefore, using the g -additivity, we can see that $\mathbf{f}(x) \equiv \mathbf{b} \pmod{\mathbf{a}}$ is equivalent to

$$\mathbf{f}((x_0 \cdots x_{t-1})_g) \equiv \mathbf{b} - \mathbf{f}(dg^t + d_{t+1}g^{t+1} + \cdots + d_mg^m) \pmod{\mathbf{a}} =: \tilde{\mathbf{b}}(d).$$

Additionally, as $d_m \neq 0$, $x = (x_0 \cdots x_{t-1}dd_{t+1} \cdots d_m)_g \in A_{\Sigma, \ell}$ is equivalent to

$$x_0 \cdots x_{t-1} \in \mathcal{L}_{\mathcal{V}, \ell}(\Sigma) \quad \text{and} \quad dd_{t+1} \cdots d_m \in F(x_0 \cdots x_{t-1}).$$

Therefore, (4.12) is equivalent to

$$\begin{aligned} x_0 \cdots x_{t-1} &\in \{w \in \mathcal{L}_{\mathcal{V}, \ell}^t(\Sigma) : \mathbf{f}((w)_g) \equiv \tilde{\mathbf{b}}(d) \pmod{\mathbf{a}}\} \quad \text{and} \\ dd_{t+1} \cdots d_m &\in F(x_0 \cdots x_{t-1}). \end{aligned}$$

Using this equivalence and the estimation (4.2) of Lemma 4.5,

$$\begin{aligned} &|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [M + dg^t, M + (d+1)g^t]\}| \\ &= \sum_{F \in \mathcal{V}} |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t(\Sigma) : \mathbf{f}((w)_g) \equiv \tilde{\mathbf{b}}(d) \pmod{\mathbf{a}}; F(w) = F\}| \cdot \mathbb{1}_F(dd_{t+1} \cdots d_m) \\ &\leq \sum_{F \in \mathcal{V}} \left(\frac{1}{|\mathbf{a}|} + C\rho^t \right) |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t(\Sigma) : F(w) = F\}| \cdot \mathbb{1}_F(dd_{t+1} \cdots d_m). \end{aligned} \quad (4.13)$$

With an analogous argument as above, we can obtain that

$$\begin{aligned} &\sum_{F \in \mathcal{V}} |\{w \in \mathcal{L}_{\mathcal{V}, \ell}^t(\Sigma) : F(w) = F\}| \cdot \mathbb{1}_F(dd_{t+1} \cdots d_m) \\ &= |A_{\Sigma, \ell} \cap [M + dg^t, M + (d+1)g^t]|. \end{aligned}$$

From the assumption $A_{\Sigma, \ell} \cap [M, M + d_tg^t] \neq \emptyset$, it is clear that $d_t \geq 1$. Thus, we can use the previous identity in (4.13), and then sum from $d = 0$ to $d_t - 1$, obtaining

$$|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv_{\mathbf{a}} \mathbf{b}\} \cap [M, M + d_t g^t]| \leq \left(\frac{1}{|\mathbf{a}|} + C\rho^t \right) |A_{\Sigma, \ell} \cap [M, M + d_t g^t]|.$$

With a similar argument, the lower inequality can be obtained, concluding the result. \square

Putting the previous results together, the distribution for the set $A_{\Sigma, \ell}$ can be estimated for an arbitrary interval $[0, N)$.

PROPOSITION 4.6. *Let $g \geq 2$, $\mathbf{a} \in \mathbb{N}^r$ and $\mathbf{f} = (f_1, \dots, f_r)$ be a collection of g -additive functions eventually periodic with respect to \mathbf{a} . Let $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ be a mixing, sofic and regular subshift, and (p, ℓ) be an eventual period such that ℓ is at least the length of the shortest synchronizing word.*

If the Markov condition holds, there exists $\gamma > 0$ such that

$$\frac{|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, N)|}{|A_{\Sigma, \ell} \cap [0, N)|} = \frac{1}{a_1 \cdot a_2 \cdots a_r} + O(N^{-\gamma})$$

for every $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$.

Proof. Since the Markov condition holds, we can apply Lemma 4.4 and Lemma 4.5 with constants $C > 0$ and $\rho \in (0, 1)$. Let $N \in \mathbb{N}$ and its base- g expansion $N = d_m g^m + \dots + d_1 g + d_0$, where $d_m \neq 0$. For every $t \in \{0, \dots, m\}$, we define $N(t) := d_m g^m + \dots + d_t g^t$. Noting

$$[0, N) = [0, g^m) \cup [g^m, d_m g^m) \cup \left(\bigcup_{t=0}^{m-1} [N(t+1), N(t)) \right), \quad (4.14)$$

we aim to estimate $|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap I|$ for each interval I in the previous decomposition of $[0, N)$. The interval $[0, g^m)$ can be estimated from Lemma 4.4 and $[N(t+1), N(t))$ from Lemma 4.5 as $N(t) = N(t+1) + d_t g^t$. However, the interval $[g^m, d_m g^m)$ can be handled in the same way as Lemma 4.5 by writing $[g^m, g^m + (d_m - 1)g^m)$. Putting these estimates together, we can obtain the upper bound

$$\frac{|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, N)|}{|A_{\Sigma, \ell} \cap [0, N)|} \leq \frac{1}{|\mathbf{a}|} + CR(m), \quad (4.15)$$

where

$$\begin{aligned} R(m) := & \rho^m \frac{|A_{\Sigma, \ell} \cap [0, g^m)|}{|A_{\Sigma, \ell} \cap [0, N)|} + \rho^m \frac{|A_{\Sigma, \ell} \cap [g^m, d_m g^m)|}{|A_{\Sigma, \ell} \cap [0, N)|} \\ & + \sum_{t=0}^{m-1} \rho^t \frac{|A_{\Sigma, \ell} \cap [N(t+1), N(t))|}{|A_{\Sigma, \ell} \cap [0, N)|}. \end{aligned}$$

If $A_{\Sigma, \ell} \cap [N(t+1), N(t)) \neq \emptyset$ (in particular, $d_t \neq 0$), each element of $A_{\Sigma, \ell} \cap [N(t+1), N(t))$ has a base- g expansion of length $m+1$, where the $(m-t)$ most significant digits are $d_{t+1} \cdots d_m$. Then, they may differ only in the $(t+1)$ least significant digits. Therefore,

$$|A_{\Sigma, \ell} \cap [N(t+1), N(t))| \leq |\mathcal{L}_{\mathcal{V}, \ell}^{t+1}(\Sigma)|.$$

Let $k \in \mathbb{N}$ be the regularity of Σ . Since $|A_{\Sigma,\ell} \cap [0, N)| \geq |\mathcal{L}_{\mathcal{V},\ell}^m(\Sigma)|$, we can get from the above estimation and Proposition 3.11 that

$$\frac{|A_{\Sigma,\ell} \cap [N(t+1), N(t))|}{|A_{\Sigma,\ell} \cap [0, N)|} \leq \frac{|\mathcal{L}_{\mathcal{V},\ell}^\ell(\Sigma)|k^{t+1-\ell}}{|\mathcal{L}_{\mathcal{V},\ell}^\ell(\Sigma)|k^{m-\ell}} = k^{t+1-m} \leq \rho^{m-t-1},$$

where we assumed that $k^{-1} \leq \rho$ (otherwise, replace ρ by k^{-1}). Therefore, using the previous estimation in (4.15),

$$\frac{|\{n \in A_{\Sigma,\ell} : \mathbf{f}(n) \equiv \mathbf{b} \bmod \mathbf{a}\} \cap [0, N)|}{|A_{\Sigma,\ell} \cap [0, N)|} - \frac{1}{|\mathbf{a}|} \leq C(2 + \rho^{-1}m)\rho^m.$$

The lower inequality can be obtained following the same argument. Considering $r = \rho^{1/2}$ and C' large enough, $C'r^m > C(2 + \rho^{-1}m)\rho^m$ for all $m \in \mathbb{N}$. The conclusion follows as $m \log(g) \leq \log(N) \leq (m+1) \log(g)$. \square

Remark 4.7. When Σ is a shift of finite type, $|A_\Sigma \setminus A_{\Sigma,\ell}|$ is finite for all ℓ large enough. In such a case, we can replace $A_{\Sigma,\ell}$ by A_Σ in Proposition 4.6. Moreover, while the choice of the initial distribution for the Markov chains is important to recover $A_{\Sigma,\ell}$ in Proposition 4.6, the convergence of the Markov chains remains independent of this choice. Therefore, different sets of integers can be considered by adjusting the initial distribution.

The next step is to transfer the distribution from $A_{\Sigma,\ell}$ to the multiplicatively invariant set A_Σ . To achieve this, it suffices to show that the difference between these two sets becomes asymptotically negligible in size. The main ingredient to conclude that is the following proposition.

PROPOSITION 4.8. *Let Σ be a transitive, sofic and regular subshift. Then,*

$$\lim_{\ell \rightarrow \infty} \frac{|\mathcal{L}^\ell(\Sigma) \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell(\Sigma)|}{|\mathcal{L}_{\mathcal{V},\ell}^\ell(\Sigma)|} = 0. \quad (4.16)$$

Proof. Let $k \geq 2$ be the regularity of Σ . For any $y \in \mathcal{L}(\Sigma)$ such that $F(y) \in \mathcal{V}$, every extension $\tilde{y} \in \mathcal{L}(\Sigma)$ of y satisfies $F(\tilde{y}) \in \mathcal{V}$. Then, for every $\ell \in \mathbb{N}$ and $i \geq \ell$, $\mathcal{L}_{\mathcal{V},\ell}^i \subseteq \mathcal{L}_{\mathcal{V},i}^i$. From Proposition 3.11, it follows that

$$|\mathcal{L}_{\mathcal{V},i}^i| \geq |\mathcal{L}_{\mathcal{V},\ell}^i| = |\mathcal{L}_{\mathcal{V},\ell}^\ell|k^{i-\ell}. \quad (4.17)$$

Since Σ is transitive, the Fischer cover is an irreducible graph, so we can find $q \in \mathbb{N}$ large enough such that, for every node $F \in \mathcal{V}$, there is a path π of length q starting at F , where $\lambda(\pi)$ is a synchronizing word. To see this, notice that every word can be extended to a synchronizing word and we can choose a q common for every node as any extension of a synchronizing word is synchronizing.

Let $n \in \mathbb{N}$ and $x \in \mathcal{L}^{\ell+nq} \setminus \mathcal{L}_{\mathcal{V},\ell+nq}^{\ell+nq}$. Then, x is in the language, but $F(x) \notin \mathcal{V}$. In particular, $F(x_0 \cdots x_{\ell-1}) \notin \mathcal{V}$ and the word $x_\ell \cdots x_{\ell+nq-1}$ is not synchronizing. Thus,

$$|\mathcal{L}^{\ell+nq} \setminus \mathcal{L}_{\mathcal{V},\ell+nq}^{\ell+nq}| \leq |\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell| \times |\{w \in \mathcal{L}^{nq}(\Sigma) : w \text{ is not synchronizing}\}|. \quad (4.18)$$

For each node $F \in \mathcal{V}$, there are k^{qn} paths of length qn in \mathcal{G} starting at F . From those paths, there are at most $(k^q - 1)^n$ paths labelled by non-synchronizing words, as for every node F , we can find at least one path starting at F of length q that represents a synchronizing word. From this idea and (4.18),

$$|\mathcal{L}^{\ell+nq} \setminus \mathcal{L}_{\mathcal{V}, \ell+nq}^{\ell+nq}| \leq |\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V}, \ell}^\ell| \times |\mathcal{V}|(k^q - 1)^n. \quad (4.19)$$

By defining

$$C := |\mathcal{V}| \max_{j=\ell, \dots, \ell+(q-1)} \frac{|\mathcal{L}^j \setminus \mathcal{L}_{\mathcal{V}, j}^j|}{|\mathcal{L}_{\mathcal{V}, j}^j|}$$

and combining (4.17) with (4.19),

$$\frac{|\mathcal{L}^i \setminus \mathcal{L}_{\mathcal{V}, i}^i|}{|\mathcal{L}_{\mathcal{V}, i}^i|} \leq C \frac{(k^q - 1)^{\lfloor i/q \rfloor}}{k^{q \lfloor i/q \rfloor}}.$$

The result follows by taking $i \rightarrow \infty$. \square

We can now state the main result of this section.

THEOREM 4.9. *Let $g \geq 2$, $\mathbf{a} = (a_1, \dots, a_r) \in \mathbb{N}^r$ and $\mathbf{f} = (f_1, \dots, f_r)$ be a collection of g -additive functions eventually periodic with respect to \mathbf{a} . Let $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ be a mixing, sofic and regular subshift, and p be an eventual period. If the Markov condition holds,*

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A_\Sigma : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [0, N]\}|}{|A_\Sigma \cap [0, N]|} = \frac{1}{a_1 \cdot a_2 \cdots a_r}$$

for every $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{Z}_{\mathbf{a}}$.

Proof. Let (p, ℓ) be an eventual period and $N \in \mathbb{N}$. For every $\mathbf{b} \in \mathbb{Z}_{\mathbf{a}}$, define

$$\begin{aligned} \alpha(\ell, N) &:= \frac{|(A_\Sigma \setminus A_{\Sigma, \ell}) \cap [0, N]|}{|A_{\Sigma, \ell} \cap [0, N]|}, \quad L(\mathbf{b}, \ell, N) \\ &:= \frac{|\{n \in A_{\Sigma, \ell} : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [0, N]\}|}{|A_{\Sigma, \ell} \cap [0, N]|}. \end{aligned}$$

A simple computation leads to the following inequality:

$$\frac{1}{1 + \alpha(\ell, N)} L(\mathbf{b}, \ell, N) \leq \frac{|\{n \in A_\Sigma : \mathbf{f}(n) \equiv \mathbf{b} \pmod{\mathbf{a}} \cap [0, N]\}|}{|A_\Sigma \cap [0, N]|} \leq L(\mathbf{b}, \ell, N) + \alpha(\ell, N).$$

Proposition 4.6 gives that $L(\mathbf{b}, \ell, N) \rightarrow 1/|\mathbf{a}|$ when $N \rightarrow \infty$. Therefore, to conclude the theorem, it will be enough to show that

$$\lim_{\ell \rightarrow \infty} \limsup_{N \rightarrow \infty} \alpha(\ell, N) = 0. \quad (4.20)$$

For this purpose, we will estimate $|(A_\Sigma \setminus A_{\Sigma, \ell}) \cap [0, N]|$. For $N \in \mathbb{N}$, let m be such that $g^m \leq N < g^{m+1}$. If $n \in (A_\Sigma \setminus A_{\Sigma, \ell}) \cap [g^t, g^{t+1})$ for some $\ell \leq t \leq m$, n can be represented by a word in $\mathcal{L}^{t+1}(\Sigma) \setminus \mathcal{L}_{\mathcal{V}, \ell}^{t+1}(\Sigma)$. Thus, we can estimate

$$\frac{|(A_\Sigma \setminus A_{\Sigma,\ell}) \cap [g^\ell, g^{m+1})|}{|A_{\Sigma,\ell} \cap [g^\ell, g^m)|} \leq \frac{\sum_{i=\ell}^{m+1} |\mathcal{L}^i \setminus \mathcal{L}_{\mathcal{V},\ell}^i|}{|\mathcal{L}_{\mathcal{V},\ell}^m|}. \quad (4.21)$$

It is easy to see that each word in $\mathcal{L}^i \setminus \mathcal{L}_{\mathcal{V},\ell}^i$ is an extension of a word in $\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell$. Let $w \in \mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell$. Despite $F(w) \notin \mathcal{V}$, $F(w)$ is union of elements in \mathcal{V} . To see this, just consider the union of all $F \in \mathcal{V}$ such that the word w can be read in some path of the Fischer cover ending at F . Since the subshift is k -regular for some $k \geq 2$, from the Fischer cover, we can deduce that there are at most $|\mathcal{V}|k^{i-\ell}$ words of length i that are extensions of w . In particular, each of those extensions is in $\mathcal{L}^i \setminus \mathcal{L}_{\mathcal{V},\ell}^i$. Therefore, each word in $\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell$ is extended to at most $|\mathcal{V}|k^{i-\ell}$ words of length i , concluding from (4.21) that

$$\frac{|(A_\Sigma \setminus A_{\Sigma,\ell}) \cap [g^\ell, g^{m+1})|}{|A_{\Sigma,\ell} \cap [g^\ell, g^m)|} \leq \frac{|\mathcal{V}| |\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell| \sum_{i=\ell}^{m+1} k^{i-\ell}}{|\mathcal{L}_{\mathcal{V},\ell}^\ell| k^{m-\ell}} \leq |\mathcal{V}| \sum_{i=-1}^{\infty} k^{-i} \frac{|\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell|}{|\mathcal{L}_{\mathcal{V},\ell}^\ell|}.$$

Therefore,

$$\begin{aligned} 0 \leq \limsup_{N \rightarrow \infty} \alpha(\ell, N) &\leq \limsup_{N \rightarrow \infty} \frac{g^\ell + |(A_\Sigma \setminus A_{\Sigma,\ell}) \cap [g^\ell, g^{m+1})|}{|A_{\Sigma,\ell} \cap [g^\ell, g^m)|} \\ &\leq \frac{|\mathcal{L}^\ell \setminus \mathcal{L}_{\mathcal{V},\ell}^\ell|}{|\mathcal{L}_{\mathcal{V},\ell}^\ell|} |\mathcal{V}| \sum_{i=-1}^{\infty} k^{-i}. \end{aligned}$$

It follows that $\lim_{\ell \rightarrow \infty} \limsup_{N \rightarrow \infty} \alpha(\ell, N) = 0$ from the previous inequality and Proposition 4.8, concluding the result. \square

Remark 4.10. Although we focus on subsets of integers, the nature of the construction allows some of the ideas presented in this paper to be immediately extended to more general contexts where elements can be uniquely represented on a certain basis (canonical numerical systems). Relevant references to these systems include [1, 34, 37]. A particular case are the Gaussian integers, where unique expansion of elements can be provided in certain bases (see [26, 28, 33]).

5. Applications to the distribution of multiplicatively invariant sets of integers

Let A be a multiplicatively invariant set of integers that admits a representation $A = A_\Sigma$, where Σ is a mixing, sofic and regular subshift. From Theorem 4.9, uniform distribution is obtained if the Markov chains are irreducible and aperiodic, which might be challenging to verify. To check the Markov condition, we can conveniently choose the period p and study the matrices M_i using the characterization provided by Proposition 3.13. For any $n \in \mathbb{N}$ and $s, s' \in \mathcal{S}$,

$$M_i^n(s, s') = k^{-pn} |\{w \in \mathcal{L}^{pn}(\Sigma) : \mathbf{f}(g^i(w)_g) \equiv \mathbf{b}' - \mathbf{b} \bmod \mathbf{a}, F \xrightarrow{w} F'\}|.$$

For M_i to be irreducible, we need to check that for all $\mathbf{b} \in \mathbb{Z}_\mathbf{a}$ and $F, F' \in \mathcal{V}$,

$$\text{there exists } w \in \mathcal{L}(\Sigma) \text{ such that } p \mid |w|, \mathbf{f}(g^i(w)_g) \equiv \mathbf{b} \bmod \mathbf{a} \text{ and } F \xrightarrow{w} F'. \quad (5.1)$$

Assuming that M_i is irreducible, aperiodicity holds if there exists $F \in \mathcal{V}$ such that

$$\gcd\{n \in \mathbb{N} : \text{there exists } w \in \mathcal{L}^{pn}(\Sigma) \text{ such that } \mathbf{f}(g^i(w)_g) \equiv 0 \pmod{a} \text{ and } F \xrightarrow{w} F\} = 1. \quad (5.2)$$

In this section, we will study the function $\mathbf{f} = (\text{id}, S_g)$, where S_g denotes the function sum of digits in base g .

5.1. Uniform distribution of integers with missing digits. Let $g \geq 2$ be an integer and let $\mathcal{D} \subseteq \{0, \dots, g-1\}$ be a subset of digits with at least two elements. We will write $\mathcal{D} = \{d_1, \dots, d_t\}$, where $d_1 < d_2 < \dots < d_t$. Recalling that $C_{g,\mathcal{D}} = A_\Sigma$ for $\Sigma = \mathcal{D}^{\mathbb{N}_0}$, it is easy to see that the Fischer cover for Σ is the labelled graph with a unique node F and for each $d \in \mathcal{D}$, there is a path from F to F labelled by d . Thus, Σ is mixing, sofic and regular, so we can apply Theorem 4.9. Moreover, the existence of a single follower set allows us to obviate this notion in the following discussion.

Definition 5.1. Let $A \subseteq \mathbb{N}_0$ be a $\times g$ -invariant set and $a, a' \in \mathbb{N}$. We say A is *jointly uniformly distributed* mod (a, a') if

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, N]|}{|A \cap [0, N]|} = \frac{1}{aa'}$$

for every $b \in \mathbb{Z}_a, b' \in \mathbb{Z}_{a'}$.

From Proposition 3.2, we can find an eventual period of the form (p, ℓ) for $\mathbf{f} = (\text{id}, S_g)$. Thus, checking that M_i is irreducible reduces to show that for all $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$,

$$\text{there is } n \in \mathbb{N} \text{ and } w \in \mathcal{D}^{np} \text{ such that } g^i(w)_g \equiv b \pmod{a} \text{ and } S_g(w) \equiv b' \pmod{a'}. \quad (5.3)$$

Similarly, M_i is aperiodic if

$$\gcd\{n \in \mathbb{N} : \text{there exists } w \in \mathcal{D}^{pn} \text{ such that } (w)_g \equiv 0 \pmod{a} \text{ and } S_g(w) \equiv 0 \pmod{a'}\} = 1. \quad (5.4)$$

Aperiodicity in the case $\gcd(g, a) = 1$ follows immediately from the following result.

LEMMA 5.2. Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$ such that $\gcd(g, a) = 1$. Let $p := a'\phi(a(g-1))$, where ϕ is the Euler's totient function. Then, for every $n \in \mathbb{N}$ and $w = d \cdots d \in \mathcal{D}^{np}$, $(w)_g \equiv 0 \pmod{a}$ and $S_g(w) \equiv 0 \pmod{a'}$.

Proof. Let $n \in \mathbb{N}$. Notice that $\gcd(g, a(g-1)) = 1$, so the Euler's theorem implies that $g^p \equiv 1 \pmod{a(g-1)}$. If $w = d \cdots d \in \mathcal{D}^{np}$,

$$(w)_g = d + dg + \dots + dg^{np-1} = d \frac{g^{np} - 1}{g - 1}. \quad (5.5)$$

Since $g^{np} - 1$ is divisible by $a(g-1)$, $(g^{np} - 1)/(g-1) \equiv 0 \pmod{a}$, concluding that $(w)_g \equiv 0 \pmod{a}$. However, $S_g(w) = npd = a'(n\phi(a(g-1))d) \equiv 0 \pmod{a'}$. \square

The following application of Bézout's identity is fundamental to check the irreducibility of M_i .

LEMMA 5.3. Let $g \geq 2$ be an integer and let $\mathcal{D} = \{d_1, \dots, d_t\}$ be a set of digits, where d_1 is the smallest digit. Consider $a, a' \in \mathbb{N}$ such that $\gcd(g, a) = 1$ and define $p := a'\phi(a(g-1))$, where ϕ is the Euler's totient function. Then, there exists a word $w \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ such that

$$|w| \equiv 0 \pmod{p}, \quad (w)_g \equiv \delta \pmod{a} \quad \text{and} \quad S_g(w) \equiv \delta \pmod{a'},$$

where $\delta := \gcd(aa', d_2 - d_1, \dots, d_t - d_1)$.

Proof. From Bézout's identity, there exist $n_1, n_2, \dots, n_t \in \mathbb{Z}$ such that

$$\delta = n_1(aa') + n_2(d_2 - d_1) + \dots + n_t(d_t - d_1).$$

For all $j \in \{2, \dots, t\}$, choose $k_j \in \mathbb{N}$ large enough such that $m_j := k_j(aa') + n_j > 0$. Then,

$$m_2(d_2 - d_1) + \dots + m_t(d_t - d_1) \equiv \delta \pmod{aa'}. \quad (5.6)$$

Let us define the word $w(j) := d_j d_1 \dots d_1 \in \mathcal{D}^p$ for $j \in \{2, \dots, t\}$. From Lemma 5.2,

$$(w(j))_g = (d_j - d_1) + d_1 + d_1 g + \dots + d_1 g^{p-1} \equiv d_j - d_1 \pmod{a}.$$

Similarly, $S_g(w(j)) = p d_1 + (d_j - d_1) \equiv d_j - d_1 \pmod{a'}$. Let us define

$$w := \underbrace{w(2) \dots w(2)}_{m_2 \text{ times}} \dots \underbrace{w(t) \dots w(t)}_{m_t \text{ times}}.$$

Notice that $|w| = p(m_2 + \dots + m_t)$. Also, the periodicity $g^p \equiv 1 \pmod{a}$ and (5.6) imply that

$$(w)_g \equiv \sum_{j=2}^t (w(j))_g m_j \equiv \sum_{j=2}^t (d_j - d_1) m_j \equiv \delta \pmod{a}.$$

Using (5.6) again, we can conclude

$$S_g(w) = \sum_{j=2}^t S_g(w(j)) m_j \equiv \sum_{j=2}^t (d_j - d_1) m_j \equiv \delta \pmod{a'}. \quad \square$$

Leveraging Theorem 4.9 and Lemma 5.3, we can establish a complete description of the phenomenon of joint distribution (Definition 5.1) for missing digits sets when $\gcd(g, a) = \gcd(a, a') = 1$. In particular, this shows Theorem A.

THEOREM 5.4. Let $g \geq 2$ be an integer and let $\mathcal{D} = \{d_1, \dots, d_t\}$ be a set of digits, where d_1 is the smallest digit. Suppose that $\gcd(g, a) = \gcd(a, a') = 1$ and recall the missing digits set $C_{g, \mathcal{D}}$ defined in (1.1). Defining $\delta := \gcd(aa', d_2 - d_1, \dots, d_t - d_1)$, we get the following.

(i) $C_{g, \mathcal{D}}$ is jointly uniformly distributed mod (a, a') if and only if $\delta = 1$.

More generally, concerning the limit

$$L(b, b') := \lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, N]|}{|C_{g, \mathcal{D}} \cap [0, N]|}, \quad (5.7)$$

we have the following.

- (ii) Let $\delta_a := \gcd(\delta, a)$ and $\delta_{a'} := \gcd(\delta, a')$. If $d_1 \equiv 0 \pmod{\delta_a}$ and $d_1 \equiv 0 \pmod{\delta_{a'}}$, then $L(b, b') = \delta/aa'$ for every (b, b') in the subgroup $\langle \delta_a \rangle \times \langle \delta_{a'} \rangle \subseteq \mathbb{Z}_a \times \mathbb{Z}_{a'}$. However, if $(b, b') \notin \langle \delta_a \rangle \times \langle \delta_{a'} \rangle$,

$$\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} = \emptyset.$$

- (iii) In any other case, for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$, the limit $L(b, b')$ either equals zero or does not exist.

Proof. Let us define $p = a'\phi(a(g-1))$, where ϕ is the Euler's totient function, and note that $(p, 1)$ is an eventual period for $\mathbf{f} = (\text{id}, S_g)$ with respect to $\mathbf{a} = (a, a')$. For $i \in \mathbb{N}$, let X^i be the Markov chain associated to $g, \mathbf{a}, \mathbf{f}, p, \ell$ and the subshift $\Sigma = \mathcal{D}^{\mathbb{N}_0}$, as defined in Definition 3.14. We denote by M_i its transition matrix and μ_i its initial distribution.

The aperiodicity of M_i follows from Lemma 5.2 and (5.4).

Let us suppose $\delta = 1$. From Lemma 5.3, there exists a word w using symbols of \mathcal{D} such that $|w| \equiv 0 \pmod{p}$, $(w)_g \equiv 1 \pmod{a}$ and $S_g(w) \equiv 1 \pmod{a'}$. Since $\gcd(a, a') = 1$, the element $(1, 1)$ is a generator for $\mathbb{Z}_a \times \mathbb{Z}_{a'}$. Then, for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$, there exists $m \in \mathbb{N}$ such that $m \equiv (g^i)^{-1}b \pmod{a}$ and $m \equiv b' \pmod{a'}$. Let $y \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ be constructed by concatenating w with itself m times. The periodicity $g^p \equiv 1 \pmod{a}$ implies that $g^i(y)_g \equiv g^i m(w)_g \equiv b \pmod{a}$. However, $S_g(y) \equiv m S_g(w) \equiv b' \pmod{a'}$. Therefore, M_i is irreducible for every $i \in \mathbb{N}$. This concludes that the Markov condition holds, then Theorem 4.9 implies that $C_{g, \mathcal{D}}$ is jointly uniformly distributed $\pmod{(a, a')}$.

Now, let us suppose that $\delta \neq 1$. Since δ divides aa' and $\gcd(a, a') = 1$, we can write $\delta = \delta_a \cdot \delta_{a'}$, where $\delta_a := \gcd(\delta, a)$ and $\delta_{a'} := \gcd(\delta, a')$. In particular, $\gcd(\delta_a, \delta_{a'}) = 1$.

From the definition of δ , note that $d_j \equiv d_1 \pmod{\delta}$ for all $j \in \{1, \dots, t\}$. Therefore, for any $m \in \mathbb{N}$ and $x = x_0 \cdots x_{m-1} \in \mathcal{D}^m$,

$$(x)_g = \sum_{j=0}^{m-1} (x_j - d_1)g^j + d_1 g^j \equiv d_1 \sum_{j=0}^{m-1} g^j \pmod{\delta_a} \quad \text{and} \quad S_g(x) \equiv m d_1 \pmod{\delta_{a'}}. \quad (5.8)$$

In particular, for every $x \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ such that $|x| \equiv 0 \pmod{p}$, Lemma 5.2 implies that $(x)_g \equiv 0 \pmod{\delta_a}$ and $S_g(x) \equiv 0 \pmod{\delta_{a'}}$. However, from Lemma 5.3, we can find a word y such that $|y| \equiv 0 \pmod{p}$, $(y)_g \equiv \delta \pmod{a}$ and $S_g(y) \equiv \delta \pmod{a'}$. By the Chinese remainder theorem, (δ, δ) is a generator for the subgroup $\mathcal{S}(\delta_a, \delta_{a'}) := \langle \delta_a \rangle \times \langle \delta_{a'} \rangle \subseteq \mathbb{Z}_a \times \mathbb{Z}_{a'}$. Therefore, recalling that $\gcd(g, \delta_a) = 1$,

$$\bigcup_{n \in \mathbb{N}} \{w \in \mathcal{D}^{np} : g^i(w)_g \equiv b \pmod{a}, S_g(w) \equiv b' \pmod{a'}\} \neq \emptyset \quad (5.9)$$

if and only if $(b, b') \in \mathcal{S}(\delta_a, \delta_{a'})$. In particular, this already ensures that M_i is not irreducible on $\mathbb{Z}_a \times \mathbb{Z}_{a'}$, so the Markov condition does not hold. However, we can show that X^i is an irreducible and aperiodic Markov chain when restricting the state space. Recall the initial distribution of X^i given by

$$\mu_i(b, b') = \frac{|\{w \in \mathcal{D}^i : (w)_g \equiv b \pmod{a}, S_g(w) \equiv b' \pmod{a'}\}|}{|\mathcal{D}|^i}.$$

Defining $\delta_a(i) := d_1 \sum_{j=0}^{i-1} g^{j-1} \pmod{\delta_a}$ and $\delta_{a'}(i) := id_1 \pmod{\delta_{a'}}$, (5.8) implies that $(x)_g \equiv \delta_a(i) \pmod{\delta_a}$ and $S_g(x) \equiv \delta_{a'}(i) \pmod{\delta_{a'}}$ for every $x \in \mathcal{D}^i$. Therefore, $\mu_i(b, b') \neq 0$ implies $b \equiv \delta_a(i) \pmod{\delta_a}$ and $b' \equiv \delta_{a'}(i) \pmod{\delta_{a'}}$. If we define

$$\mathcal{S}(i) := (\delta_a(i), \delta_{a'}(i)) + \mathcal{S}(\delta_a, \delta_{a'}),$$

the equivalence in (5.9) indicates that X^i never visits the states in $\mathbb{Z}_a \times \mathbb{Z}_{a'} \setminus \mathcal{S}(i)$, and also X^i is irreducible and aperiodic when restricted to the state space $\mathcal{S}(i)$. Note that M^i restricted to $\mathcal{S}(i)$ is still doubly stochastic, since we only ignore zero-value entries. Therefore, as $|\mathcal{S}(i)| = |\mathcal{S}(\delta_a, \delta_{a'})| = (a/\delta_a) \cdot (a'/\delta_{a'})$,

$$\lim_{N \rightarrow \infty} \mu_{i+Np}(b, b') = \begin{cases} \delta/aa' & \text{if } (b, b') \in \mathcal{S}(i), \\ 0 & \text{if } (b, b') \notin \mathcal{S}(i). \end{cases} \quad (5.10)$$

If $d_1 \equiv 0 \pmod{\delta_a}$ and $d_1 \equiv 0 \pmod{\delta_{a'}}$, $\mathcal{S}(i) = \mathcal{S}(\delta_a, \delta_{a'})$ for every $i \in \mathbb{N}$. This shows that each Markov chain X^i is irreducible and aperiodic in a common state space $\mathcal{S}(\delta_a, \delta_{a'})$. Note that the Markov condition means that every X^i is irreducible and aperiodic over the state space $\mathbb{Z}_a \times \mathbb{Z}_{a'}$, but it is easy to see that the procedure of §4 is identical if we assume that each X^i is irreducible and aperiodic over a subgroup of $\mathbb{Z}_a \times \mathbb{Z}_{a'}$ (and the Markov chain never visits the other states). Thus, by replicating the argument of §4, we can conclude

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv a b, S_g(n) \equiv_{a'} b'\} \cap [0, N)|}{|C_{g, \mathcal{D}} \cap [0, N)|} = \begin{cases} \delta/aa' & \text{if } (b, b') \in \mathcal{S}(\delta_a, \delta_{a'}), \\ 0 & \text{if } (b, b') \notin \mathcal{S}(\delta_a, \delta_{a'}). \end{cases}$$

This shows (5.4). If either $d_1 \not\equiv 0 \pmod{\delta_a}$ or $d_1 \not\equiv 0 \pmod{\delta_{a'}}$, notice that $\mathcal{S}(1) \cap \mathcal{S}(2) = \emptyset$. Then, for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$, there exists $j \in \mathbb{N}$ such that the Markov chain X^j does not visit the state (b, b') . From this fact, we want to show that (5.7) either equals zero or the limit does not exist. Let $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$ and let $j \in \mathbb{N}$ such that X^j never visits the state (b, b') . By contradiction, suppose that $L(b, b') > 0$. Since X^j does not visit the state (b, b') ,

$$\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [g^{j+mp-1}, g^{j+mp}) = \emptyset$$

for every $m \in \mathbb{N}$. The contradiction follows by noting that

$$\begin{aligned} L(b, b') &= \lim_{m \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, g^{j+mp-1})|}{|C_{g, \mathcal{D}} \cap [0, g^{j+mp-1})|} \\ &> \liminf_{m \rightarrow \infty} \frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, g^{j+mp})|}{|C_{g, \mathcal{D}} \cap [0, g^{j+mp})|}. \end{aligned}$$

This concludes case (5.4). □

Note that in the cases (5.4) and (5.4), we can conclude from Proposition 4.6 instead of using Theorem 4.9, as Σ is a full shift. Hence, the order of convergence $O(N^{-\gamma})$ for some $\gamma > 0$ can be provided for the limit (5.7). An immediate consequence follows by taking

$a' = 1$, which allows us to fully describe the phenomenon of uniform distribution mod a when $\gcd(g, a) = 1$. This shows Corollary A.

COROLLARY 5.5. *Let $g \geq 2$ be an integer and let $\mathcal{D} = \{d_1, \dots, d_t\}$ be a set of digits, where d_1 is the smallest digit. Suppose $\gcd(g, a) = 1$ and denote $\delta := \gcd(a, d_2 - d_1, \dots, d_t - d_1)$. Then:*

- (i) $C_{g, \mathcal{D}}$ is uniformly distributed mod a if and only if $\delta = 1$;
- (ii) if $\delta \neq 1$ and $d_1 \equiv 0 \pmod{\delta}$, $C_{g, \mathcal{D}}$ is uniformly distributed in the subgroup $\langle \delta \rangle \subseteq \mathbb{Z}_a$;
- (iii) in any other case, the limit of

$$\frac{|\{n \in C_{g, \mathcal{D}} : n \equiv b \pmod{a}\} \cap [0, N)|}{|C_{g, \mathcal{D}} \cap [0, N)|}$$

when $N \rightarrow \infty$ is either zero or does not exist.

Note that the condition $\gcd(g, a) = 1$ is commonly assumed in results of this kind. In the framework of Corollary 5.5, when $\gcd(g, a) \neq 1$, the matrices M_i are no longer irreducible over \mathbb{Z}_a . However, the Markov chains constructed can still capture the behaviour of the distribution by breaking X^i into pieces determined by the communication classes of the Markov chain (see [44, §1.2] for details about this notion). The disadvantage of this situation is that the convergence of X^i depends on the initial distribution.

The following simple example shows that when $\gcd(a, g) \neq 1$, the limiting behaviour depends heavily on the initial distribution.

Example 5.6. Let $a = 12$, $g = 6$ and $\mathcal{D} = \{1, 2, 4\}$. Notice that a divides g^i for every $i \geq 2$, then $(w_0 \cdots w_n)_g \equiv w_0 + w_1 g \pmod{a}$. Therefore,

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{6, \mathcal{D}} : n \equiv b \pmod{12}\} \cap [0, N)|}{|C_{6, \mathcal{D}} \cap [0, N)|} = \begin{cases} 2/9 & \text{if } b \in \{1, 2, 4\}, \\ 1/9 & \text{if } b \in \{7, 8, 10\}, \\ 0 & \text{if } b \in \{0, 3, 5, 6, 9, 11\}. \end{cases}$$

Notice that this distribution agrees with the distribution when considering $[0, g^2)$. Also, we can see that the limit is different and non-zero for some classes, behaviour which is not observed under the assumption $\gcd(g, a) = 1$.

Finally, we will present an equivalent condition for uniform distribution mod (a, a') without the assumption $\gcd(a, a') = 1$, however, at the cost of having a more complicated condition to check.

PROPOSITION 5.7. *Let $g \geq 2$ be an integer and let $\mathcal{D} = \{d_1, \dots, d_t\}$ be a set of digits, where d_1 is the smallest digit. Consider $a, a' \in \mathbb{N}$ such that $\gcd(g, a) = 1$ and define $p := a' \phi(a(g-1))$. Therefore, $C_{g, \mathcal{D}}$ is jointly uniformly distributed mod (a, a') if and only if $\gcd(a, d_2 - d_1, \dots, d_t - d_1) = 1$, and there exists $w \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ such that $(w)_g \equiv 0 \pmod{a}$, $S_g(w) \equiv 1 \pmod{a'}$ and $|w| \equiv 0 \pmod{p}$.*

Proof. If $C_{g, \mathcal{D}}$ is jointly uniformly distributed mod (a, a') , in particular, it is uniformly distributed mod a . From Corollary 5.5, it follows that $\gcd(a, d_2 - d_1, \dots, d_t - d_1) = 1$.

In addition, there exist infinitely many words w such that $(w)_g \equiv 0 \pmod{a}$, $S_g((w)_g) \equiv 1 \pmod{a'}$ and $|w| \equiv 0 \pmod{p}$. Otherwise, for every $m \in \mathbb{N}$,

$$\{n \in C_{g,\mathcal{D}} : n \equiv 0 \pmod{a}, S_g(n) \equiv 1 \pmod{1}\} \cap [g^{mp-1}, g^{mp}) = \emptyset,$$

contradicting that

$$\lim_{N \rightarrow \infty} \frac{|\{n \in C_{g,\mathcal{D}} : n \equiv 0 \pmod{a}, S_g(n) \equiv 1 \pmod{a'}\} \cap [0, N)|}{|C_{g,\mathcal{D}} \cap [0, N)|} = \frac{1}{aa'} > 0.$$

However, let us suppose that $\gcd(a, d_2 - d_1, \dots, d_t - d_1) = 1$ and that there exists $w \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ such that $(w)_g \equiv 0 \pmod{a}$, $S_g(w) \equiv 1 \pmod{a'}$ and $|w| \equiv 0 \pmod{p}$. From Lemma 5.3, there exists $y \in \mathcal{L}(\mathcal{D}^{\mathbb{N}_0})$ such that $(y)_g \equiv 1 \pmod{a}$ and $|y| \equiv 0 \pmod{p}$. To prove that M_i is irreducible, we will check (5.3). Let $b \in \mathbb{Z}_a$ and $b' \in \mathbb{Z}_{a'}$. If \tilde{y} be the concatenation $((g^i)^{-1}b \pmod{a})$ -times of the word y (in the case where $b = 0$, we concatenate a times), the periodicity $g^p \equiv 1 \pmod{a}$ implies that $g^i(\tilde{y})_g \equiv g^i(g^i)^{-1}b \equiv b \pmod{a}$. If $s := S_g(\tilde{y})$, we denote by \tilde{w} the concatenation $(b' - s \pmod{a'})$ -times of w . Thus, we can conclude that $g^i(\tilde{y}\tilde{w})_g \equiv b \pmod{a}$ and $S_g(\tilde{y}\tilde{w}) \equiv b' \pmod{a'}$. Clearly, $|\tilde{y}\tilde{w}| \equiv 0 \pmod{p}$. Therefore, M_i is irreducible. The matrix M_i is also aperiodic as a consequence of Lemma 5.2. Thus, the Markov condition holds and the result follows by Theorem 4.9. \square

5.2. Revisiting Gelfond's theorem: an equivalent condition. Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$. As presented in Corollary 1.2, Gelfond [25] studied the long-term behaviour of

$$|\{n \leq N : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\}|, \quad (5.11)$$

showing that, under the assumption $\gcd(g-1, a') = 1$, the set $A = \mathbb{N}$ is jointly uniformly distributed $\pmod{(a, a')}$. In this section, we present an extension of Corollary 1.2, obtaining an equivalent condition to uniform distribution $\pmod{(a, a')}$.

Notice that the obstruction in applying our method arises from the lack of the condition $\gcd(g, a) = 1$, which implies that the Markov chains associated with this problem are not irreducible. However, we can still break each Markov chain X^i into pieces, generating irreducible and aperiodic Markov chains. In that way, the behaviour of (5.11) can be understood using our method.

THEOREM 5.8. *Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$. Then,*

$$\lim_{N \rightarrow \infty} \frac{|\{n \leq N : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\}|}{N} = \frac{1}{aa'} \quad (5.12)$$

for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$ if and only if there exists $n \in \mathbb{N}$ such that $n \equiv 0 \pmod{a}$ and $S_g(n) \equiv 1 \pmod{a'}$.

Proof. If (5.12) holds, there are infinitely many $n \in \mathbb{N}$ such that $n \equiv 0 \pmod{a}$ and $S_g(n) \equiv 1 \pmod{a'}$.

For the other direction, let $n \in \mathbb{N}$ be such that $n \equiv 0 \pmod{a}$ and $S_g(n) \equiv 1 \pmod{a'}$. For $\mathbf{f} = (\text{id}, S_g)$, let $p \in \mathbb{N}$ be given by Proposition 3.2. Then, $(p, a-1)$ is an eventual period for \mathbf{f} with respect to $(\mathbf{a}, \mathbf{a}')$ and notice that $\Sigma = \{0, \dots, g-1\}^{\mathbb{N}_0}$ is such that $\mathbb{N} = A_\Sigma$.

For $i \geq a - 1$, let X^i be the Markov chain of Definition 3.14, with transition matrix M_i and initial distribution μ_i .

Let $w = 10 \cdots 0$ such that $|w| = p$ and clearly $g^i(w)_g \equiv g^i \pmod{a}$. If y is the base- g expansion of n (by adding 0s, we can suppose that y has length divisible by p), then $(y)_g \equiv 0 \pmod{a}$ and $S_g(y) \equiv 1 \pmod{a'}$. Hence, for any $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$, there exists $k \in \mathbb{N}$ such that $(w^b y^k)_g \equiv b \pmod{a}$ and $S_g(w^b y^k)_g \equiv b' \pmod{a'}$ (where z^j denotes the concatenation of z j -times). Therefore,

$$\bigcup_{m \in \mathbb{N}} \{w \in \mathcal{D}^{mp} : g^i(w)_g \equiv b \pmod{a}, S_g(w) \equiv b' \pmod{a'}\} \neq \emptyset \quad (5.13)$$

if and only if $(b, b') \in \langle g^i \rangle \times \mathbb{Z}_{a'} \subseteq \mathbb{Z}_a \times \mathbb{Z}_{a'}$. Denoting $\delta_i := \gcd(g^i, a)$, it follows that $\sum_{m \in \mathbb{N}} M_i^m((b_0, b_1), (b'_0, b'_1)) > 0$ if and only if $b_0 \equiv b'_0 \pmod{\delta_i}$. Thus, if the Markov chain X^i starts at $X_0^i = (b, b')$, that trajectory will only visit classes in $(b + \langle \delta_i \rangle) \times \mathbb{Z}_{a'} \subseteq \mathbb{Z}_a \times \mathbb{Z}_{a'}$. In this way, we can split X^i into pieces depending on the starting point.

For each $j \in \{0, \dots, \delta_i - 1\}$, we define $\mathcal{S}(i, j) := j + \langle \delta_i \rangle$ and

$$\mu_{i,j}(b, b') := \begin{cases} \frac{\mu_i(b, b')}{\bar{\mu}_{i,j}} & \text{if } b \equiv j \pmod{\delta_i}, \\ 0 & \text{otherwise,} \end{cases}$$

where μ_i is the initial distribution of X^i and

$$\bar{\mu}_{i,j} := \sum_{b' \in \mathbb{Z}_{a'}} \sum_{b \in \mathcal{S}(j)} \mu_i(b, b') = \frac{|w \in \{0, \dots, g-1\}^i : (w)_g \equiv j \pmod{\delta_i}|}{g^i} = \frac{1}{\delta_i}.$$

Let $X^{i,j}$ be a Markov chain on the state space $\mathcal{S}(i, j)$, with initial distribution $\mu_{i,j}$ and transition matrix M_i , both restricted to indexes of $\mathcal{S}(i, j)$. The aperiodicity follows directly as $0 \cdots 0$ is in the language of $\{0, \dots, g-1\}^{\mathbb{N}_0}$, and M_i restricted to $\mathcal{S}(i, j)$ is irreducible by (5.13). Therefore, the Markov chain $X^{i,j}$ converges to its invariant distribution given by $1/|\mathcal{S}(i, j)| = \delta_i/aa'$. It is not difficult to see that

$$\mu_{i+np} = \sum_{j=0}^{\delta_i-1} \bar{\mu}_{i,j}(\mu_{i,j} M_i^n) = \frac{1}{\delta_i} \sum_{j=0}^{\delta_i-1} \mu_{i,j} M_i^n. \quad (5.14)$$

From Theorem 2.1, there exist $C_{i,j} > 0$ and $\rho_{i,j} \in (0, 1)$ such that

$$\left| (\mu_{i,j} M_i^n)(b, b') - \frac{\delta_i}{aa'} \right| \leq C_{i,j} \rho_{i,j}^n \quad (5.15)$$

for every $(b, b') \in \mathcal{S}(i, j)$. Clearly, $(\mu_{i,j} M_i^n)(b, b') = 0$ if $(b, b') \notin \mathcal{S}(i, j)$. Since $\mathbb{Z}_a \times \mathbb{Z}_{a'} = \bigsqcup_{j \in \{0, \dots, \delta_i-1\}} \mathcal{S}(i, j)$, from (5.14) and (5.15), we can obtain that, for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$,

$$\left| \mu_{i+np}(b, b') - \frac{1}{aa'} \right| \leq C_i \rho_i^n, \quad (5.16)$$

where $C_i := \max_{j=0, \dots, \delta_i-1} C_{i,j}$ and $\rho_i := \max_{j=0, \dots, \delta_i-1} \rho_{i,j}$. Equivalently,

$$\left| \mathbb{P}(X_n^i = (b, b')) - \frac{1}{aa'} \right| \leq C_i \rho_i^n. \quad (5.17)$$

Despite each X^i being not irreducible, we still have convergence to the uniform measure $1/aa'$. Note that (5.17) is equivalent to the estimation (4.3) in the case of Σ being a full shift, and essentially, this is what is required to retrieve the results of §4 and, in particular, Theorem 4.9. Therefore, we can conclude that \mathbb{N} is jointly uniformly distributed mod (a, a') by replicating the arguments of §4. \square

Remark 5.9. The obstruction when trying to use the previous idea in the context of missing digits sets lies in the fact that for some $i \in \mathbb{N}$, the normalization factor $\bar{\mu}_{i,j}$ could be not the same for each Markov chain $X^{i,j}$.

5.3. Forbidding combinations of digits. So far, we have focused on applying Theorem 4.9 in restricted digit sets, where restrictions come from forbidding some digits (missing digits sets). Given the flexibility in choosing digits and restrictions in the general case, it does not seem possible to establish general conditions like those in §5.1 within the context of general $\times g$ -invariant sets. In such a case, Theorem 4.9 is still a useful tool on a case-by-case basis. Seeking greater generality, we construct a class of sets with more flexible digit restrictions that exhibit joint uniform distribution. While we focus on restrictions of length 2, some of these ideas can be extended to longer lengths.

Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$. Let $\mathcal{D} \subseteq \{0, \dots, g-1\}$ be a set of digits and let T be a matrix on \mathcal{D} such that $T(d, d') \in \{0, 1\}$ for every $d, d' \in \mathcal{D}$.

Definition 5.10. We say that the matrix T is k -regular if for every $d \in \mathcal{D}$,

$$\sum_{d' \in \mathcal{D}} M(d, d') = \sum_{d' \in \mathcal{D}} M(d', d) = k.$$

Also, we say that T is *irreducible* if for every $d, d' \in \mathcal{D}$, there exists $n \in \mathbb{N}$ such that $M^n(d, d') > 0$.

As pointed out in §2.2, such a matrix T induces a shift of finite type

$$\Sigma_T := \{x \in \mathcal{D}^{\mathbb{N}_0} : T(x_i, x_{i+1}) = 1 \text{ for all } i \in \mathbb{N}_0\}.$$

In fact, any 1-step subshift of finite type can be written in this form. Similarly, we can define the $\times g$ -invariant set of integers generated by the subshift Σ_T as

$$A_T := A_{\Sigma_T} = \left\{ \sum_{i=0}^n d_i g^i : T(d_i, d_{i+1}) = 1 \text{ for all } i \in \mathbb{N}_0 \right\}.$$

As discussed in Remark 3.15, for a 1-step SFT, we can apply Theorem 4.9 when the Markov chains X^i (Definition 3.14) are constructed using the cover of Remark 2.7 instead of the Fischer cover. It is also straightforward that the graph of Remark 2.7 is regular and irreducible if the matrix T is regular and irreducible.

THEOREM 5.11. *Let $g \geq 2$ be an integer and $a, a' \in \mathbb{N}$ such that $\gcd(g, a) = \gcd(a, a') = 1$. Let $\mathcal{D} \subseteq \{0, \dots, g-1\}$ be a set of digits and T be a matrix over \mathcal{D} to values in $\{0, 1\}$, and assume that T is k -regular (for $k \geq 2$) and irreducible. If there exist $d, d' \in \mathcal{D}$ such that $T(d, d) = T(d, d') = T(d', d) = 1$ and $\gcd(d' - d, aa') = 1$, it follows that*

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A_T : n \equiv b \pmod{a}, S_g(n) \equiv b' \pmod{a'}\} \cap [0, N]|}{|A_T \cap [0, N]|}$$

for every $b \in \mathbb{Z}_a$ and $b' \in \mathbb{Z}_{a'}$.

Proof. We will check the conditions to apply Theorem 4.9. Denote by $\Sigma \subseteq \mathcal{D}^{\mathbb{N}_0}$ the subshift of finite type associated to the matrix T . Since T is regular and irreducible, the cover \mathcal{G}_T constructed in Remark 2.7 is regular and irreducible. Moreover, \mathcal{G}_T is aperiodic, as there exist $d \in \mathcal{D}$ such that $T(d, d)$ (that is, there is a cycle of length 1 in \mathcal{G}_T). Let $p = a'\phi(a(g-1))$ and notice that $(p, 1)$ is an eventual period for $\mathbf{f} = (\text{id}, S_g)$ with respect to (a, a') . By considering the Markov chains X^i of Definition 3.14 (with the consideration of Remark 3.15), the transition matrix M_i of X^i has state space $\mathcal{S} = (\mathbb{Z}_a \times \mathbb{Z}_{a'}) \times \mathcal{V}$, where $\mathcal{V} = \{F(\tilde{d}) : \tilde{d} \in \mathcal{D}\}$.

Since $T(d, d) = 1$, the word $x = d \cdots d \in \mathcal{D}^p$ satisfies $(w)_g \equiv 0 \pmod{a}$ and $S_g(w) \equiv 0 \pmod{a'}$ (Lemma 5.2). In addition, $F(d) \xrightarrow{x} F(d)$ as the last symbol of x is d . Thus, the state $(0, 0, F(d))$ is aperiodic.

Let $(b_0, b'_0, F_0), (b_1, b'_1, F_1) \in \mathcal{S}$. To show that M_i is irreducible for every i , we have to find $n(i) \in \mathbb{N}$ such that

$$M_i^{n(i)}((b_0, b'_0, F_0), (b_1, b'_1, F_1)) > 0.$$

Let $d_0, d_1 \in \mathcal{D}$ such that $F_0 = F(d_0)$ and $F_1 = F(d_1)$. Since the cover \mathcal{G}_T is aperiodic, there exists a path (of length divisible by p) from F_0 to $F(d)$ labelled by some word y . Similarly, there exists a path labelled by z (of length divisible by p) from $F(d)$ to F_1 .

Since $T(d, d) = T(d, d') = T(d', d) = 1$, the word $w = d'd \cdots d \in \mathcal{D}^p$ can be read in a path from $F(d)$ to $F(d)$, $(w)_g \equiv d' - d \pmod{a}$ and $S_g(w) \equiv d' - d \pmod{a'}$. Since $\gcd(d' - d, aa') = 1$ and $\gcd(a, a') = 1$, $(d' - d, d' - d)$ generates $\mathbb{Z}_a \times \mathbb{Z}_{a'}$, so we can find $m \in \mathbb{N}$ such that $m(d' - d) \equiv b_1 - b_0 - (y)_g - (z)_g \pmod{a}$ and $m(d' - d) \equiv b'_1 - b'_0 - S_g(y) - S_g(z) \pmod{a'}$. Let $\tilde{w} \in \mathcal{L}(\Sigma)$ be the concatenation m -times of w . It holds that $F_0 \xrightarrow{y\tilde{w}z} F_1$, the length of $y\tilde{w}z$ is divisible by p , $(y\tilde{w}z)_g \equiv b_1 - b_0 \pmod{a}$ and $S_g(y\tilde{w}z) \equiv b'_1 - b'_0 \pmod{a'}$. Thus, if $n(i) = |y\tilde{w}z|$, $M_i^{n(i)}((b_0, b'_0, F_0), (b_1, b'_1, F_1)) > 0$. Therefore, M_i is irreducible and aperiodic for every $i \geq a - 1$, so the conclusion follows from Theorem 4.9. \square

Remark 5.12. Notice that the condition in the Theorem 5.11 about the existence of $d, d' \in \mathcal{D}$ such that $T(d, d) = T(d, d') = T(d', d) = 1$ and $\gcd(d' - d, aa') = 1$ can be easily relaxed. However, we state it in that way because it allows us to check the Markov condition in a simple way, providing a tool to generate many examples of sets with digit constraints of length 2 that are jointly uniformly distributed.

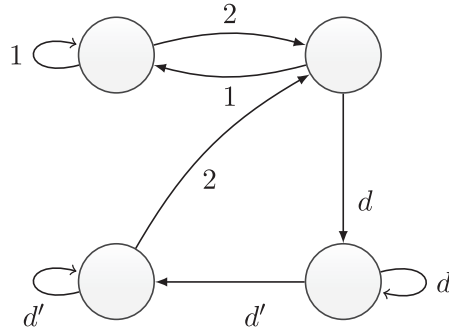


FIGURE 1. Graph that represents the 2-regular and transitive shift of finite type associated to the matrix T in Example 5.14.

Example 5.13. Let $g = 10$ and consider the set of integers

$$A := \left\{ \sum_{j=0}^m w_j 10^j : w_i \in \{0, \dots, 9\}, w_{i+1} \in \{w_i - 1, w_i, w_i + 1\} \bmod 10 \right\}.$$

Notice that $A = A_T$, where $T(d, d') = 1$ if and only if $d' \in \{d - 1, d, d + 1\} \bmod 10$. Clearly, T is 3-regular and irreducible. Also, $T(0, 0) = T(0, 1) = T(1, 0) = 1$.

Let $a, a' \in \mathbb{N}$ such that $\gcd(10, a) = \gcd(a, a') = 1$. From Theorem 5.11,

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A_T : n \equiv b \bmod a, S_{10}(n) \equiv b' \bmod a'\} \cap [0, N]|}{|A_T \cap [0, N]|} = \frac{1}{aa'}$$

for every $(b, b') \in \mathbb{Z}_a \times \mathbb{Z}_{a'}$.

Example 5.14. Let $g = 10$ and $\mathcal{D} = \{1, 2, d, d'\}$ for any $d, d' \in \{0, 3, 4, \dots, 9\}$. Let T be a matrix indexed by $\mathcal{D} \times \mathcal{D}$ and defined by

$$T = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

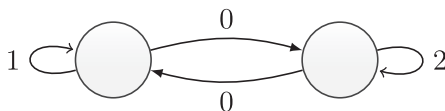
Since $T(1, 1) = T(1, 2) = T(2, 1)$, Theorem 5.11 gives that A_T is jointly uniformly distributed mod (a, a') whenever $\gcd(10, a) = \gcd(a, a') = 1$. This set can be understood as all the integers which base-10 expansion can be read in paths of the graph of Figure 1.

5.4. Example involving a sofic subshift. When Σ is a sofic subshift, more general restriction can be placed, but at the same time, it is harder to give general conditions to ensure uniform distribution. We will provide a sofic shift Σ , which is not of finite type, whose associated multiplicatively invariant set is uniformly distributed.

Example 5.15. Let \mathcal{G} be the graph presented in Figure 2 and let Σ be the sofic shift $\Sigma := \Sigma_{\mathcal{G}}$ (recall Definition 2.5). We can see that

$$\Sigma = \{x \in \{0, 1, 2\}^{\mathbb{N}_0} : \text{there is an even number of zeros between two 1s, or two 2s}\},$$

and the cover \mathcal{G} is the Fischer cover for Σ .

FIGURE 2. Fischer cover of the sofic shift Σ presented in Example 5.15.

Notice that the graph \mathcal{G} is 2-regular and aperiodic, so that Σ is a mixing, sofic and 2-regular subshift. For $g = 3$, we define the $\times 3$ -invariant set

$$A_{\Sigma} := \left\{ \sum_{j=0}^m d_j 3^j : \text{between two consecutive 1s or 2s, there is an even number of 0s} \right\}.$$

Let $a, a' \in \mathbb{N}$ such that $\gcd(3, a) = \gcd(a, a') = 1$. We can find a period $p \in \mathbb{N}$ such that $g^p \equiv 1 \pmod{a(g-1)}$. Without loss of generality, we can suppose p is even (otherwise, change p by $2p$). The aperiodicity of the matrix M_i follows because $0 \cdots 0 \in \mathcal{L}^p(\Sigma)$ is a cycle. Similarly, the word $w = 110 \cdots 0 \in \mathcal{L}^p(\Sigma)$ is a cycle in the graph \mathcal{G} . Also, $(w)_3 \equiv 4 \pmod{a}$ and $S_3(w) \equiv 2 \pmod{a'}$. By concatenating w with itself, we can generate every state in $\langle 4 \rangle \times \langle 2 \rangle \subseteq \mathbb{Z}_a \times \mathbb{Z}_{a'}$.

For instance, if $a = 5$ and $a' = 7$, $\langle 4 \rangle \times \langle 2 \rangle = \mathbb{Z}_5 \times \mathbb{Z}_7$. Considering that the word w is read in a cycle of the Fischer cover, $((w)_3, S_3(w))$ is a generator for $\mathbb{Z}_5 \times \mathbb{Z}_7$ and \mathcal{G} is aperiodic, it is possible to see that M_i is irreducible for every $i \in \mathbb{N}$. Theorem 4.9 implies that

$$\lim_{N \rightarrow \infty} \frac{|\{n \in A_{\Sigma} : n \equiv b \pmod{5}, S_3(n) \equiv b' \pmod{7}\} \cap [0, N]|}{|A_{\Sigma} \cap [0, N]|} = \frac{1}{35}$$

for every $(b, b') \in \mathbb{Z}_5 \times \mathbb{Z}_7$.

6. Transversality between multiplicatively invariant sets and arithmetic progressions

In this section, we provide a partial answer to the open question [27, Question 5.6]. Recalling the notion of mass dimension provided in Definition 2.11, the question is stated as follows. *Let $A \subseteq \mathbb{N}_0$ be a multiplicatively invariant set and P be an arithmetic progression. Is it true that $\dim_{\mathbb{M}}(A \cap P)$ is either zero or $\dim_{\mathbb{M}}(A)$?*

Let A be $\times g$ -invariant and recall the existence of a subshift Σ such that $A = A_{\Sigma}$. If A satisfies the conditions of Theorem 4.9 with $f = \text{id}$ and $a \in \mathbb{N}$ (including the Markov condition), it is possible to see that $\dim_{\mathbb{M}}(A \cap (a\mathbb{N} + b)) = \dim_{\mathbb{M}}(A)$ for every $b \in \mathbb{N}$. Therefore, this provides a positive answer to the open question in those cases. As the notion of dimension is less sensitive to rough estimations than the relative density, we will modify the construction presented in §3 to address this question for every $A = A_{\Sigma}$ when Σ is any transitive sofic subshift.

6.1. Counterexample for general cases. We will show that the answer is negative if we only suppose that the subshift Σ is only transitive or only sofic.

Let $g = a > 4$, and let $\mathcal{D} = \{d_0, d_1\}$ and $\mathcal{D}' = \{d_2, d_3, d_4\}$ be disjoint subsets of $\{0, \dots, g-1\}$. We define the subshift $\Sigma = \mathcal{D}^{\mathbb{N}_0} \cup \mathcal{D}'^{\mathbb{N}_0}$. Notice that Σ is a sofic subshift

(in fact, it is an SFT). From Proposition 2.13, it is easy to see that $\dim_{\mathbf{M}}(A_{\Sigma}) = \log(3)/\log(g)$. Since $a = g$, $n \in A_{\Sigma} \cap (a\mathbb{N} + d_0)$ if and only if d_0 is the least significant digit of n in base g , so we can conclude that

$$0 < \dim_{\mathbf{M}}(A_{\Sigma} \cap (a\mathbb{N} + d_0)) = \log(2)/\log(g) < \dim_{\mathbf{M}}(A_{\Sigma}).$$

More generally, other examples can be constructed even when $\gcd(g, a) = 1$. For this, we can use the same idea of considering a disjoint set of digits with different size and choose the elements to apply condition (5.5) of Corollary 5.5. This concludes that the sofic condition by itself is not enough to provide a positive answer to the open question. Now, we will see that assuming only transitivity is not enough either.

PROPOSITION 6.1. *For any integer $g \geq 3$, there exists a transitive subshift Σ and an arithmetic progression P such that $0 < \dim_{\mathbf{M}}(A_{\Sigma} \cap P) < \dim_{\mathbf{M}}(A_{\Sigma})$.*

Proof. Let us choose a set of digits $\mathcal{D} \subsetneq \{0, \dots, g-1\}$ with at least two elements, fixing some digits $h \notin \mathcal{D}$ and $d \in \mathcal{D}$. For every $k \in \mathbb{N}_0$, we define the set of words

$$W_k := \{hd^k z : z \in \mathcal{D}^k\} \subseteq \mathcal{L}(\{0, \dots, g-1\}^{\mathbb{N}_0}),$$

where $d^k := d \cdots d$ is the concatenation k times of d . The length of any word in W_k is $2k+1$ and $|W_k| = |\mathcal{D}|^k$. Giving an enumeration to the $|\mathcal{D}|^k$ words of W_k , we can write

$$W_k = \{w_k^j : j = 0, \dots, |\mathcal{D}|^k - 1\}.$$

By concatenating all the words in the sets W_k , we define the infinite sequence

$$w := (w_0^0)(w_1^0 w_1^1 \cdots w_1^{|\mathcal{D}|-1}) \cdots (w_k^0 \cdots w_k^{|\mathcal{D}|^k-1}) \cdots,$$

and we refer to the subword $w_k^0 \cdots w_k^{|\mathcal{D}|^k-1}$ as the block induced by W_k . Considering the full shift $(\{0, \dots, g-1\}^{\mathbb{N}_0}, \sigma)$, we can define the transitive subshift

$$\Sigma = \overline{\text{Orb}_{\sigma}(w)} := \overline{\{\sigma^n(w) : n \in \mathbb{N}_0\}}.$$

Also, it is straightforward that

$$\mathcal{L}(\Sigma) = \{x = x_0 \cdots x_m : m \in \mathbb{N}_0, x \text{ appears in the sequence } w\}.$$

We claim that $\dim_{\mathbf{M}}(A_{\Sigma}) = \log(|\mathcal{D}|)/\log(g)$. To show it, notice that $C_{g,\mathcal{D}} \subseteq A_{\Sigma}$ as $\mathcal{D}^{\mathbb{N}_0} \subseteq \Sigma$. Thus, Corollary 2.14 implies that

$$\dim_{\mathbf{M}}(A_{\Sigma}) \geq \dim_{\mathbf{M}}(C_{g,\mathcal{D}}) = \log(|\mathcal{D}|)/\log(g). \quad (6.1)$$

For the other direction, we aim to estimate the number of elements in A_{Σ} whose base- g expansion has length at most $i \in \mathbb{N}$. After the block W_i , no new words of length at most i appear in the sequence w . Therefore, any $n \in A_{\Sigma} \cap [0, g^i)$ can be generated by a subword with length at most i of the sequence

$$w(i) := w_0^0 w_1^0 \cdots w_1^{|\mathcal{D}|-1} \cdots w_i^0 \cdots w_i^{|\mathcal{D}|^i-1}.$$

In $w(i)$, there are $|w(i)| - (i - 1)$ possible starting symbols for words of length i , so there are at most $|w(i)| - (i - 1)$ words of length i . A quick estimation leads to

$$|w(i)| = \sum_{k=0}^i (2k+1)|\mathcal{D}|^k \leq (i+1)(2i+1)|\mathcal{D}|^i.$$

Therefore, $i(i+1)(2i+1)|\mathcal{D}|^i$ gives a rough estimation for the number of words of length at most i that can be read in $w(i)$. Then,

$$\frac{\log(|A_\Sigma \cap [0, g^i]|)}{i \log(g)} \leq \frac{\log(i(i+1)(2i+1)) + i \log |\mathcal{D}|}{i \log(g)}.$$

Taking $i \rightarrow \infty$ and using Lemma 2.15, $\log(|\mathcal{D}|)/\log(g) \leq \dim_M(A_\Sigma)$. This concludes that

$$\dim_M(A_\Sigma) = \log(|\mathcal{D}|)/\log(g).$$

Let $P := g\mathbb{N} + h$ be an arithmetic progression. Notice that $n \in A_\Sigma \cap P$ if and only if $n \in A_\Sigma$ and the least significant digit of n in base g is h . Thus, let us define $\Sigma(h) := \{v \in \mathcal{L}(\Sigma) : v_0 = h\}$, $A(h) := \{(v)_g : v \in \Sigma(h)\} \subseteq A_\Sigma$ and $r = g^2$. Notice that for any $i \in \mathbb{N}$ and $z \in \mathcal{D}^{i-1}$, $(hy^{i-1}z)_g \in A(h) \cap [0, r^i)$. Then, $|\mathcal{D}|^{i-1} \leq |A(h) \cap [0, r^i)|$.

However, after the appearance of the word w_{2i-1}^0 in the sequence w , the only word of length $2i$ that starts with h and can be read is hy^{2i-1} . For $j < i$, there are $|\mathcal{D}|^j$ appearances of h in the block j of w , and each occurrence can be associated with one $2i$ -long word. In the block $(w_i^0 \cdots w_i^{|\mathcal{D}|^{i-1}})$, every word of length $2i$ starting with h is given by $hy^i z$ for some $z \in \mathcal{D}^{i-1}$, and hence, there are $|\mathcal{D}|^{i-1}$ of those words in W_i . More generally, for every $j \geq i$, every word of length $2i$ starting with h in the block W_j has the form $hy^j z$, where $z \in \mathcal{D}^{2i-j-1}$, so there are $|\mathcal{D}|^{2i-j-1}$ of such words in W_j . Using these estimates,

$$|\{w \in \Sigma(h) : |w| = 2i\}| \leq \sum_{j=0}^{i-1} |\mathcal{D}|^j + \sum_{j=i}^{2i-1} |\mathcal{D}|^{2i-j-1} \leq 2i|\mathcal{D}|^{i-1}.$$

Using the same estimate for words of length less than $2i$, it follows that

$$|A(h) \cap [0, r^i)| \leq (2i)^2 |\mathcal{D}|^{i-1}.$$

Hence,

$$|\mathcal{D}|^{i-1} \leq |A(h) \cap [0, r^i)| \leq (2i)^2 |\mathcal{D}|^{i-1}.$$

It follows that

$$\frac{i-1}{2i} \frac{\log(|\mathcal{D}|)}{\log(g)} \leq \frac{\log(|A(h) \cap [0, r^i]|)}{2i \log(g)} \leq \frac{2 \log(2i) + i \log(|\mathcal{D}|)}{2i \log(g)},$$

and invoking Lemma 2.15, we conclude

$$\dim_M(A(h)) = \frac{1}{2} \frac{\log(|\mathcal{D}|)}{\log(g)}. \quad (6.2)$$

Since the arithmetic progression $P = g\mathbb{N} + h$ satisfies $P \cap A_\Sigma = A(h)$, from (6.2), it follows that

$$\dim_{\mathbb{M}}(A_\Sigma \cap P) = \frac{\dim_{\mathbb{M}}(A_\Sigma)}{2}. \quad \square$$

Notice that for every rational $\alpha \in (0, 1)$, it is possible to modify the construction of the sets W_j to construct a transitive subshift Σ such that $\dim_{\mathbb{M}}(A_\Sigma \cap P) = \alpha \dim_{\mathbb{M}}(A_\Sigma) > 0$.

6.2. Positive answer for transitive sofic subshifts. In this section, we will prove that the question [27, Question 5.6] has a positive answer when the multiplicatively invariant set can be represented by a subshift that is transitive and sofic.

Remark 6.2. For the rest of this section, consider an integer $g \geq 2$ and $a \in \mathbb{N}$. Let $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ be a transitive and sofic subshift, where $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \lambda)$ is the Fischer cover of Σ . Let $p \in \mathbb{N}$ be as given by Proposition 3.2, so that $(p, a-1)$ forms an eventual period for the function $\mathbf{f} = \text{id}$ with respect to a .

We use the sets E_i from Definition 3.4 to define a non-normalized version of the matrix M and the measure μ introduced in §3.

Definition 6.3. For any $i \in \mathbb{N}$, $b, b' \in \mathbb{Z}_a$ and $F, F' \in \mathcal{V}$, we define

$$\mathcal{M}_i((b, F), (b', F')) = |E_i(b' - b, F, F')|$$

and

$$v_i(b, F) = |\{w \in \mathcal{L}^i(\Sigma) : (w)_g \equiv b \pmod{a}, F(w) = F\}|.$$

In a straightforward way, we can rewrite Proposition 3.13 in terms of \mathcal{M}_i and v_i .

COROLLARY 6.4. *If $i \geq a-1$ and $n \in \mathbb{N}$,*

$$\mathcal{M}_i^n((b_0, F_0), (b_1, F_1)) = |\{w \in \mathcal{L}^{np}(\Sigma) : g^i(w)_g \equiv b' - b \pmod{a} \text{ and } F \xrightarrow{w} F'\}|$$

for any $b, b' \in \mathbb{Z}_a$ and $F, F' \in \mathcal{V}$. Also, $v_{i+np} = v_i \mathcal{M}_i^n$.

The main idea can be summarized as follows. First, we want to find a word $w \in \mathcal{L}(\Sigma)$ such that $i := |w| \geq a-1$, $(w)_g \equiv b \pmod{a}$ and $F(w) \in \mathcal{V}$. By studying the extensions of w to words of length $i + pn$ for some $n \in \mathbb{N}$, we will see that a significant number of these extensions \tilde{w} satisfy $(\tilde{w})_g \equiv b \pmod{a}$. As a conclusion, we aim to show $\dim_{\mathbb{M}}(A_\Sigma \cap (a\mathbb{N} + b)) = \dim_{\mathbb{M}}(A_\Sigma)$. If such a word w does not exist, then the intersection will be finite (hence, zero-dimensional).

For a given word w , it is easy to compute $(w)_g \pmod{a}$; however, deciding whether $F(w) \in \mathcal{V}$ or not could be challenging if we do not know the Fischer cover explicitly. The following lemma shows that, for our purposes, such a check is not required.

LEMMA 6.5. *Let $w \in \mathcal{L}(\Sigma)$ such that $|w| \geq a-1$ and $F(w) \notin \mathcal{V}$. Then, there exists an extension \tilde{w} of w such that $(w)_g \equiv (\tilde{w})_g \pmod{a}$, $F(\tilde{w}) \in \mathcal{V}$ and $|\tilde{w}| = |w| + pn$ for some $n \in \mathbb{N}$.*

Proof. Since Σ is transitive and sofic, there exists a synchronizing word $y \in \mathcal{L}(\Sigma)$ and also there exists $x \in \mathcal{L}(\Sigma)$ such that $wxy \in \mathcal{L}(\Sigma)$. Since y is synchronizing, any extension of y is also synchronizing. Thus, we can suppose $|xy| = mp$ for some $m \in \mathbb{N}$ (otherwise, we can extend y to meet that condition). Since y is synchronizing, the follower set of any word that ends with the subword y is equal to $F(y)$. In particular, $F := F(wxy) = F(xy) = F(y) \in \mathcal{V}$. Furthermore, every path in \mathcal{G} labelled by xy has F as terminal node and we can find a path π labelled by xy starting at some $F' \in \mathcal{V}$. As \mathcal{G} is an irreducible graph, there is a path from F to F' , labelled by some $z \in \mathcal{L}(\Sigma)$. For any $k \in \mathbb{N}$, denote

$$(xyz)^k = \underbrace{(xyz) \cdots (xyz)}_{k \text{ times}} \in \mathcal{L}(\Sigma).$$

Let $\tilde{w} = w(xyz)^{ap}$ and $\ell := |xyz|$. Using that $g^{p+j} \equiv g^p \pmod{a}$ for all $j \geq a-1$,

$$\begin{aligned} (\tilde{w})_g &= (w)_g + g^{|w|}(xyz)^p + \cdots + g^{|w|+(a-1)p\ell}(xyz)^p \\ &= (w)_g + ag^{|w|}(xyz)^p \equiv (w)_g \pmod{a}. \end{aligned}$$

As the path associated to z ends at F' , it follows that $F(\tilde{w}) = F' \in \mathcal{V}$. Also, notice that $|\tilde{w}| = |w| + \ell ap$. \square

We now establish a sufficient condition to ensure that extensions of a certain word w will visit certain congruence classes and follower sets.

LEMMA 6.6. *Let $w \in \mathcal{L}(\Sigma)$ such that $|w| \geq a-1$, $(w)_g \equiv b \pmod{a}$ and $F(w) \in \mathcal{V}$. If $\mathcal{M}_{|w|}^n((b', F'), (b, F(w))) > 0$ for some $(b', F') \in \mathbb{Z}_a \times \mathcal{V}$ and $n \in \mathbb{N}$, there exists $\tilde{w} \in \mathcal{L}(\Sigma)$ such that $(\tilde{w})_g \equiv b' \pmod{a}$, $F(\tilde{w}) = F'$ and $|\tilde{w}| = |w| + mp$ for some $m \in \mathbb{N}$.*

Proof. As $\mathcal{M}_{|w|}^n((b', F'), (b, F(w))) > 0$, Corollary 6.4 provides the existence of a word y with length np such that $g^{|w|}(y)_g \equiv b - b' \pmod{a}$ and $F' \xrightarrow{y} F(w)$. Equivalently, we can find a path π_1 in \mathcal{G} with length np , labelled by y and from F' to $F(w)$. Since the Fischer cover is irreducible, there is a path π_2 from $F(w)$ to F' ; thus, the path $\tilde{\pi} := (\pi_2\pi_1)^{p-1}\pi_2$ goes from $F(w)$ to F' and $|\tilde{\pi}| = p|\pi_2| + (p-1)|\pi_1|$ is divisible by p . If $\tilde{\pi}$ is labelled by some $x \in \mathcal{L}(\Sigma)$, we can define $\tilde{w} := w(xy)^{a-1}x \in \mathcal{L}(\Sigma)$. Since the terminal node for $\tilde{\pi}$ is F' , $F(\tilde{w}) = F'$. As $|x|$ and $|y|$ are divisible by p , it follows that $|\tilde{w}| = |w| \pmod{p}$ and the periodicity implies

$$\begin{aligned} (\tilde{w})_g &\equiv (w)_g + (a-1)g^{|w|}(x)_g + (a-1)g^{|w|}(y)_g + g^{|w|}(x)_g \\ &\equiv b + (a-1)g^{|w|}(y)_g \pmod{a}. \end{aligned}$$

Since $g^{|w|}(y)_g \equiv b - b'$, it follows that $(\tilde{w})_g \equiv b + (a-1)(b - b') \equiv b' \pmod{a}$. \square

By combining Lemmas 6.5, 6.6 and ideas of previous sections, we can show a more general version of Theorem C.

THEOREM 6.7. Let $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ be a transitive sofic subshift and $a \in \mathbb{N}$. If there exists $w \in \mathcal{L}(\Sigma)$ such that $|w| \geq a-1$ and $(w)_g \equiv b \pmod{a}$, then

$$\dim_{\mathbb{M}}(A_{\Sigma} \cap (a\mathbb{N} + b)) = \dim_{\mathbb{M}}(A_{\Sigma}). \quad (6.3)$$

In particular, either (6.3) holds or $|A_{\Sigma} \cap (a\mathbb{N} + b)| \leq g^a$.

Proof. Consider $p \in \mathbb{N}$ given by Proposition 3.2. From Lemma 6.5, we can suppose $F(w) \in \mathcal{V}$ (otherwise, use \tilde{w} instead) and define $i := |w| \geq a-1$. In particular, (p, i) is an eventual period. Consider the set

$$\mathcal{R} := \{(b', F') \in \mathbb{Z}_a \times \mathcal{V} : \text{there exists } n \in \mathbb{N} \text{ such that } \mathcal{M}_i^n((b', F'), (b, F)) > 0\}.$$

From Lemma 6.6, for each $(b', F') \in \mathcal{R}$, there is $m' \in \mathbb{N}$ such that $v_{i+m'p}(b', F') \neq 0$.

In particular, we can obtain some $m \in \mathbb{N}$ large enough such that

$$\sum_{j=0}^m v_{i+jp}(b', F') \neq 0 \quad (6.4)$$

for every $(b', F') \in \mathcal{R}$. Let $n \in \mathbb{N}$ and $(b, F) \in \mathbb{Z}_a \times \mathcal{V}$. Using Corollary 6.4,

$$\begin{aligned} \sum_{j=0}^m v_{i+(j+n)p}(b, F) &= \sum_{j=0}^m v_{i+jp} \mathcal{M}_i^n(\cdot, (b, F)) \\ &= \sum_{j=0}^m \sum_{b' \in \mathbb{Z}_a, F' \in \mathcal{V}} v_{i+jp}(b', F') \mathcal{M}_i^n((b', F'), (b, F)) \\ &= \sum_{b' \in \mathbb{Z}_a, F' \in \mathcal{V}} \left(\sum_{j=0}^m v_{i+jp}(b', F') \right) \mathcal{M}_i^n((b', F'), (b, F)) \\ &\geq \sum_{b' \in \mathbb{Z}_a, F' \in \mathcal{V}} \mathcal{M}_i^n((b', F'), (b, F)), \end{aligned}$$

where the last inequality follows from the definition of \mathcal{R} and (6.4). Combining the previous inequality, Proposition 3.6 and the characterization of Corollary 6.4, we can conclude

$$\sum_{j=0}^m v_{i+(j+n)p}(b, F) \geq |\{\pi = (\pi_0 \cdots \pi_{np-1}) : \pi \text{ is a path of } \mathcal{G} \text{ and } t(\pi) = F\}|. \quad (6.5)$$

Since \mathcal{G} is an irreducible graph, for each $F' \in \mathcal{V}$, there exists $\ell(F') \in \mathbb{N}$ such that there is a path of length $\ell(F')$ from F' to F . Thus, defining $\ell := \max_{F' \in \mathcal{V}} \ell(F')$, for every n large enough,

$$\begin{aligned} &|\{\pi = (\pi_0 \cdots \pi_{np-1}) : \pi \text{ is a path of } \mathcal{G} \text{ and } t(\pi) = F\}| \\ &\geq \sum_{F' \in \mathcal{V}} |\{\pi = (\pi_0 \cdots \pi_{np-1-\ell(F')}) : \pi \text{ is a path of } \mathcal{G} \text{ and } t(\pi) = F'\}| \\ &\geq \sum_{F' \in \mathcal{V}} |\{\pi = (\pi_0 \cdots \pi_{np-1-\ell}) : \pi \text{ is a path of } \mathcal{G} \text{ and } t(\pi) = F'\}| \\ &= |\{\pi = (\pi_0 \cdots \pi_{np-1-\ell}) : \pi \text{ is a path of } \mathcal{G}\}|. \end{aligned} \quad (6.6)$$

Since \mathcal{G} is a cover for Σ , for every $x \in \mathcal{L}^{np-1-\ell}(\Sigma)$, there exists at least one path in \mathcal{G} labelled by x . From this idea, (6.5) and (6.6), it follows that

$$\sum_{j=0}^m v_{i+(j+n)p}(b, F) \geq |\mathcal{L}^{np-\ell-1}(\Sigma)|. \quad (6.7)$$

However, by definition of $v_{i+(j+n)p}$,

$$\sum_{j=0}^m v_{i+(j+n)p}(b, F) \leq \left| \left\{ w \in \bigcup_{j=0}^m \mathcal{L}^{i+(j+n)p}(\Sigma) : (w)_g \equiv b \pmod{a} \right\} \right|. \quad (6.8)$$

Since two words of different length can represent the same integer, the right-hand side of (6.8) contains at most $m+1$ representations for each integer, then

$$\left| \left\{ w \in \bigcup_{j=0}^m \mathcal{L}^{i+(j+n)p}(\Sigma) : (w)_g \equiv b \pmod{a} \right\} \right| \leq (m+1) |A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+(m+n)p})|.$$

Combining this estimation with (6.7) and (6.8), it follows that

$$\frac{1}{m+1} |\mathcal{L}^{np-\ell-1}(\Sigma)| \leq |A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+(m+n)p})|.$$

Taking $\log(\cdot)$ and rearranging terms,

$$\begin{aligned} & \frac{np - \ell - 1}{(i + mp + np) \log(g)} \frac{\log(|\mathcal{L}^{np-\ell-1}(\Sigma)|)}{np - \ell - 1} - \frac{\log(m+1)}{(i + mp + np) \log(g)} \\ & \leq \frac{\log(|A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+(m+n)p})|)}{(i + mp + np) \log(g)}. \end{aligned}$$

Taking \liminf , we conclude that

$$h(\Sigma)/\log(g) \leq \liminf_{n \rightarrow \infty} \frac{\log(|A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+mp+np})|)}{(i + mp + np) \log(g)},$$

where $h(\Sigma)$ is the topological entropy of Σ . Since $A_\Sigma \cap (a\mathbb{N} + b) \subseteq A_\Sigma$ and recalling that $\dim_{\mathbb{M}}(A_\Sigma) = h(\Sigma)/\log(g)$, we can obtain

$$\limsup_{n \rightarrow \infty} \frac{\log(|A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+mp+np})|)}{(i + mp + np) \log(g)} \leq h(\Sigma)/\log(g).$$

Therefore,

$$\lim_{n \rightarrow \infty} \frac{\log(|A_\Sigma \cap (a\mathbb{N} + b) \cap [0, g^{i+mp+np})|)}{(i + mp + np) \log(g)} = h(\Sigma)/\log(g).$$

From Lemma 2.15 (in particular, we use a slight variant of this lemma that follows directly from the proof of [27, Lemma 3.2]),

$$\dim_{\mathbb{M}}(A_\Sigma \cap (a\mathbb{N} + b)) = h(\Sigma)/\log(g).$$

Finally, if there is no word $w \in \mathcal{L}(\Sigma)$ such that $|w| \geq a - 1$ and $(w)_g \equiv b \pmod{a}$, $A_\Sigma \cap (a\mathbb{N} + b) \cap [g^a, \infty) = \emptyset$. Consequently, $|A_\Sigma \cap (a\mathbb{N} + b)| \leq g^a$ and thus,

$$\dim_{\mathbf{M}}(A_\Sigma \cap (a\mathbb{N} + b)) = 0. \quad \square$$

From the discussion in §6.1 and Theorem 6.7, the conclusion is that the Open question 1.5 is affirmative if the multiplicatively invariant set can be represented by a transitive sofic shift. However, this is not necessarily true if the shift is merely transitive or sofic.

Nevertheless, we can extend the positive answer beyond transitive sofic subshifts.

Definition 6.8. Let $S \subseteq \mathbb{N}_0$. We define the S -gap subshift as

$$X_S := \{x \in \{0, 1\}^{\mathbb{N}_0} : \text{the number of 0s between two consecutive 1s is in } S\}.$$

For $S \subseteq \mathbb{N}_0$, we denote $A_S := A_{\Sigma_S}$.

PROPOSITION 6.9. Let $S \subseteq \mathbb{N}_0$. For any arithmetic progression P , $\dim_{\mathbf{M}}(A_S \cap P)$ is either $\dim_{\mathbf{M}}(A_S)$ or zero.

Proof. If $|S| < \infty$, the result follows from Theorem 6.7 as Σ_S is transitive and sofic. Let us suppose $|S| = \infty$. Consider s_n as the n th element of S in increasing order and define $S_n := \{s_1, \dots, s_n\}$. The subshift X_{S_n} is a transitive and sofic shift for every $n \in \mathbb{N}$, and notice that $\mathcal{L}^n(X_{S_n}) = \mathcal{L}^n(X_S)$. Consequently, it is not difficult to see that

$$h(X_S) = \sup_{n \in \mathbb{N}} h(X_{S_n}). \quad (6.9)$$

Suppose that there exists $w \in \mathcal{L}(X_S)$ such that $|w| \geq a - 1$ and $(w)_g \equiv b \pmod{a}$. Then, there exists $n_0 \in \mathbb{N}$ such that $w \in \mathcal{L}(X_{S_n})$ for every $n \geq n_0$. Since X_{S_n} is sofic and transitive, we can apply Theorem 6.7 to obtain that

$$\dim_{\mathbf{M}}(A_{S_n} \cap (a\mathbb{N} + b)) = \dim_{\mathbf{M}}(A_{S_n}) = h(X_{S_n})/\log(g). \quad (6.10)$$

However,

$$h(X_S)/\log(g) = \dim_{\mathbf{M}}(A_S) \geq \dim_{\mathbf{M}}(A_S \cap (a\mathbb{N} + b)) \geq \dim_{\mathbf{M}}(A_{S_n} \cap (a\mathbb{N} + b)). \quad (6.11)$$

From (6.10) and (6.11),

$$h(X_{S_n})/\log(g) \leq \dim_{\mathbf{M}}(A_S \cap (a\mathbb{N} + b)) \leq h(X_S)/\log(g).$$

The conclusion follows from (6.9). \square

If S is the set of prime numbers, it is easy to see that X_S is transitive but not sofic. Moreover, note that the property can be extended to subshifts that can be approximated (in terms of entropy) from within by transitive sofic subshifts. In particular, the class of transitive sofic subshifts is not optimal to provide an affirmative answer.

One possible refinement is to look at subshifts with a unique measure of maximal entropy (which includes all transitive sofic subshifts [52]). However, it is not difficult to see that the subshift constructed in Proposition 6.1 also has a unique measure of maximal

entropy. Thus, this class is not sufficient to settle the question either. Seeking a general answer to the open question, we will consider the following class.

Definition 6.10. A subshift Σ is said to be *entropy minimal* if for every subshift $\Sigma' \subsetneq \Sigma$, $h(\Sigma) > h(\Sigma')$.

The classes of transitive sofic subshifts and S -gap subshifts are examples of entropy minimal subshifts (see [10, Remark 2.4] and [24]). However, the counterexample provided in Proposition 6.1 is not entropy minimal. Based on the techniques used in this paper, we think that it should be possible to extend the affirmative answer to the open question to subshifts having left almost specification with bounded function (see [9, Definition 2.14]), which is a subclass of the entropy minimal subshifts. More generally, we propose the following conjecture.

Conjecture 6.11. Let $A \subseteq \mathbb{N}_0$ be a $\times g$ -invariant set. There exists an entropy minimal subshift $\Sigma \subseteq \{0, \dots, g-1\}^{\mathbb{N}_0}$ such that $A = A_\Sigma$ if and only if, for every arithmetic progression P , it holds that $\dim_M(A \cap P)$ is either 0 or $\dim_M(A)$.

Acknowledgements. The author thanks his advisor, Joel Moreira, for introducing him to the problem, and for his continuous support and guidance throughout this project. The author also expresses gratitude to Vaughn Climenhaga for providing useful references and to the anonymous referee for their valuable feedback and comments.

REFERENCES

- [1] S. Akiyama and A. Pethő. On canonical number systems. *Theor. Comput. Sci.* **270**(1–2) (2002), 921–933.
- [2] K. Aloui. Sur les entiers ellipsépiques: somme des chiffres et répartition dans les classes de congruence. *Period. Math. Hungar.* **70**(2) (2015), 171–208.
- [3] K. Aloui, C. Mauduit and M. Mkaouer. Somme des chiffres et répartition dans les classes de congruence pour les palindromes ellipsépiques. *Acta Math. Hungar.* **151**(2) (2017), 409–455.
- [4] T. Austin. A new dynamical proof of the Shmerkin–Wu theorem. *J. Mod. Dyn.* **18** (2022), 1–11.
- [5] W. D. Banks, A. Conflitti and I. E. Shparlinski. Character sums over integers with restricted g -ary digits. *Illinois J. Math.* **46**(3) (2002), 819–836.
- [6] W. D. Banks and I. E. Shparlinski. Arithmetic properties of numbers with restricted digits. *Acta Arith.* **112**(4) (2004), 313–332.
- [7] R. Bellman and H. N. Shapiro. On a problem in additive number theory. *Ann. of Math. (2)* **2**(49) (1948), 333–340.
- [8] J. Bésineau. Indépendance statistique d’ensembles liés à la fonction “somme des chiffres”. *Acta Arith.* **20** (1972), 401–416.
- [9] V. Climenhaga and R. Pavlov. One-sided almost specification and intrinsic ergodicity. *Ergod. Th. & Dynam. Sys.* **39**(9) (2019), 2456–2480.
- [10] V. Climenhaga and D. J. Thompson. Beyond Bowen’s specification property. *Thermodynamic Formalism (Lecture Notes in Mathematics, 2290)*. Ed. M. Pollicott and S. Vaienti. Springer, Cham, 2021, pp. 3–82.
- [11] S. Col. Diviseurs des nombres ellipsépiques. *Period. Math. Hungar.* **58**(1) (2009), 1–23.
- [12] C. Dartyge and C. Mauduit. Nombres presque premiers dont l’écriture en base r ne comporte pas certains chiffres. *J. Number Theory* **81**(2) (2000), 270–291.
- [13] C. Dartyge and C. Mauduit. Ensembles de densité nulle contenant des entiers possédant au plus deux facteurs premiers. *J. Number Theory* **91**(2) (2001), 230–255.
- [14] C. Dartyge, C. Mauduit and A. Sárközy. Polynomial values and generators with missing digits in finite fields. *Funct. Approx. Comment. Math.* **52**(1) (2015), 65–74.

- [15] C. Dartyge and A. Sárközy. The sum of digits function in finite fields. *Proc. Amer. Math. Soc.* **141**(12) (2013), 4119–4124.
- [16] M. Drmota and C. Mauduit. Weyl sums over integers with affine digit restrictions. *J. Number Theory* **130**(11) (2010), 2404–2427.
- [17] M. Drmota, C. Mauduit and J. Rivat. The sum-of-digits function of polynomial sequences. *J. Lond. Math. Soc.* (2) **84**(1) (2011), 81–102.
- [18] P. Erdős, C. Mauduit and A. Sárközy. On arithmetic properties of integers with missing digits. I. Distribution in residue classes. *J. Number Theory* **70**(2) (1998), 99–120.
- [19] P. Erdős, C. Mauduit and A. Sárközy. On arithmetic properties of integers with missing digits. II. Prime factors. *Discrete Math.* **200** (1999), 149–164; Paul Erdős memorial collection.
- [20] N. J. Fine. The distribution of the sum of digits (mod p). *Bull. Amer. Math. Soc.* **71** (1965), 651–652.
- [21] R. Fischer. Sofic systems and graphs. *Monatsh. Math.* **80**(3) (1975), 179–186.
- [22] R. Fischer. Graphs and symbolic dynamics. *Topics in Information Theory (Second Colloquium, Keszthely, 1975) (Colloquia Mathematica Societatis János Bolyai, 16)*. Ed. I. Csizsár and P. Elias. János Bolyai Mathematical Society, Budapest, 1977, pp. 229–244.
- [23] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49.
- [24] F. García-Ramos and R. Pavlov. Extender sets and measures of maximal entropy for subshifts. *J. Lond. Math. Soc.* (2) **100**(3) (2019), 1013–1033.
- [25] A. O. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.* **13** (1967/68), 259–265.
- [26] B. Gittenberger and J. M. Thuswaldner. Asymptotic normality of b -additive functions on polynomial sequences in the Gaussian number field. *J. Number Theory* **84**(2) (2000), 317–341.
- [27] D. Glasscock, J. Moreira and F. K. Richter. Additive and geometric transversality of fractal sets in the integers. *J. Lond. Math. Soc.* (2) **109**(5) (2024), Paper no. e12902, 55 pp.
- [28] I. Kátai. Construction of number systems in algebraic number fields. *Ann. Univ. Sci. Budapest. Sect. Comput.* **18** (1999), 103–107.
- [29] A. J. Kempner. A curious convergent series. *Amer. Math. Monthly* **21**(2) (1914), 48–50.
- [30] D.-H. Kim. On the joint distribution of q -additive functions in residue classes. *J. Number Theory* **74**(2) (1999), 307–336.
- [31] S. Konyagin. Arithmetic properties of integers with missing digits: distribution in residue classes. *Period. Math. Hungar.* **42**(1–2) (2001), 145–162.
- [32] S. Konyagin, C. Mauduit and A. Sárközy. On the number of prime factors of integers characterized by digit properties. *Period. Math. Hungar.* **40**(1) (2000), 37–52.
- [33] A. Kovács. On expansions of Gaussian integers with non-negative digits. *Math. Pannon.* **10**(2) (1999), 177–191.
- [34] B. Kovács. Canonical number systems in algebraic number fields. *Acta Math. Acad. Sci. Hungar.* **37**(4) (1981), 405–407.
- [35] D. A. Levin and Y. Peres. *Markov Chains and Mixing Times*, 2nd edn. American Mathematical Society, Providence, RI, 2017; with contributions by Elizabeth L. Wilmer, with a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
- [36] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding (Cambridge Mathematical Library)*, 2nd edn. Cambridge University Press, 2021.
- [37] M. G. Madritsch. The sum-of-digits function of canonical number systems: distribution in residue classes. *J. Number Theory* **132**(12) (2012), 2756–2772.
- [38] K. Mahler. On the generating functions of integers with a missing digit. *K'o Hsüeh (Science)* **29** (1947), 265–267.
- [39] C. Mauduit and J. Rivat. Sur un problème de Gelfond: la somme des chiffres des nombres premiers. *Ann. of Math.* (2) **171**(3) (2010), 1591–1646.
- [40] C. Mauduit and A. Sárközy. On the arithmetic structure of the integers whose sum of digits is fixed. *Acta Arith.* **81**(2) (1997), 145–173.
- [41] J. Maynard. Primes with restricted digits. *Invent. Math.* **217**(1) (2019), 127–218.
- [42] J. Maynard. Primes and polynomials with restricted digits. *Int. Math. Res. Not. IMRN* **2022**(14) (2022), 1–23 [10626–10648 on table of contents].
- [43] N. G. Moshchevitin and I. D. Shkredov. On the multiplicative properties modulo m of numbers with missing digits. *Mat. Z.* **81**(3) (2007), 385–404.
- [44] J. R. Norris. Discrete-time Markov chains. *Markov Chains (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, 1997, pp. 1–59.
- [45] J. S. Rosenthal. Convergence rates for Markov chains. *SIAM Rev.* **37**(3) (1995), 387–405.
- [46] J. Sándor and B. Crstici. *Handbook of Number Theory. II*. Kluwer Academic Publishers, Dordrecht, 2004.

- [47] P. Shmerkin. On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions. *Ann. of Math. (2)* **189**(2) (2019), 319–391.
- [48] C. Spandl. Computing the topological entropy of shifts. *MLQ Math. Log. Q.* **53**(4–5) (2007), 493–510.
- [49] C. Swaenepoel. Prescribing digits in finite fields. *J. Number Theory* **189** (2018), 97–114.
- [50] J. M. Thuswaldner. The sum of digits function in number fields: distribution in residue classes. *J. Number Theory* **74**(1) (1999), 111–125.
- [51] A. Walker and A. Walker. Arithmetic progressions with restricted digits. *Amer. Math. Monthly* **127**(2) (2020), 140–150.
- [52] B. Weiss. Subshifts of finite type and sofic systems. *Monatsh. Math.* **77** (1973), 462–474.
- [53] M. Wu. A proof of Furstenberg's conjecture on the intersections of $\times p$ - and $\times q$ -invariant sets. *Ann. of Math. (2)* **189**(3) (2019), 707–751.
- [54] H. Yu. Additive properties of numbers with restricted digits. *Algebra Number Theory* **15**(5) (2021), 1283–1301.