

Variable-Length Coding of Two-Sided Asymptotically Mean Stationary Measures

Łukasz Dębowski

debowski@cwi.nl

Centrum Wiskunde & Informatica

Science Park 123, NL-1098 XG Amsterdam,

The Netherlands

Abstract

We collect several observations that concern variable-length coding of two-sided infinite sequences in a probabilistic setting. Attention is paid to images and preimages of asymptotically mean stationary measures defined on subsets of these sequences. We point out sufficient conditions under which the variable length coding and its inverse preserve asymptotic mean stationarity. Moreover, conditions for preservation of shift-invariant σ -fields and the finite-energy property are discussed, and the block entropies for stationary means of coded processes are related in some cases. Subsequently, we apply certain of these results to construct a stationary nonergodic process with a desired linguistic interpretation.

Key words: asymptotically mean stationary processes, variable-length coding, synchronization, shift-invariant algebras, complete fix-free sets, finite-energy processes, block entropy

MSC 2000: 60G10, 28D99, 94A45, 94A17

Running head: Variable-Length Coding of Two-Sided AMS Measures

1 Introduction

Let \mathbb{X} and \mathbb{Y} be a pair of countable sets, called here alphabets. Fixing λ as the empty string, denote the set of nonempty strings over an alphabet \mathbb{X} as $\mathbb{X}^+ := \bigcup_{n \in \mathbb{N}} \mathbb{X}^n$ and the set of all strings as $\mathbb{X}^* := \mathbb{X}^+ \cup \{\lambda\}$. The set of one-sided infinite sequences $x^{\mathbb{N}} = (x_i)_{i \in \mathbb{N}} = x_1 x_2 x_3 \dots$ is written $\mathbb{X}^{\mathbb{N}}$ and the set of two-sided infinite sequences $x^{\mathbb{Z}} = (x_i)_{i \in \mathbb{Z}} = \dots x_{-1} x_0 x_1 x_2 \dots$ is denoted by $\mathbb{X}^{\mathbb{Z}}$. (Mind the bold-face dot between the 0-th and the first symbol.) Shorthands $x^n := (x_i)_{1 \leq i \leq n}$ and $x_k^l := (x_i)_{k \leq i \leq l}$ denote substrings, whereas $|x|$ is the length of a string x .

Subsequently, consider a function $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ that maps single symbols into strings. We will extend it to $f^* : \mathbb{X}^* \rightarrow \mathbb{Y}^*$, $f^{\mathbb{N}} : \mathbb{X}^{\mathbb{N}} \rightarrow \mathbb{Y}^{\mathbb{N}} \cup \mathbb{Y}^*$, and $f^{\mathbb{Z}} : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}} \cup (\mathbb{Y}^* \times \mathbb{Y}^*)$ defined as

$$f^*(x^n) := f(x_1)f(x_2)\dots f(x_n), \quad (1)$$

$$f^{\mathbb{N}}(x^{\mathbb{N}}) := f(x_1)f(x_2)f(x_3)\dots, \quad (2)$$

$$f^{\mathbb{Z}}(x^{\mathbb{Z}}) := \dots f(x_{-1})f(x_0) \cdot f(x_1)f(x_2)\dots, \quad (3)$$

where $x_i \in \mathbb{X}$. These extensions are known in literature under several names, such as “variable-length coding” [23] or “sequence morphisms” [3]. The finite extension (1) plays a fundamental role in the definition of instantaneous codes in information theory [9]. On the other hand, probabilistic analyses that involve strong laws and ergodic theorems necessarily operate on infinite sequences, cf., e.g., [24, 19, 21]. For these analyses, extensions (2) and (3) seem more natural, and the variable-length coding (3) has been discussed by communication engineers for a few decades [7, 17, 27].

Fix a sufficiently rich probability space (Ω, \mathcal{J}, P) , and let $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ and $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ denote the standard measurable spaces of two-sided infinite sequences. Let us consider a “shrunk” stochastic process $(X_i)_{i \in \mathbb{Z}} : (\Omega, \mathcal{J}) \rightarrow (\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ and an “expanded” process $(Y_i)_{i \in \mathbb{Z}} : (\Omega, \mathcal{J}) \rightarrow (\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ related through an almost sure equality

$$(Y_i)_{i \in \mathbb{Z}} = f^{\mathbb{Z}}((X_i)_{i \in \mathbb{Z}}), \quad (4)$$

assuming that $\lim_n |f^*(X_{-m}^n)| = \lim_m |f^*(X_{-m}^n)| = \infty$ almost surely. Throughout the article, the distributions of these processes will be written as

$$\mu = P((X_i)_{i \in \mathbb{Z}} \in \cdot) \quad \text{and} \quad \nu = P((Y_i)_{i \in \mathbb{Z}} \in \cdot) = \mu \circ (f^{\mathbb{Z}})^{-1}. \quad (5)$$

Having denoted the shift operation as $T(x^{\mathbb{Z}}) := \dots x_0 x_1 x_2 x_3 \dots = (x_{i+1})_{i \in \mathbb{Z}}$, a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ is called asymptotically mean stationary (AMS) if the limits

$$\bar{\mu}(A) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu \circ T^{-i}(A) \quad (6)$$

exist for all $A \in \mathcal{X}^{\mathbb{Z}}$, see [17, 21]. The limit $\bar{\mu}$, if it exists as a total function $\mathcal{X}^{\mathbb{Z}} \rightarrow \mathbb{R}$, forms a stationary measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$, i.e., $\bar{\mu} \circ T^{-1} = \bar{\mu}$, and is called the stationary mean of μ . It is possible that limits (6) exist for a computable measure μ and all cylinder sets but they do not exist for some other sets, see Example 6.3 later.

As we shall show, under mild conditions, the transported measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ is AMS if μ is AMS. A weaker proposition, assuming a stationary μ , was obtained in [17, Example 6]. Besides coding theory, stationary means $\mu \circ (f^{\mathbb{Z}})^{-1}$ of variable-length coded measures appear in disguise in statistical applications such as length-biased sampling [10] or philosophical probabilistic puzzles such as the Sleeping Beauty problem [14].

An application at the interface between information theory and linguistics has drawn our attention to the question whether a few specific properties of a stationary measure μ can be simultaneously preserved by the stationary mean $\mu \circ (f^{\mathbb{Z}})^{-1}$ for a certain injection $f^{\mathbb{Z}} : \mathbb{X} \rightarrow \mathbb{Y}^*$, where \mathbb{X} is infinite, and \mathbb{Y} is finite. In this article, we gather several results of independent interest that concern partly relaxed and partly more general cases of our original problem. The question that stimulated our research will be presented at the end of this section and answered in positive later.

We shall not discuss measures on one-sided sequences, see [22, 17], since they do not arise naturally in the methods and applications considered here. However, there appear a few more specific conditions on the coding function $f^{\mathbb{Z}}$, which appeal to two-sidedness of coded sequences. The first condition has to do with various concepts of synchronization, see [26, 6, 2].

Definition 1.1 *A function $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$ is called a synchronizable injection if π is an injection and $T^i \pi(x^{\mathbb{Z}}) = \pi(b^{\mathbb{Z}})$ for an $i \in \mathbb{Z}$ implies $T^j x^{\mathbb{Z}} = b^{\mathbb{Z}}$ for some $j \in \mathbb{Z}$.*

For example, $f^{\mathbb{Z}}$ is a synchronizable injection for a comma-separated code $f(x) = g(x)c$, where $c \in \mathbb{Y}^m$, and $g : \mathbb{X} \rightarrow (\mathbb{Y}^m \setminus \{c\})^*$ is an injection.

Other conditions considered are more local. Let us recall that a set of strings $\mathcal{L} \subset \mathbb{Y}^*$ is called (i) *prefix-free* if $w \neq zs$ for $w, z \in \mathcal{L}$ and $s \in \mathbb{Y}^+$, (ii) *suffix-free* if $w \neq sz$ for $w, z \in \mathcal{L}$ and $s \in \mathbb{Y}^+$, (iii) *fix-free* if it is both prefix-free and suffix-free, and (iv) *complete* if it satisfies the Kraft equality $\sum_{w \in \mathcal{L}} |\mathbb{Y}|^{-|w|} = 1$, where $|\mathbb{Y}|$ is the cardinality of \mathbb{Y} .

Definition 1.2 *A function $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is called (complete) prefix/suffix/fix-free if f is an injection and the image $f(\mathbb{X})$ is respectively (complete) prefix/suffix/fix-free. For finite $f(\mathbb{X})$, f is called finite.*

For instance, the set $\{01, 000, 100, 110, 111, 0010, 0011, 1010, 1011\}$ is complete fix-free with respect to $\mathbb{Y} = \{0, 1\}$ [16, 1]. The aforementioned comma-separated code $f(x) = g(x)c$ is prefix-free but it is not complete.

The main results of this paper are as follows:

- (i) The measure $\mu \circ (f^{\mathbb{Z}})^{-1}$ is AMS for an AMS measure μ , provided that the expansion rate $\bar{l}(x^{\mathbb{Z}}) := \lim_n n^{-1} \sum_{i=1}^n |f(x_i)|$ is in the range $(0, \infty)$ μ -almost everywhere (Section 2). This result generalizes [17, Example 6], where $\mu \circ (f^{\mathbb{Z}})^{-1}$ was shown AMS provided that μ is stationary, $f^{\mathbb{Z}}$ is an injection, and the μ -expectation of \bar{l} is finite.
- (ii) The shift-invariant algebras for processes $(X_i)_{i \in \mathbb{Z}}$ and $f^{\mathbb{Z}}((X_i)_{i \in \mathbb{Z}})$ remain in one-to-one correspondence, and their distributions coincide on these algebras if $f^{\mathbb{Z}}$ is a synchronizable injection (Section 3).

- (iii) The measure $\nu \circ f^{\mathbb{Z}}$ is stationary or AMS respectively for a stationary or an AMS measure ν if f is complete fix-free (Section 4).
- (iv) Write the cylinder sets as $[u] := \{x^{\mathbb{Z}} : x^{|u|} = u\}$. As defined in [25], a measure μ has finite energy if conditional probabilities of cylinder sets are uniformly exponentially damped, i.e., if

$$\mu([uv]) \leq Kc^{|v|}\mu([u]) \quad (7)$$

for certain $c < 1$ and $K < \infty$. (Condition (7) may be only satisfied for $c \geq |\mathbb{X}|^{-1}$ and, for a finite alphabet \mathbb{X} , (7) implies that the length of the longest nonoverlapping repeat in the μ -distributed block of length n is almost surely bounded by $O(\log n)$ [25].) We will show that the stationary mean $\bar{\mu}$ has also finite energy if (7) holds. Moreover, $\mu \circ (f^{\mathbb{Z}})^{-1}$ has finite energy if μ has finite energy and f is finite prefix-free (Section 5).

- (v) Block entropy for a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ is the function

$$H_{\mu}(i; n) := - \sum_{u \in \mathbb{X}^n} \mu(T^{-i}[u]) \log \mu(T^{-i}[u]) = \int \mu(T^{-i}[x_{i+1}^{i+n}]) d\mu(x^{\mathbb{Z}}), \quad (8)$$

where we also use the shorthand $H_{\mu}(n) := H_{\mu}(0; n)$. We will demonstrate that for a fixed length injection $f : \mathbb{X} \rightarrow \mathbb{Y}^K$, a finite \mathbb{X} , and $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$, block entropies $H_{\bar{\mu}}(n)$ and $H_{\bar{\nu}}(nK)$ of the stationary means do not differ more than a constant (Section 6).

We have researched these topics while seeking for a class of nonergodic processes $(\bar{Y}_i)_{i \in \mathbb{Z}}$ that satisfy four conditions:

- (a) $(\bar{Y}_i)_{i \in \mathbb{Z}}$ is a process over a finite alphabet $\mathbb{Y} = \{0, 1, \dots, D-1\}$,
- (b) $(\bar{Y}_i)_{i \in \mathbb{Z}}$ is stationary,
- (c) $(\bar{Y}_i)_{i \in \mathbb{Z}}$ has finite energy, and
- (d) there exist independent equidistributed binary random variables $(\bar{Z}_k)_{k \in \mathbb{N}}$, $\bar{P}(\bar{Z}_k = z) = 1/2$, $z \in \{0, 1\}$, measurable against the shift-invariant σ -field of $(\bar{Y}_i)_{i \in \mathbb{Z}}$ such that

$$\liminf_{n \rightarrow \infty} n^{-\beta} |\bar{U}_{\bar{\delta}}(n)| > 0 \quad (9)$$

holds for a certain $\beta \in (0, 1)$, all $\bar{\delta} \in (1/2, 1)$, and the sets $\bar{U}_{\bar{\delta}}(n) := \{k \in \mathbb{N} : \bar{P}(\bar{s}_k(\bar{Y}^n) = \bar{Z}_k) \geq \bar{\delta}\}$ of well-predictable \bar{Z}_k 's, where functions \bar{s}_k satisfy

$$\lim_{n \rightarrow \infty} \bar{P}(\bar{s}_k(\bar{Y}_{i+1}^{i+n}) = \bar{Z}_k) = 1, \quad \forall i \in \mathbb{Z}. \quad (10)$$

As demonstrated in [12], properties (a)–(d) imply a power-law growth of the number of distinct nonterminal symbols in the shortest grammar-based compression of the block \bar{Y}_{i+1}^{i+n} , see [20, 8]. In a linguistic interpretation posited in manuscript [12], variables \bar{Y}_i stand for consecutive letters of an infinitely long text, whereas the values of variables \bar{Z}_i stand for random facts repetitively described in the text. Since nonterminal symbols of grammar-based compressed

texts in natural language often correspond to words of a particular language (we mean words in the common sense of strings of letters separated by spaces), the demonstrated implication forms a new explanation of a power-law growth of text vocabulary (known as Zipf's law in linguistics [28]). Precisely, the explanation takes the form of the statement: If an n -letter long text describes n^β independent facts in a repetitive way, then the text contains at least $n^\beta / \log n$ different words.

Properties (b)–(d), but not (a), are satisfied by the following process.

Example 1.3 ([13]) *Let $(X_i)_{i \in \mathbb{Z}}$ be a process on (Ω, \mathcal{J}, P) , where variables*

$$X_i = (K_i, Z_{K_i}) \quad (11)$$

assume values from an infinite alphabet $\mathbb{X} = \mathbb{N} \times \{0, 1\}$, variables K_i and Z_k are probabilistically independent, K_i are distributed according to a power law, $P(K_i = k) = k^{-\alpha} / \zeta(\alpha)$, $\alpha > 1$, $\zeta(\alpha) := \sum_{k=1}^{\infty} k^{-\alpha}$, and Z_k are equidistributed, $P(Z_k = z) = 1/2$, $z \in \{0, 1\}$.

Let us write $u \sqsubseteq v$ when a sequence or a string v contains a string u as a substring. For $\mathbb{X} = \mathbb{N} \times \{0, 1\}$ and $v \in \mathbb{X}^{\mathbb{Z}} \cup \mathbb{X}^*$, define the predictors

$$s_k(v) := \begin{cases} 0 & \text{if } (k, 0) \sqsubseteq v \text{ and } (k, 1) \not\sqsubseteq v, \\ 1 & \text{if } (k, 1) \sqsubseteq v \text{ and } (k, 0) \not\sqsubseteq v, \\ 2 & \text{else.} \end{cases}$$

Variables Z_k are measurable against the shift-invariant σ -field of $(X_i)_{i \in \mathbb{Z}}$ since they satisfy $Z_k = s_k((X_i)_{i \in \mathbb{Z}})$ almost surely. Moreover,

$$\lim_{n \rightarrow \infty} P(s_k(X_{i+1}^{i+n}) = Z_k) = 1, \quad i \in \mathbb{Z}, \quad (12)$$

and

$$|U_\delta(n)| \geq \left[\frac{n}{-\zeta(\alpha) \log(1 - \delta)} \right]^{1/\alpha} \quad (13)$$

for $\delta \in (1/2, 1)$ and $U_\delta(n) := \{k \in \mathbb{N} : P(s_k(X^n) = Z_k) \geq \delta\}$, as shown in [12].

We have supposed that a suitable distribution over a finite alphabet can be constructed as the stationary mean of a certain encoding of the process (11). The results of Sections 2 through 6 suggest the following statement:

Proposition 1.4 *Let $\mu = P((X_i)_{i \in \mathbb{Z}} \in \cdot)$ be the distribution of the process from Example 1.3 and put $\mathbb{Y} = \{0, 1, 2\}$. Consider the coding function $f : \mathbb{X} \mapsto \mathbb{Y}^+$ given as*

$$f(k, z) = b(k)z2, \quad (14)$$

where $1b(k) \in \{0, 1\}^+$ is the binary representation of a natural number k . The process $(\bar{Y}_i)_{i \in \mathbb{Z}}$ distributed according to the stationary mean $\bar{P}((\bar{Y}_i)_{i \in \mathbb{Z}} \in \cdot) = \mu \circ (f^{\mathbb{Z}})^{-1}$ satisfies conditions (a)–(d) for $\beta = \alpha^{-1}$ and $\zeta(\alpha) > 4$. Variables \bar{Z}_k may be constructed as $\bar{Z}_k = \bar{s}_k((\bar{Y}_i)_{i \in \mathbb{Z}})$, where

$$\bar{s}_k(w) := \begin{cases} 0 & \text{if } 2b(k)02 \sqsubseteq w \text{ and } 2b(k)12 \not\sqsubseteq w, \\ 1 & \text{if } 2b(k)12 \sqsubseteq w \text{ and } 2b(k)02 \not\sqsubseteq w, \\ 2 & \text{else} \end{cases} \quad (15)$$

for $w \in \mathbb{Y}^{\mathbb{Z}} \cup \mathbb{Y}^$.*

This proposition is proved in the final Section 7. The inequality $\zeta(\beta^{-1}) > 4$ holds for $\beta > 0.7728\dots$. Mind that processes $(\bar{Y}_i)_{i \in \mathbb{Z}}$ and $(X_i)_{i \in \mathbb{Z}}$ live on different probability spaces, say $(\Omega, \mathcal{J}, \bar{P})$ and (Ω, \mathcal{J}, P) , respectively. Obviously, the coding function (14) is prefix-free, and its extension $f^{\mathbb{Z}}$ is a synchronizable injection.

2 AMS measures and finite expansion rate

Previous accounts of AMS measures on two-sided sequences can be found in [17, 15]. Let us recall a few useful facts. First of all, for the shift $T(x^{\mathbb{Z}}) := \dots x_0 x_1 x_2 x_3 \dots = (x_{i+1})_{i \in \mathbb{Z}}$, an AMS measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ can be equivalently characterized as such that the almost sure ergodic theorem is satisfied, i.e., the limit $\lim_n n^{-1} \sum_{i=0}^{n-1} g \circ T^i$ exists μ -almost everywhere for every nonnegative measurable function $g : (\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}}) \rightarrow (\mathbb{R}, \mathcal{R})$ [17, Theorem 1]. Trivially, $\bar{\mu} = \mu$ for a stationary μ . However, the equality

$$\bar{\mu}(A) = \mu(A), \quad A \in \mathcal{I}_{\mathbb{X}}, \quad (16)$$

is also satisfied for the T -invariant algebra $\mathcal{I}_{\mathbb{X}} := \{A \in \mathcal{X}^{\mathbb{Z}} : T^{-1}A = A\}$ in the general AMS case. This follows directly from (6), see [17]. Extending the concept of ergodicity, usually discussed for stationary measures, an AMS measure μ is called *ergodic* if $\mu(A) \in \{0, 1\}$ for all $A \in \mathcal{I}_{\mathbb{X}}$.

The lemma below is mostly a well known fact:

Lemma 2.1 (cf. [15, Theorem 0]) *A measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ is AMS if and only if there exists a stationary measure τ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ such that $\tau \gg \mu$. In the latter case, we have $\tau \gg \bar{\mu} \gg \mu$.*

Remark: The notation $\tau \gg \mu$ stands for measure dominance, i.e., $\tau(A) = 0$ implies that $\mu(A) = 0$ for all sets A in the domain of μ .

The proof in [15] does not cover the inequality $\tau \gg \bar{\mu}$. To justify it, let us observe that $\tau \gg \mu$ and $\tau(A) = 0$ imply $\mu(T^{-i}A) = 0$. Hence $\bar{\mu}(A) = 0$ as well. Moreover, the proof in [15] cannot be carried to the one-sided case since it applies invariant sets of form $\bigcup_{i \in \mathbb{Z}} T^i A$. The same trick resurfaces in Proposition 2.3 below and in Section 3, where synchronizable injections are considered.

By the definition, $\bar{\mu}(A) = \lim_n \int \left[n^{-1} \sum_{i=0}^{n-1} \mathbf{1}_{\{T^i x^{\mathbb{Z}} \in A\}} \right] d\mu(x^{\mathbb{Z}})$. Hence a useful frequency interpretation follows by the dominated convergence.

Lemma 2.2 *For an AMS measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$,*

$$\bar{\mu}(A) = \int \left[\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{\{T^i x^{\mathbb{Z}} \in A\}} \right] d\mu(x^{\mathbb{Z}}), \quad A \in \mathcal{X}^{\mathbb{Z}}. \quad (17)$$

Remark: In the ergodic case, the integrated expression is almost everywhere constant. Symbol $\mathbf{1}_{\{\varphi\}}$ therein denotes the indicator function, i.e., $\mathbf{1}_{\{\varphi\}} := 1$ if φ is true and $\mathbf{1}_{\{\varphi\}} := 0$ otherwise.

Now we move on to variable-length coding of an AMS measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$. Let $f : \mathbb{X} \rightarrow \mathbb{Y}^*$, $l_i(x^{\mathbb{Z}}) := |f(x_i)|$, and $S(x^{\mathbb{Z}}, n) := \sum_{i=1}^n l_i(x^{\mathbb{Z}})$. By the ergodic theorem [19, Theorem 9.6] and Lemma 2.1, the limit

$$\bar{l}(x^{\mathbb{Z}}) := \lim_{n \rightarrow \infty} \frac{S(x^{\mathbb{Z}}, n)}{n} = \mathbf{E}_{\bar{\mu}}(l_1 \mid \mathcal{I}_{\mathbb{X}})(x^{\mathbb{Z}}) \quad (18)$$

exists both $\bar{\mu}$ - and μ -almost everywhere. We will call the function $\bar{l}(\cdot)$ the expansion rate, whereas its expectation will be denoted by

$$L := \int \bar{l} d\mu = \int \bar{l} d\bar{\mu} = \int l_1 d\bar{\mu}. \quad (19)$$

Let $T : (y_i)_{i \in \mathbb{Z}} \mapsto (y_{i+1})_{i \in \mathbb{Z}}$ also denote the shift $\mathbb{Y}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$. Put

$$\begin{aligned} F(A, k, x^{\mathbb{Z}}) &:= \mathbf{1}_{\{T^k f^{\mathbb{Z}}(x^{\mathbb{Z}}) \in A\}}, \\ G(A, x^{\mathbb{Z}}) &:= \sum_{k=0}^{|f(x_1)|-1} F(A, k, x^{\mathbb{Z}}), \end{aligned}$$

where $\sum_{k=0}^{-1} := 0$. By the quasiperiodic identity

$$T^{|f(x_1)|} f^{\mathbb{Z}}(x^{\mathbb{Z}}) = f^{\mathbb{Z}}(Tx^{\mathbb{Z}}),$$

we have

$$F(T^{-1}A, k, x^{\mathbb{Z}}) = F(A, k+1, x^{\mathbb{Z}}), \quad (20)$$

$$F(A, |f(x_1)|, x^{\mathbb{Z}}) = F(A, 0, Tx^{\mathbb{Z}}), \quad (21)$$

$$G(T^{-1}A, k, x^{\mathbb{Z}}) = G(A, k, x^{\mathbb{Z}}) - F(A, 0, x^{\mathbb{Z}}) + F(A, 0, Tx^{\mathbb{Z}}). \quad (22)$$

The following proposition is a direct consequence of the above identities and the two previous lemmas.

Proposition 2.3 *Let μ be an AMS measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ and suppose that the expansion rate (18) is μ -almost surely in the range $(0, \infty)$ for an $f : \mathbb{X} \rightarrow \mathbb{Y}^*$. Then $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ is an AMS measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ with the stationary mean*

$$\bar{\nu}(A) = \int \left[\lim_{n \rightarrow \infty} \frac{1}{S(x^{\mathbb{Z}}, n)} \sum_{k=0}^{S(x^{\mathbb{Z}}, n)-1} F(A, k, x^{\mathbb{Z}}) \right] d\mu(x^{\mathbb{Z}}) \quad (23)$$

$$= \int [\bar{l}(x^{\mathbb{Z}})]^{-1} \left[\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} G(A, T^i x^{\mathbb{Z}}) \right] d\mu(x^{\mathbb{Z}}) \quad (24)$$

$$= \int [\bar{l}(x^{\mathbb{Z}})]^{-1} \left[\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} G(A, T^i x^{\mathbb{Z}}) \right] d\bar{\mu}(x^{\mathbb{Z}}), \quad A \in \mathcal{Y}^{\mathbb{Z}}. \quad (25)$$

Proof: The transported measure ν is a measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ if and only if $\lim_n S(x^{\mathbb{Z}}, n) = \infty$ μ -almost surely. This condition is satisfied. Observe that the limits in the brackets in (23)–(25) exist μ - and $\bar{\mu}$ -almost surely. Consequently, the integrals are equal. Denote the right-hand side of (23) as $\tau(A)$. The function τ is a stationary measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ by the dominated convergence and the Vitali-Hahn-Saks theorem. Suppose that there exists a set $A \in \mathcal{Y}^{\mathbb{Z}}$ such that $\nu(A) > \tau(A) = 0$. Then we would have $\nu(B) > \tau(B) = 0$ for $B = \bigcup_{i \in \mathbb{Z}} T^i A$. But B is shift-invariant, so $\tau(B) = \nu(B)$ by formula (23). Thus, our assumption was false, and we rather have $\tau \gg \nu$. Hence ν is AMS in view of Lemma 2.1. Moreover, τ coincides with the expression for $\bar{\nu}$ given by Lemma 2.2. \square

Corollary 2.4 *If the expansion rate (18) is μ -almost surely in the range $(0, \infty)$ and $\bar{\mu} = \bar{\tau}$ for two AMS measures μ and τ , then*

$$\overline{\mu \circ (f^{\mathbb{Z}})^{-1}} = \overline{\tau \circ (f^{\mathbb{Z}})^{-1}}.$$

Putting $\tau = \bar{\mu}$, we obtain $\bar{\nu} = \overline{\bar{\mu} \circ (f^{\mathbb{Z}})^{-1}}$ and $\bar{\nu} \gg \bar{\mu} \circ (f^{\mathbb{Z}})^{-1}$ for $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$.

Under much stronger assumptions, there exists a “finite-sum” expression for the stationary mean $\bar{\nu}$ in terms of $\bar{\mu}$, noticed by Kieffer and Gray for $\bar{\mu} = \mu$ [17, Example 6]. Namely, we can construct a stationary measure ρ by averaging the stationary mean $\bar{\mu} \circ (f^{\mathbb{Z}})^{-1}$ over a randomized shift within the quasiperiod $|f(x_1)|$. This idea is more generic, see [7, 18].

Proposition 2.5 *Let μ be an AMS measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ and suppose that the expected expansion rate (19) is in the range $(0, \infty)$ for an $f : \mathbb{X} \rightarrow \mathbb{Y}^*$. Then there exists a stationary measure*

$$\rho(A) = \frac{1}{L} \int G(A, x^{\mathbb{Z}}) d\bar{\mu}(x^{\mathbb{Z}}), \quad A \in \mathcal{Y}^{\mathbb{Z}}. \quad (26)$$

Remark: Under the above assumptions, the expansion rate $\bar{l}(\cdot)$ may vanish on a set of positive measure.

Proof: Stationarity of ρ was discussed in [17] for an injective $f^{\mathbb{Z}}$. The following proof is more general. First of all, we have $\rho(\mathbb{Y}^{\mathbb{Z}}) = 1$, whereas the countable additivity follows by the dominated convergence theorem. As for stationarity, we obtain

$$\rho(T^{-1}A) - \rho(A) = L^{-1} \int (F(A, 0, Tx^{\mathbb{Z}}) - F(A, 0, x^{\mathbb{Z}})) d\bar{\mu}(x^{\mathbb{Z}}) = 0$$

from (22) and $\bar{\mu} \circ T^{-1} = \bar{\mu}$. \square

Although ρ does not necessarily equal $\bar{\nu}$, it dominates the measure ν .

Corollary 2.6 *Suppose that the hypothesis of Proposition 2.5 holds true and $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ is a measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$. Then ν is AMS and $\rho \gg \bar{\nu}, \bar{\mu} \circ (f^{\mathbb{Z}})^{-1} \gg \nu$.*

Proof: Observe that $\rho(A) \geq L^{-1} \bar{\mu}((f^{\mathbb{Z}})^{-1}A)$. Hence $\rho \gg \bar{\mu} \circ (f^{\mathbb{Z}})^{-1}$. Since $\bar{\mu} \circ (f^{\mathbb{Z}})^{-1} \gg \mu \circ (f^{\mathbb{Z}})^{-1} = \nu$ follows from $\bar{\mu} \gg \mu$, we obtain $\rho \gg \nu$. Measure ρ is stationary by Proposition 2.5, so ν is AMS, and $\rho \gg \bar{\nu}$ by Lemma 2.1. \square

The next proposition states that ρ is the stationary mean of the transported measure if the expansion rate is almost surely constant.

Proposition 2.7 *Suppose that the hypothesis of Proposition 2.5 holds true and $\bar{l}(\cdot) = L$ μ -almost everywhere. Then $\rho = \bar{\nu}$ for $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$.*

Proof: By stationarity of $\bar{\mu}$, identity (22), and the dominated convergence,

$$\rho(A) = L^{-1} \int \left[\lim_n n^{-1} \sum_{i=0}^{n-1} G(A, T^i x^{\mathbb{Z}}) \right] d\bar{\mu}(x^{\mathbb{Z}}).$$

This expression equals $\bar{\nu}(A)$ by Proposition 2.3 if $\bar{l}(\cdot) = L$ almost surely. \square

Example 2.8 *The equality $\bar{l}(\cdot) = L$ holds almost everywhere for the nonergodic process (11) and $f : \mathbb{N} \times \{0, 1\} \rightarrow \{0, 1\}^*$ if $f(k, z) = g(k)w(z)$ for $k \in \mathbb{N}$, $z \in \{0, 1\}$, $|g(k)| = O(\log k)$, and $|w(z)| = A$. Code (14) falls under that case.*

3 Synchronization and shift-invariant σ -fields

For an injection $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$, the transported shift

$$T_{\pi} := \pi \circ T \circ \pi^{-1}, \quad (27)$$

considered in [17, Example 6], constitutes an injection $\pi(\mathbb{X}^{\mathbb{Z}}) \rightarrow \pi(\mathbb{X}^{\mathbb{Z}})$. In that case, $\nu := \mu \circ \pi^{-1}$ is stationary with respect to T_{π} for a stationary measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$, i.e., $\nu \circ T_{\pi}^{-1} = \nu$.

We will demonstrate that ergodic properties of measures μ and $\nu = \mu \circ \pi^{-1}$ may be further related in the more specific case of a synchronizable injection. Some apparent technical difficulty is that the set $\pi(\mathbb{X}^{\mathbb{Z}})$ usually does not belong to the T -invariant algebra $\mathcal{I}_{\mathbb{Y}} := \{A \in \mathcal{Y}^{\mathbb{Z}} : T^{-1}A = A\}$. However, this can be overcome easily given certain care.

Lemma 3.1 *For an injection $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$, consider pseudo-invariant algebras*

$$\begin{aligned} \mathcal{Q} &:= \{A \in \mathcal{Y}^{\mathbb{Z}} : A = B \cap \pi(\mathbb{X}^{\mathbb{Z}}), T^{-1}B = B\}, \\ \mathcal{Q}_{\pi} &:= \{A \in \mathcal{Y}^{\mathbb{Z}} : A = B \cap \pi(\mathbb{X}^{\mathbb{Z}}), T_{\pi}^{-1}B = B\}, \end{aligned}$$

where T_{π} is defined by (27). We have

$$\mathcal{Q} \subset \mathcal{Q}_{\pi} = \pi(\mathcal{I}_{\mathbb{X}}).$$

Proof: The right equality is obvious. As for the left relation, observe that $\bigcup_{i \in \mathbb{Z}} T^i B \supset \bigcup_{i \in \mathbb{Z}} T_{\pi}^i B \supset B$. If $T^{-1}B = B$ then $\bigcup_{i \in \mathbb{Z}} T^i B = B$. Hence $B \cap \pi(\mathbb{X}^{\mathbb{Z}}) \in \mathcal{Q}_{\pi}$ since formula $\bigcup_{i \in \mathbb{Z}} T_{\pi}^i B$ defines a T_{π} -invariant set. \square

Proposition 3.2 *For a synchronizable injection $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$,*

$$\mathcal{Q} = \mathcal{Q}_{\pi}.$$

Proof: By Lemma 3.1, $\mathcal{Q} \subset \mathcal{Q}_{\pi}$. Thus it suffices to show that $\mathcal{Q}_{\pi} \subset \mathcal{Q}$ or, equivalently, that $\mathcal{I}_{\mathbb{X}} \subset \pi^{-1}(\mathcal{Q})$. We will demonstrate the latter. Consider an $A \in \mathcal{I}_{\mathbb{X}}$ and construct the set $E = \pi(\mathbb{X}^{\mathbb{Z}}) \cap \bigcup_{i \in \mathbb{Z}} T^i \pi(A) \in \mathcal{Q}$. Since π is synchronizable and A is T -invariant, we have that $\pi^{-1}(E) = A$. \square

Proposition 3.3 *Consider a synchronizable injection $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$, a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$, and its image $\nu = \mu \circ \pi^{-1}$ on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$. For each $E \in \mathcal{I}_{\mathbb{Y}}$, there exists such an $A \in \mathcal{I}_{\mathbb{X}}$, and for each $A \in \mathcal{I}_{\mathbb{X}}$, there exists an $E \in \mathcal{I}_{\mathbb{Y}}$ such that*

$$\nu(E) = \nu(E \cap \pi(\mathbb{X}^{\mathbb{Z}})) = \mu(A). \quad (28)$$

Remark: As a further corollary, either both measures μ and ν are ergodic or neither of them has this property. Let us recall that $\bar{\mu}(A) = \mu(A)$ for $A \in \mathcal{I}_{\mathbb{X}}$ and an AMS μ . The analogical equality $\bar{\nu}(E) = \nu(E)$ holds for $E \in \mathcal{I}_{\mathbb{Y}}$ and an AMS ν . Some oddness of (28) is buried in the fact that $\bar{\nu}(E)$ does not necessarily equal $\bar{\nu}(E \cap \pi(\mathbb{X}^{\mathbb{Z}}))$. It is only the support of ν that is confined to $\pi(\mathbb{X}^{\mathbb{Z}})$, and $\pi(\mathbb{X}^{\mathbb{Z}})$ need not be T -invariant, as it has been remarked.

Proof: For $E \in \mathcal{I}_{\mathbb{Y}}$, take $A = \pi^{-1}(E)$. For $A \in \mathcal{I}_{\mathbb{X}}$, take $E = \bigcup_{i \in \mathbb{Z}} T^i \pi(A)$. Then the equality follows immediately from Proposition 3.2. \square

Example 3.4 Process (11) has a nonatomic shift-invariant sub- σ -field [13]. Hence the expanded process $(Y_i)_{i \in \mathbb{Z}} = f^{\mathbb{Z}}((X_i)_{i \in \mathbb{Z}})$ distributed $P((Y_i)_{i \in \mathbb{Z}} \in \cdot) = \nu$ and its stationary mean $(\bar{Y}_i)_{i \in \mathbb{Z}}$ distributed $\bar{P}((\bar{Y}_i)_{i \in \mathbb{Z}} \in \cdot) = \bar{\nu}$ have the same property if we use a comma-separated code $f : \mathbb{X} \rightarrow \mathbb{Y}^*$, like (14). Moreover, $(\bar{Y}_i)_{i \in \mathbb{Z}}$ has a nonatomic shift-invariant sub- σ -field if and only if (10) holds true for certain functions \bar{s}_k and independent equidistributed binary random variables $(\bar{Z}_k)_{k \in \mathbb{N}}$ [13, Theorem 9].

4 Complete fix-free codes and stationarity

This section contains a result of independent interest, loosely related to the setting of our initial problem. For any injection $\pi : \mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$, measures may also be transported in the opposite direction. That is, for any measure ν on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$, there exists a measure $\mu := \nu \circ \pi$ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$. A condition opposite to synchronization appears when this mapping is required to preserve stationarity. We came across the following proposition, which seemingly has not been noticed so far, cf. [4]:

Proposition 4.1 *Suppose that \mathbb{X} is finite and $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is complete fix-free. Then*

- (i) $f^{\mathbb{Z}}$ is a bijection $\mathbb{X}^{\mathbb{Z}} \rightarrow \mathbb{Y}^{\mathbb{Z}}$, and
- (ii) $\mu = \nu \circ f^{\mathbb{Z}}$ is a stationary measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ if ν is a stationary measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$.

Remark: Statement (i) may be false for a complete infinite prefix-free set $f(\mathbb{X})$. For instance, the set $(f^{\mathbb{Z}})^{-1}(\{\dots 00.00\dots\})$ is empty for $f(n) = \{0^{n-1}1 : n \in \mathbb{N}\}$, $\mathbb{X} = \mathbb{N}$, and $\mathbb{Y} = \{0, 1\}$. However, we do not know of any complete infinite set of strings that would be both prefix- and suffix-free, cf. [16, 1].

Proof: Let $\mathcal{L} = f(\mathbb{X})$.

(i) Clearly, $|w| \geq 1$ for $w \in \mathcal{L}$ if f is a bijection. Thus $f^{\mathbb{Z}}(x^{\mathbb{Z}})$ is a two-sided sequence for $x^{\mathbb{Z}} \in \mathbb{X}^{\mathbb{Z}}$. Moreover, given a $y^{\mathbb{Z}} \in \mathbb{Y}^{\mathbb{Z}}$, we can reconstruct the unique $x^{\mathbb{Z}} \in \mathbb{X}^{\mathbb{Z}}$ with $f^{\mathbb{Z}}(x^{\mathbb{Z}}) = y^{\mathbb{Z}}$ by cutting off the consecutive suffixes or prefixes belonging to \mathcal{L} from $y_{-\infty}^0$ and y_1^{∞} . By the following reasoning, this parsing process is guaranteed not to stop after a finite number of steps.

On the contrary, assume that there is such an infinite sequence y_k^{∞} (the mirrorlike argumentation applies to $y_{-\infty}^k$) such that no $w \in \mathcal{L}$ is a prefix of y_k^{∞} . Let v be a prefix of y_k^{∞} that is longer than any $w \in \mathcal{L}$. The set $\{v\} \cup \mathcal{L}$ is prefix-free, and so $\sum_{w \in \mathcal{L}} |\mathbb{Y}|^{-|w|} \leq 1 - |\mathbb{Y}|^{-|v|} < 1$ by the Kraft inequality. We have arrived at a contradiction, so the assumption was false.

(ii) By the Kolmogorov process theorem and the π - λ theorem, stationarity of μ is equivalent to the set of equalities

$$\sum_{z \in \mathcal{L}} \nu([wz]) = \nu([w]) = \sum_{z \in \mathcal{L}} \nu([zw]), \quad w \in \mathcal{L}^*. \quad (29)$$

On the other hand, stationarity of ν is equivalent to

$$\sum_{s \in \mathbb{Y}} \nu([ws]) = \nu([w]) = \sum_{s \in \mathbb{Y}} \nu([sw]), \quad w \in \mathbb{Y}^*. \quad (30)$$

The following auxiliary fact is useful to derive (29) from (30): Let $l(\mathcal{M}) \geq 1$ be the length of the longest string in a set \mathcal{M} . Any finite complete prefix-free set $\mathcal{M} \subset \mathbb{Y}^*$ may be decomposed as $\mathcal{M} = \mathcal{M}_r \cup (\mathcal{M}_p \times \mathbb{Y})$, where $l(\mathcal{M}) = l(\mathcal{M}_p \times \mathbb{Y}) > l(\mathcal{M}_r)$, and $\mathcal{M}_m = \mathcal{M}_r \cup \mathcal{M}_p$ is a complete prefix-free set. This decomposition may be proved by contradiction with the Kraft inequality applied to \mathcal{M}_m .

From this and from (30), it follows that for any complete prefix-free \mathcal{M} with $l(\mathcal{M}) \geq 1$, there exists a complete prefix-free \mathcal{M}_m such that $l(\mathcal{M}) = l(\mathcal{M}_m) - 1$ and

$$\sum_{z \in \mathcal{M}} \nu([wz]) = \sum_{z \in \mathcal{M}_m} \nu([wz]).$$

Using this, the left equality in (29) may be proved by induction on $l(\mathcal{M})$ starting with $\mathcal{M}_m = \{\lambda\}$. The proof of the right equality is mirrorlike. \square

Corollary 4.2 *Suppose that \mathbb{X} is finite, $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is complete fix-free, and ν is AMS on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$. Then $\mu = \nu \circ f^{\mathbb{Z}}$ is AMS on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ with the stationary mean $\bar{\mu} \ll \bar{\nu} \circ f^{\mathbb{Z}}$.*

Proof: We have $\mu = \nu \circ f^{\mathbb{Z}} \ll \bar{\nu} \circ f^{\mathbb{Z}}$, where the last measure is stationary by Proposition 4.1. Hence the claim follows by Lemma 2.1. \square

5 Preservation of the finite-energy property

We supposed that both $f^{\mathbb{Z}}$ and $(f^{\mathbb{Z}})^{-1}$ preserve the finite-energy property if the coding function f is sufficiently nice, prefix-free in particular. The proofs are a bit more complicated than we expected, but convenient sufficient conditions can be formulated.

Definition 5.1 *More specifically, we will say that (i) a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ has (K, c) -energy if $c < 1$, $K < \infty$, and condition (7) holds and (ii) the measure μ has (K, c, f) -energy for a coding function $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ if $c < 1$, $K < \infty$, and*

$$\mu([uv]) \leq Kc^{|f^{*(v)}|} \mu([u]). \quad (31)$$

Remark: If a function $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is prefix-free, then, by the Kraft inequality $\sum_{x \in \mathbb{X}} |\mathbb{Y}|^{-|f(x)|} \leq 1$, condition (31) may be only satisfied for $c \geq |\mathbb{Y}|^{-1}$. In particular, the inequality $c > |\mathbb{Y}|^{-1}$ must be strict for a noncomplete coding function, i.e., when $\sum_{x \in \mathbb{X}} |\mathbb{Y}|^{-|f(x)|} < 1$.

Proposition 5.2 *If $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is prefix-free and a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ has (K, c, f) -energy, then μ has also (K, c) -energy.*

Proof: If f is prefix-free, then $|f^*(u)| \geq |u|$. Hence,

$$\mu([vu]) \leq Kc^{|f^*(u)|} \mu([v]) \leq Kc^{|u|} \mu([v]).$$

\square

Proposition 5.3 *If $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is prefix-free and a measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ has (K, c) -energy, then the measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ has (K, c, f) -energy.*

Proof: Let $\nu([zw]) \leq Kc^{|w|}\nu([z])$. The extension f^* is an injection for a prefix-free f . Moreover, if both z and w belong to $f^*(\mathbb{X}^*)$, then

$$(f^{\mathbb{Z}})^{-1}([zw]) = [(f^*)^{-1}(z)(f^*)^{-1}(w)].$$

In particular, for $z = f^*(u)$ and $w = f^*(v)$, we obtain

$$\mu([uv]) = \nu([zw]) \leq Kc^{|w|}\nu([z]) = Kc^{|f^*(v)|}\mu([u]).$$

□

The converse of Proposition 5.3 is valid under additional restrictions. Denote the difference of strings $w, z \in \mathbb{Y}^*$ as

$$w - z := \begin{cases} s & \text{if } w = zs, s \in \mathbb{Y}^+, \\ \lambda & \text{if } z = ws, s \in \mathbb{Y}^*, \\ w & \text{else.} \end{cases}$$

We choose this definition to have $p^{|w-z|} \leq p^{|w|-|z|}$ for $p \in (0, 1)$. The set of z 's that fall under the first two cases is denoted as

$$\langle w \rangle := \{z \in \mathbb{Y}^* : \exists s \in \mathbb{Y}^* : w = zs \vee z = ws\}.$$

Moreover, for a set $\mathcal{L} \subset \mathbb{Y}^*$, we define the remainder $w_{\mathcal{L}}$ of a string $w \in \mathbb{Y}^*$ as the shortest element of the set $\{w - s : s \in \mathcal{L}^*, |s| \leq |w|\}$. The completion set of the string w with respect to the set \mathcal{L} is defined as

$$\mathcal{L}_w := \{s \in \mathbb{Y}^* : w_{\mathcal{L}}s \in \mathcal{L}\}.$$

This set is prefix-free for a prefix-free \mathcal{L} .

Proposition 5.4 *Suppose that $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is prefix-free and a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ has (K, c, f) -energy. For $\mathcal{L} = f(\mathbb{X})$, put*

$$M_f(p) := \sup_{w \in \mathbb{Y}^* : \mathcal{L}_w \neq \emptyset} \sum_{s \in \mathcal{L}_w} p^{|s|},$$

$$N_{f,\mu}(p) := \sup_{w \in \mathbb{Y}^* : \mathcal{L}_w \neq \emptyset} \frac{\sum_{s \in \mathcal{L}_w} p^{-|s|} \mu([(f^*)^{-1}(ws)])}{\sum_{s \in \mathcal{L}_w} \mu([(f^*)^{-1}(ws)])}.$$

If $M_f(c) < \infty$ and $N_{f,\mu}(c_2) < \infty$ for a certain $c_2 \in [c, 1)$, then the measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ has (\tilde{K}, c_2) -energy, where $\tilde{K} = N_{f,\mu}(c_2)M_f(c)K$.

Proof: If $z = f^*(u)$ and $w = f^*(v)$ for some u and v , then

$$\nu([zw]) = \mu([uv]) \leq Kc^{|f^*(v)|}\mu([u]) = Kc^{|w|}\nu([z]).$$

Notice that $\nu([z]) = \sum_{s \in \mathcal{L}_z} \nu([zs])$ for any $z \in \mathbb{Y}^*$. Assume now that $z \in \mathcal{L}^*$ and let w be arbitrary. By $\mathcal{L}_{zw} = \mathcal{L}_w$ we obtain

$$\nu([zw]) = \sum_{s \in \mathcal{L}_w} \nu([zws]) \leq \sum_{s \in \mathcal{L}_w} Kc^{|ws|}\nu([z]) \leq M_f(c)Kc^{|w|}\nu([z]).$$

Eventually, consider arbitrary z and w . We have

$$\begin{aligned}\nu([zw]) &= \sum_{s \in \mathcal{L}_z \cap (w)} \nu([zs(w-s)]) \leq \sum_{s \in \mathcal{L}_z \cap (w)} M_f(c) K c^{|w-s|} \nu([zs]) \\ &\leq \sum_{s \in \mathcal{L}_z} M_f(c) K c_2^{|w-s|} \nu([zs]) \leq M_f(c) K c_2^{|w|} \sum_{s \in \mathcal{L}_z} c_2^{-|s|} \nu([zs]) \\ &\leq \tilde{K} c_2^{|w|} \nu([z]).\end{aligned}$$

□

Corollary 5.5 *If $f : \mathbb{X} \rightarrow \mathbb{Y}^*$ is finite prefix-free and a measure μ on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ has finite energy, then the measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ has finite energy.*

Proof: Assume that μ has (K, q) -energy. We have $|\mathcal{L}_w| \leq |\mathcal{L}| < \infty$ and $\sup_{z \in \mathcal{L}_w} |z| \leq \sup_{z \in \mathcal{L}} |z| < \infty$. Hence we obtain inequalities $M_f(p) < \infty$, $N_{f, \mu}(p) < \infty$, and (31) for $p < 1$ and $c \in [\max_{z \in \mathcal{L}} q^{1/|z|}, 1)$. In consequence, the claim follows by Proposition 5.4. □

Below we present a more specific example with an infinite image $f(\mathbb{X})$.

Corollary 5.6 *Let $f : \mathbb{X} = \mathbb{N} \mapsto \mathbb{Y}^+ = \{0, 1, 2\}^+$ be given as*

$$f(k) = b(k)w(k)2,$$

where $1b(k) \in \{0, 1\}^+$ is the binary representation of a natural number k , and $w(k) \in \{0, 1\}^*$ is a string of fixed length, $|w(k)| = A$. Let also μ be a measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ that satisfies $\mu([vu]) = \mu([v])\mu([u])$ and

$$\mu([k]) = \frac{k^{-\alpha}}{\zeta(\alpha)}, \quad k \in \mathbb{N},$$

for some $\alpha > 1$. If $\zeta(\alpha) > 2^{A+1}$, then μ has $(1, c, f)$ -energy for $c \in [\max\{2^{-\alpha}, (\zeta(\alpha))^{-1/(A+1)}\}, 2^{-1})$, whereas $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ has (\tilde{K}, c_2) -energy for $\tilde{K} = N_{f, \mu}(c_2)M_f(c)$ and $c_2 \in (\max\{c, 2^{1-\alpha}\}, 1)$.

Proof: We have $|f(k)| = \lfloor \log_2 k \rfloor + 1 + |w(k)|$ and $k^{-\alpha} = (2^{-\alpha})^{\log_2 k}$. Thus,

$$(\zeta(\alpha))^{-|u|} (2^{-\alpha})^{|f^*(u)| - A|u|} \leq \mu([u]) \leq (\zeta(\alpha))^{-|u|} (2^{-\alpha})^{|f^*(u)| - (A+1)|u|} \leq c^{|f^*(u)|}.$$

In particular, (31) follows for $K = 1$. Consider a string $w \in \mathbb{Y}^*$ and let a_l be the number of strings of length l in the set \mathcal{L}_w . We can see that $a_l \leq 1$ for $l \leq A + 1$, whereas $a_l = 2^{l-(A+1)}$ for $l > A + 1$ if \mathcal{L}_w is not empty. Hence, $M_f(c) \leq \sum_{l=0}^{\infty} \max\{1, 2^{l-(A+1)}\} c^l < \infty$ and

$$N_{f, \mu}(c_2) \leq \frac{\sum_{l=0}^{\infty} \max\{1, 2^{l-(A+1)}\} c_2^{-l} (2^{-\alpha})^{l-1}}{\sum_{l=A+1}^{\infty} 2^{l-(A+1)} (2^{-\alpha})^l} < \infty.$$

So the claim holds by Proposition 5.4. □

By means of the following two simple statements, the above result can be extended to certain nonergodic measures, including the distribution of process (11) and its stationary variable-length coding.

Proposition 5.7 Consider a measure P on (Ω, \mathcal{J}) and a probability kernel τ from (Ω, \mathcal{J}) to $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ (i.e., $\tau(\cdot, \omega)$ is a measure on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ for P -almost all $\omega \in \Omega$, and the function $\tau(A, \cdot)$ is measurable \mathcal{J} for each $A \in \mathcal{Y}^{\mathbb{Z}}$). If $\tau(\cdot, \omega)$ has (K, c) -energy for P -almost all $\omega \in \Omega$, then so does the measure $\int \tau(\cdot, \omega) dP(\omega)$.

Proof:

$$\int \tau([zw], \omega) dP(\omega) \leq \int Kc^{|w|} \tau([w], \omega) dP(\omega) \leq Kc^{|w|} \int \tau([w], \omega) dP(\omega).$$

□

Proposition 5.8 If an AMS measure ν on $(\mathbb{Y}^{\mathbb{Z}}, \mathcal{Y}^{\mathbb{Z}})$ has (K, c) -energy, then so does the measure $\bar{\nu}$.

Proof:

$$\begin{aligned} \bar{\mu}([zw]) &= \lim_n n^{-1} \sum_{i=0}^{n-1} \sum_{s \in \mathbb{Y}^i} \mu([szw]) \\ &\leq Kc^{|w|} \lim_n n^{-1} \sum_{i=0}^{n-1} \sum_{s \in \mathbb{Y}^i} \mu([sz]) \leq Kc^{|w|} \bar{\mu}([z]). \end{aligned}$$

□

6 Block entropies of stationary means

For a stationary measure μ , block entropy $H_\mu(n) = H_\mu(0; n)$ defined in (8) is a nonnegative, growing, and concave function of n , see [11]. Hence the limit

$$h_\mu := \lim_{n \rightarrow \infty} \frac{H_\mu(n)}{n}, \quad (32)$$

known as the entropy rate, exists in that case. Whereas block entropy behaves less regularly in a general AMS case, we can bound the block entropy of the stationary mean in the following way.

Proposition 6.1 For an AMS measure μ ,

$$H_{\bar{\mu}}(m) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} H_\mu(i; m). \quad (33)$$

Proof: The claim

$$-\sum_{u \in \mathbb{X}^m} \bar{\mu}([u]) \log \bar{\mu}([u]) \geq \limsup_{n \rightarrow \infty} \left[-\frac{1}{n} \sum_{i=0}^{n-1} \sum_{u \in \mathbb{X}^m} \mu(T^{-i}[u]) \log \mu(T^{-i}[u]) \right]$$

follows by the Jensen inequality for the function $p \mapsto -p \log p$. □

Proposition 6.2 Let μ be AMS with $h_{\bar{\mu}} < \infty$ and $H_\mu(n) < \infty$ for all n . Then

$$h_{\bar{\mu}} \leq \liminf_{n \rightarrow \infty} \frac{H_\mu(n)}{n}. \quad (34)$$

Proof: By the generalized Shannon-McMillan-Breiman theorem [5, Theorem 2] and (16), the $L^1(\bar{\mu})$ convergence

$$h_{\bar{\mu}} = \int \lim_{n \rightarrow \infty} \frac{\log \bar{\mu}([x^n])}{n} d\bar{\mu}(x^{\mathbb{Z}}) = \int \lim_{n \rightarrow \infty} \frac{\log \bar{\mu}([x^n])}{n} d\mu(x^{\mathbb{Z}}) \quad (35)$$

holds for the stationary measure $\bar{\mu}$ if $h_{\bar{\mu}} < \infty$. On the other hand, by [5, Theorem 3], an AMS measure μ with $H_{\mu}(n) < \infty$ satisfies

$$\lim_{n \rightarrow \infty} \frac{\log \bar{\mu}([x^n])}{n} = \lim_{n \rightarrow \infty} \frac{\log \mu([x^n])}{n} \quad (36)$$

for μ - and $\bar{\mu}$ -almost all $x^{\mathbb{Z}}$. Hence the claim follows by the Fatou lemma. \square

Example 6.3 *Using Proposition 6.2, we can show a simple example of a measure μ such that limits (6) exist for all cylinder sets but μ is not AMS. Consider the Champernowne sequence $b^{\mathbb{N}} = 12345678910111213\dots$ (i.e., the concatenation of decimal representations of natural numbers) and put $\mu([b^k]) = 1$ for $k \in \mathbb{N}$. We have $\bar{\mu}([u]) = 10^{-|u|}$ since $b^{\mathbb{N}}$ is normal. If $\bar{\mu}$ is extended from these values to a measure on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$, then $H_{\bar{\mu}}(m) = m \log 10$ and $h_{\bar{\mu}} = \log 10$ but $H_{\mu}(i; m) = 0$. Hence μ cannot be AMS.*

Block entropies of two stationary means linked through variable-length coding can be related as well. The link for the entropy rate is very simple if the expansion rate is constant and the coding function is uniquely decodable.

Proposition 6.4 (cf. [27, Theorem 1]) *Let μ be AMS on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ with $h_{\bar{\mu}} < \infty$ and $H_{\mu}(n) < \infty$ for all n and suppose that the expansion rate satisfies $\bar{l}(\cdot) = L \in (0, \infty)$ μ -almost everywhere for a prefix-free $f : \mathbb{X} \rightarrow \mathbb{Y}^*$. Then we have*

$$h_{\bar{\nu}} = L^{-1} h_{\bar{\mu}} \quad (37)$$

for the measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ if $h_{\bar{\nu}} < \infty$.

Remark: We have $h_{\bar{\mu}} < \infty$ and $H_{\mu}(n) < \infty$ if μ is stationary and $H_{\mu}(1) < \infty$, whereas $h_{\bar{\nu}} < \infty$ if the alphabet \mathbb{Y} is finite. Formula (37) is a special case of [27, Theorem 1], but their proof is partly flawed. It uses the version of the Shannon-McMillan-Breiman theorem [17, Corollary 4], which is true only for finite \mathbb{X} and \mathbb{Y} . A correct proof for any \mathbb{X} and \mathbb{Y} , invoking the already mentioned Theorems 2 and 3 from [5], is given below. As noticed in [27, Theorem 1], $\lim_n n^{-1} \log \mu([x^n])$ converges almost surely to the entropy rate of the $x^{\mathbb{Z}}$ -typical ergodic component of the measure μ also when the expansion rate is not constant, see the ergodic decomposition theorems in [19, Chapter 9 until Theorem 9.12].

Proof: The measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$ is AMS by Proposition 2.3, whereas the extension f^* is an injection for a prefix-free f . Hence $\mu([x^n]) = \nu([f^*(x^n)])$ and $|f^*(x)| \geq |x|$. As a result, $H_{\nu}(n) \leq H_{\mu}(n)$, whereas (35), (36), and their analogues for ν imply

$$\begin{aligned} h_{\bar{\nu}} &= \int \left[\lim_{n \rightarrow \infty} n^{-1} \log \nu([y^n]) \right] d\nu(y^{\mathbb{Z}}) \\ &= \int \left[\lim_{n \rightarrow \infty} |f^*(x^n)|^{-1} \log \mu([x^n]) \right] d\mu(x^{\mathbb{Z}}) \\ &= \int [l(x^{\mathbb{Z}})]^{-1} \left[\lim_{n \rightarrow \infty} n^{-1} \log \mu([x^n]) \right] d\mu(x^{\mathbb{Z}}) = L^{-1} h_{\bar{\mu}}. \end{aligned}$$

□

In the following, we wish to obtain some bounds for $H_{\bar{\nu}}(n)$ in terms of $H_{\bar{\mu}}(n)$ for finite n . We shall observe that formula (26) can be interpreted in terms of random variables if $|f(x)| > 0$ for all $x \in \mathbb{X}$. To simplify the notation, we shall assume here that $\bar{\mu} = \mu := P((X_i)_{i \in \mathbb{Z}} \in \cdot)$ is the measure of a stationary process $(X_i)_{i \in \mathbb{Z}}$. Then $\rho = \bar{P}((\bar{Y}_i)_{i \in \mathbb{Z}} \in \cdot)$ is the stationary distribution of

$$(\bar{Y}_i)_{i \in \mathbb{Z}} = T^N f^{\mathbb{Z}}((\bar{X}_i)_{i \in \mathbb{Z}}), \quad (38)$$

where the random shift N and the nonstationary process $(\bar{X}_i)_{i \in \mathbb{Z}}$ are conditionally independent given \bar{X}_1 , their distribution being

$$\bar{P}(\bar{X}_k^l = x_k^l) = P(X_k^l = x_k^l) \cdot \frac{|f(x_1)|}{L}, \quad k \leq 1 \leq l, \quad (39)$$

$$\bar{P}(N = n | \bar{X}_1 = x_1) = \frac{\mathbf{1}_{\{0 \leq n \leq |f(x_1)|-1\}}}{|f(x_1)|}, \quad n \in \mathbb{N} \cup \{0\}. \quad (40)$$

Suppose that $\bar{l}(\cdot) = L$ holds μ -almost surely. As shown in Proposition 2.7, this guarantees that $\rho = \bar{\nu}$ for the AMS measure $\nu := P((Y_i)_{i \in \mathbb{Z}} \in \cdot)$ of the expanded process (4). Thus we have

$$H_{\bar{\nu}}(n) = H_{\bar{P}}(\bar{Y}_k^{k+n-1}) \text{ and } H_{\bar{\mu}}(n) = H_P(X_k^{k+n-1}), \quad (41)$$

where $H_P(U) := \mathbf{E}_P[-\log P(U = \cdot)]$ is the entropy of a discrete variable U .

Denote the conditional entropy $H_P(U|V) := H_P(U, V) - H_P(V)$ and covariance $\text{Cov}_P(U, V) := \mathbf{E}_P(UV) - \mathbf{E}_P U \mathbf{E}_P V$. Entropies of blocks drawn from the above introduced processes can be linked easily when blocks of random length are allowed.

Proposition 6.5 *Suppose that a process $(X_i)_{i \in \mathbb{Z}}$ is stationary and $L = \mathbf{E}_P |f(X_i)| < \infty$ for a prefix-free $f : \mathbb{X} \rightarrow \mathbb{Y}^*$. Consider then processes $(Y_i)_{i \in \mathbb{Z}}$, $(\bar{X}_i)_{i \in \mathbb{Z}}$, and $(\bar{Y}_i)_{i \in \mathbb{Z}}$ that satisfy (4), (38), (39), and (40). Put also $M_n := \sum_{i=1}^n |f(X_i)|$, $\bar{M}_n := \sum_{i=1}^n |f(\bar{X}_i)|$, and $\eta := \mathbf{E}_P \left[\frac{|f(X_1)|}{L} \log \frac{|f(X_1)|}{L} \right] \geq 0$. Then we have*

$$(i) \quad H_P(Y^{M_n}) = H_P(X^n),$$

$$(ii) \quad H_{\bar{P}}(\bar{X}_k^l) \geq H_P(X_k^l) - \eta \text{ if and only if } \text{Cov}_P(|f(X_1)|, -\log P(X_k^l = \cdot)) \geq 0 \text{ for } k \leq 1 \leq l,$$

$$(iii) \quad H_{\bar{P}}(N | \bar{X}_k^l) = \log L + \eta, \quad k \leq 1 \leq l, \text{ whereas}$$

$$(iv) \quad H_{\bar{P}}(\bar{Y}^{\bar{M}_n - \bar{M}_1}, N) \leq H_{\bar{P}}(\bar{X}^n, N) \text{ and } H_{\bar{P}}(\bar{X}_2^n, N) \leq H_{\bar{P}}(\bar{Y}^{\bar{M}_n}, N).$$

Proof: In (i) and (iv), we use that f^* is an injection for a prefix-free f .

$$(i) \quad \text{The claim is true since } Y^{M_n} = f^*(X^n) \text{ and } X^n = (f^*)^{-1}(Y^{M_n}).$$

$$(ii) \quad \text{Whenever } \text{Cov}_P(|f(X_1)|, -\log P(X_k^l = \cdot)) \geq 0, \text{ we observe}$$

$$\begin{aligned} H_{\bar{P}}(\bar{X}_k^l) + \eta &= \mathbf{E}_P \left[-\frac{|f(X_1)|}{L} \log P(X_k^l = \cdot) \right] \\ &\geq \mathbf{E}_P \left[\frac{|f(X_1)|}{L} \right] \mathbf{E}_P [-\log P(X_k^l = \cdot)] = H_P(X_k^l). \end{aligned}$$

(iii) By (40).

(iv) By (38), the string $\bar{Y}^{\bar{M}_n - \bar{M}_1}$ is a function of \bar{X}^n and N , whereas string \bar{X}_2^n is a function of $\bar{Y}^{\bar{M}_n}$ and N . Hence the claimed inequalities follow.

□

Corollary 6.6 *Let μ be AMS on $(\mathbb{X}^{\mathbb{Z}}, \mathcal{X}^{\mathbb{Z}})$ and $f : \mathbb{X} \rightarrow \mathbb{Y}^L$ be a prefix-free fixed-length coding. Then we have*

$$|H_{\bar{\nu}}(nL) - H_{\bar{\mu}}(n)| \leq H_{\bar{\mu}}(2) + \log L \quad (42)$$

for the measure $\nu = \mu \circ (f^{\mathbb{Z}})^{-1}$.

Proof: In view of Corollary 2.4, we may assume without loss of generality that μ is stationary. Let us repeat the construction of processes that precedes Proposition 6.5. Observe that the processes \bar{X}_k^l and X_k^l share the same distribution and $\bar{M}_n = nL$. Thus Proposition 6.5(iv) yields $H_{\bar{P}}(\bar{Y}^{(n-1)L}) \leq H_{\bar{P}}(\bar{X}^n, N) \leq H_P(X^{n-1}) + H_{\bar{P}}(\bar{X}^2, N)$ and $H_P(X^{n-1}) = H_{\bar{P}}(\bar{X}_2^n) \leq H_{\bar{P}}(\bar{Y}^{Ln}, N) \leq H_{\bar{P}}(\bar{Y}^{L(n-1)}) + H_{\bar{P}}(\bar{Y}^L, N) \leq H_{\bar{P}}(\bar{Y}^{L(n-1)}) + H_{\bar{P}}(\bar{X}^2, N)$. To complete the proof, notice that $H_{\bar{P}}(\bar{X}^2, N) = H_P(X^2) + \log L$. □

7 Encoding of the process $X_i = (K_i, Z_{K_i})$

We have not managed to produce an analogue of Corollary 6.6 for the processes discussed in Proposition 1.4. But, as shown in [12], then we have

$$H_P(X^n) \geq h_{\mu}n + [\log 2 - \eta(\delta)] \cdot |U_{\delta}(n)|, \quad (43)$$

$$H_{\bar{P}}(\bar{Y}^m) \geq h_{\bar{\nu}}m + [\log 2 - \eta(\bar{\delta})] \cdot |\bar{U}_{\bar{\delta}}(m)|, \quad (44)$$

where $\nu := \mu \circ (f^{\mathbb{Z}})^{-1}$ and $\eta(p) := -p \log p - (1-p) \log(1-p)$. Whereas $h_{\bar{\nu}} = L^{-1}h_{\mu}$ by Proposition 6.4, point (d) of the proof below demonstrates that $\bar{U}_{\bar{\delta}}(m) \supset U_{\delta}(n)$ for $\delta > \bar{\delta}/a$, $n = \lfloor (\delta - \bar{\delta}/a)L^{-1}(m - C_a) \rfloor$, $a \in (\bar{\delta}, 1)$, and a certain constant C_a .

Proof of the Proposition 1.4:

(a)-(b) Process $(|f(X_i)|)_{i \in \mathbb{Z}}$ is ergodic. Thus the expansion rate equals its expectation almost surely:

$$\bar{l}((X_i)_{i \in \mathbb{Z}}) = L = \mathbf{E}_P |f(X_i)| = \sum_{k=1}^{\infty} ([\log_2 k] + 2) \frac{k^{-\alpha}}{\zeta(\alpha)} \in (0, \infty).$$

Hence the stationary mean $\overline{\mu \circ (f^{\mathbb{Z}})^{-1}}$ exists by Proposition 2.3 and constitutes a measure over a finite alphabet.

(c) Consider the process $(Y_i)_{i \in \mathbb{Z}} = f^{\mathbb{Z}}((X_i)_{i \in \mathbb{Z}})$ and the probability kernel $\tau(\cdot, \omega) = P((Y_i)_{i \in \mathbb{Z}} \in \cdot | (Z_k)_{k \in \mathbb{N}})(\omega)$. For $\zeta(\alpha) > 4$ and P -almost all ω , $\tau(\cdot, \omega)$ takes form of the measure ν considered in Corollary 5.6. Hence the distribution of the process $(Y_i)_{i \in \mathbb{Z}}$ has finite energy by Proposition 5.7 and, consequently, $(\bar{Y}_i)_{i \in \mathbb{Z}}$ has finite energy by Proposition 5.8.

(d) Define $\bar{Z}_k := \bar{s}_k((\bar{Y}_i)_{i \in \mathbb{Z}})$. The functions \bar{s}_k are shift-invariant. Notice that $Z_k = \bar{s}_k((Y_i)_{i \in \mathbb{Z}})$ almost surely on the space (Ω, \mathcal{J}, P) . Hence, by Eq. (16) applied to the AMS measure $\mu \circ (f^{\mathbb{Z}})^{-1}$, the process $(\bar{Z}_k)_{k \in \mathbb{N}}$ also consists of independent equidistributed binary variables measurable against the shift-invariant σ -field of $(\bar{Y}_i)_{i \in \mathbb{Z}}$.

Repeat the construction of processes that precedes Proposition 6.5, putting $M_n := \sum_{i=1}^n |f(X_i)|$ and $\bar{M}_n := \sum_{i=1}^n |f(\bar{X}_i)|$. Recalling that $\mathbf{E}_P M_1 = L$, fix such a $C_a > 0$ that

$$\mathbf{E}_P [M_1 \mathbf{1}_{\{M_1 \leq C_a\}}] \geq aL$$

for some $a \in (\bar{\delta}, 1)$. Observe that $\bar{s}_k(\bar{Y}^m) = z$ if $s_k(\bar{X}_2^n) = z \in \{0, 1\}$ and $\bar{M}_n \leq m$. Hence,

$$\bar{P}(\bar{s}_k(\bar{Y}^m) = \bar{Z}_k) \geq \bar{P}(s_k(\bar{X}_2^n) = \bar{Z}_k, \bar{M}_1 \leq C_a, \bar{M}_n - \bar{M}_1 \leq m - C_a).$$

The event on the right-hand side is measurable $(\bar{X}_i)_{i \in \mathbb{Z}}$ since $\bar{Z}_k = s_k((\bar{X}_i)_{i \in \mathbb{Z}})$. On the other hand, $s_k((X_i)_{i \in \mathbb{Z}}) = Z_k$. Thus by (39) and further by the independence of the variable M_1 from (X_2^n, Z_k) we obtain

$$\begin{aligned} \bar{P}(s_k(\bar{X}_2^n) = \bar{Z}_k, \bar{M}_1 \leq C_a, \bar{M}_n - \bar{M}_1 \leq m - C_a) \\ &= L^{-1} \mathbf{E}_P \left[M_1 \mathbf{1}_{\{s_k(X_2^n) = Z_k, M_1 \leq C_a, M_n - M_1 \leq m - C_a\}} \right] \\ &= L^{-1} \mathbf{E}_P [M_1 \mathbf{1}_{\{M_1 \leq C_a\}}] \mathbf{E}_P \left[\mathbf{1}_{\{s_k(X_2^n) = Z_k, M_n - M_1 \leq m - C_a\}} \right]. \end{aligned}$$

But $(X_i)_{i \in \mathbb{Z}}$ is stationary, so the last expression yields simply

$$\bar{P}(\bar{s}_k(\bar{Y}^m) = \bar{Z}_k) \geq aP(s_k(X^n) = Z_k, M_n \leq m - C_a).$$

Now, by $P(A \cap B) \geq P(A) - P(B^c)$ and by the Markov inequality,

$$\begin{aligned} P(s_k(X^n) = Z_k, M_n \leq m - C_a) \\ &\geq P(s_k(X^n) = Z_k) - P(M_n > m - C_a) \\ &\geq P(s_k(X^n) = Z_k) - \frac{Ln}{m - C_a}. \end{aligned}$$

Taking $\delta > \bar{\delta}/a$ and $n = \lfloor (\delta - \bar{\delta}/a)L^{-1}(m - C_a) \rfloor$, we obtain

$$\begin{aligned} k \in \bar{U}_{\bar{\delta}}(m) &\iff a(P(s_k(X^n) = Z_k) - (\delta - \bar{\delta}/a)) \geq \bar{\delta} \\ &\iff P(s_k(X^n) = Z_k) \geq \delta \iff k \in U_{\delta}(n), \end{aligned}$$

so (9) follows for $\beta = \alpha^{-1}$ from (13).

□

Acknowledgements

The author thanks Peter Harremoës, Peter Grünwald, Jan Mielniczuk, and an anonymous referee for remarks that helped to improve the quality of this paper. The manuscript was completed during the author's leave from the Institute of Computer Science, Polish Academy of Sciences, whereas the research was supported by the grant no. 1/P03A/045/28 of the Polish Ministry of Scientific Research and Information Technology and under the PASCAL II Network of Excellence, IST-2002-506778.

References

- [1] R. Ahlswede, B. Balkenhol, and L. H. Khachatrian. Some properties of fix-free codes. In *Proceedings of the First INTAS International Seminar on Coding Theory and Combinatorics, 1996, Thakkadzor, Armenia*, pages 20–33. 1996.
- [2] R. Ahlswede, B. Balkenhol, C. Deppe, H. Mashurian, and T. Partner. T -shift synchronization codes. *Electr. Not. Disc. Math.*, 21:119–123, 2005.
- [3] J.-P. Allouche and J. Shallit. *Automatic Sequences. Theory, Applications, Generalizations*. Cambridge University Press, 2003.
- [4] D. Bajic, C. Stefanovic, and D. Vukobratovic. Search process and probabilistic bifix approach. In *Proceedings of the International Symposium on Information Theory, 2005*, pages 19–22. 2005.
- [5] A. R. Barron. The strong ergodic theorem for densities: Generalized Shannon-McMillan-Breiman theorem. *Ann. Probab.*, 13:1292–1303, 1985.
- [6] R. Capocelli, A. D. Santis, L. Gargano, and U. Vaccaro. The construction of statistically synchronizable codes. *IEEE Trans. Inform. Theor.*, 38:407–414, 1992.
- [7] G. Cariolaro and G. Pierobon. Stationary symbol sequences from variable-length word sequences. *IEEE Trans. Inform. Theor.*, 23:243–253, 1977.
- [8] M. Charikar, E. Lehman, A. Lehman, D. Liu, R. Panigrahy, M. Prabhakaran, A. Sahai, and A. Shelat. The smallest grammar problem. *IEEE Trans. Inform. Theor.*, 51:2554–2576, 2005.
- [9] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [10] D. R. Cox. *Renewal Theory*. London: Methuen and Company, 1962.
- [11] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: The entropy convergence hierarchy. *Chaos*, 15:25–54, 2003.
- [12] Ł. Dębowski. On the vocabulary of grammar-based codes and the logical consistency of texts. 2008. URL <http://arxiv.org/abs/0810.3125>.
- [13] Ł. Dębowski. A general definition of conditional information and its application to ergodic decomposition. *Statist. Probab. Lett.*, 79:1260–1268, 2009.
- [14] A. Elga. Self-locating belief and the Sleeping Beauty problem. *Analysis*, 60:143–147, 2000.
- [15] R. Fontana, R. Gray, and J. Kieffer. Asymptotically mean stationary channels. *IEEE Trans. Inform. Theor.*, 27:308–316, 1981.
- [16] D. Gillman and R. L. Rivest. Complete variable-length “fix-free” codes. *Designs Cod. Cryptogr.*, 5:109–114, 1995.

- [17] R. M. Gray and J. C. Kieffer. Asymptotically mean stationary measures. *Ann. Probab.*, 8:962–973, 1980.
- [18] H. L. Hurd. Stationarizing properties of random shifts. *SIAM J. Appl. Math.*, 26:203–212, 1974.
- [19] O. Kallenberg. *Foundations of Modern Probability*. Springer, 1997.
- [20] J. C. Kieffer and E. Yang. Grammar-based codes: A new class of universal lossless source codes. *IEEE Trans. Inform. Theor.*, 46:737–754, 2000.
- [21] U. Krengel. *Ergodic theorems*. Walter de Gruyter, 1985.
- [22] O. W. Reichard. Invariant measures for many-one transformations. *Duke Math. J.*, 23:477–488, 1956.
- [23] D. Salomon. *Variable-length Codes for Data Compression*. Springer, 2007.
- [24] P. C. Shields. *The Ergodic Theory of Discrete Sample Paths*. American Mathematical Society, 1996.
- [25] P. C. Shields. String matching bounds via coding. *Ann. Probab.*, 25:329–336, 1997.
- [26] J. Stiffler. *Theory of Synchronous Communications*. Prentice Hall, 1971.
- [27] R. Timo, K. Blackmore, and L. Hanlen. On the entropy rate of word-valued sources. In *Proceedings of the Telecommunication Networks and Applications Conference, ATNAC 2007*, pages 377–382. 2007.
- [28] G. K. Zipf. *The Psycho-Biology of Language: An Introduction to Dynamic Philology, 2nd ed.* The MIT Press, 1965.