View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Queen Mary Research Online

STOCHASTIC DOMINANCE FOR SHIFT-INVARIANT MEASURES

Vasso Anagnostopoulou*

Queen Mary University of London
Mile End Road
London, E1 4NS, UK

Abstract. Let X be the full shift on two symbols. The lexicographic order induces a partial order known as first-order stochastic dominance on the collection $\mathcal{M}_X$ of its shift-invariant probability measures. We present a study of the fine structure of this dominance order, denoted by $\prec$, and give criteria for establishing comparability or incomparability between measures in $\mathcal{M}_X$. The criteria also give an insight to the complicated combinatorics of orbits in the shift. As a by-product, we give a direct proof that Sturmian measures are totally ordered with respect to $\prec$.

1. Introduction

Let X be the full shift $\{0,1\}^{\mathbb{N}}$ on two symbols, and consider $\mathcal{M}_X$, the collection of shift-invariant Borel probability measures on X. If X is equipped with the lexicographic order, then $\mathcal{M}_X$ can be equipped with the partial order of first-order stochastic dominance: if μ and ν are shift-invariant Borel probability measures on X, then ν (first-order stochastically) dominates μ (written $\mu \prec \nu$), if $\int_X f d\mu \leq \int_X f d\nu$ for all increasing functions $f : X \to \mathbb{R}$.

The concept arises in decision theory and decision analysis (see, for example, [17]). In the setting of this article, first-order stochastic dominance is used to make precise the notion of one probability measure being larger than another. The study of this order in such a setting is motivated by interesting questions that arise in ergodic optimization: the study of the smallest and largest possible ergodic averages of a given function and of the invariant measures which attain these extrema, known as minimizing and maximizing measures, respectively. For example, in [3], β-shifts X_β (subsets of $\{0,\dots,[\beta]\}^{\mathbb{N}}$ defined as the closure with respect to the product topology of the set of sequences arising as a β-expansion) were considered and the set $\mathcal{M}_{X_\beta}$ of shift-invariant Borel probability measures was ordered with first-order stochastic dominance. Then it was natural to ask, for which $\beta > 1$, the corresponding β-shift has a largest invariant measure, i.e. a measure which dominates all other measures in $\mathcal{M}_{X_\beta}$ or equivalently, a measure that is simultaneously f-maximizing for every increasing function $f : X_\beta \to \mathbb{R}$. The authors proved that for $1 < \beta < 2$, the β-shift X_β has a largest shift-invariant measure if and only if the lexicographically largest element in X_β is the periodic sequence given by repeating the length-$(ap+1)$ word $(10^{a-1})^p0$, for some integers $a, p \geq 1$ (a more stringent restriction is needed for $\beta > 2$). In that case, the largest shift-invariant measure on X_β was the unique one supported by the periodic orbit generated by the largest

The author was partially supported by EPSRC grant EP/L02246X/1.

element of X_β. This measure also belongs to the remarkable one parameter family of Sturmian measures: measures which are supported on orbits whose permutation under the shift map is a rotation by angle ϱ. Sturmian sequences and measures have been studied extensively (e.g. [1, 4, 8, 19, 21, 24]), and appear naturally in many problems in ergodic optimization (e.g. [2, 5, 7, 11, 14, 15, 10]). In [14, 15], second-order stochastic dominance (also known as majorization) is studied, and Sturmian measures are shown to simultaneously optimise the integral of all convex functions.

To understand the partially ordered set $(\mathcal{M}_X, \prec)$, it is natural to study its fine structure. In this article, we examine the dominance relations between any two measures in $\mathcal{M}_X$ and present results which yield dominance, or incomparability, between them.

Useful reformulations of first-order stochastic dominance are detailed in Lemmas 3.1 and 3.2, a simplified version of which is the following. For atomic measures $\mu, \nu \in \mathcal{M}_X$, we have that $\mu \prec \nu$ if and only if $\mu[x, \overline{1}] \leq \nu[x, \overline{1}]$, for all $x \in \operatorname{supp}(\mu) \cup \operatorname{supp}(\nu)$. This allows to quickly establish dominance relations between atomic measures supported on periodic orbits of small period. In Theorem 3.4, we consider two atomic measures $\mu, \hat{\mu} \in \mathcal{M}_X$ and assume μ is dominated by $\hat{\mu}$. We prove that the measure supported on the orbit generated by the concatenation $\mu * \hat{\mu}$ (see Definition 2.7) of the two orbits carrying μ and $\hat{\mu}$, dominates μ, and is dominated by $\hat{\mu}$. This result also gives an insight to the combinatorics of periodic orbits that are constructed via concatenations.

For example, consider the periodic sequences $\overline{110}$ and $\overline{110010}$ (and the orbits they generate) and note that using the aforementioned reformulation of dominance, it is easy to establish that $\mu_{110010} \prec \mu_{110}$. Their concatenation is the orbit generated by $\overline{110} * \overline{110010} = \overline{110110010}$ and by Theorem 3.4, the measure $\mu_{110110010}$ carried by this orbit dominates μ_{110010} and is dominated by μ_{110}. The three orbits are illustrated in Figure 1, below.

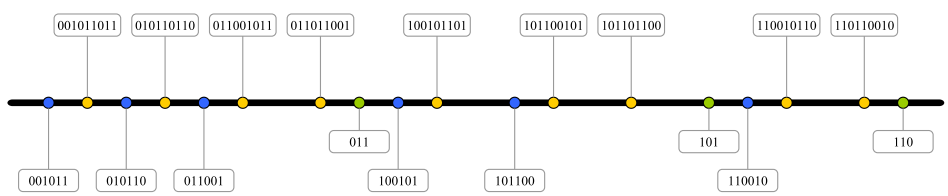

FIGURE 1. Orbits supporting measures $\mu_{110010} \prec \mu_{110110010} = \mu_{110} * \mu_{110010} \prec \mu_{110}$

An immediate use of this result is the possibility of establishing the dominance relation between all Sturmian measures i.e. the Sturmian measure S_{ϱ_1} is dominated by S_{ϱ_2} if and only if $\varrho_1 \leq \varrho_2$. This is due to the fact that Sturmian (periodic) orbits are generated by a concatenation procedure analogous to the Farey construction of rational numbers, see Figure 3, and Proposition 4.3(iv). Hence, for example,

$$S_0 \prec S_{\frac{1}{8}} \prec S_{\frac{1}{7}} \prec S_{\frac{1}{6}} \prec S_{\frac{1}{5}} \prec S_{\frac{1}{4}} \prec S_{\frac{2}{7}} \prec S_{\frac{1}{3}} \prec S_{\frac{3}{8}} \prec S_{\frac{2}{5}} \prec S_{\frac{3}{7}} \prec S_{\frac{1}{2}} \\ \prec S_{\frac{4}{7}} \prec S_{\frac{3}{5}} \prec S_{\frac{5}{8}} \prec S_{\frac{2}{3}} \prec S_{\frac{5}{7}} \prec S_{\frac{3}{4}} \prec S_{\frac{4}{5}} \prec S_{\frac{5}{6}} \prec S_{\frac{6}{7}} \prec S_{\frac{7}{8}} \prec S_1 \, .$$

This result is extended to all Sturmian measures with rotation number $\varrho \in [0, 1]$ and is presented in Corollary 4.4.

So, when are measures incomparable? In Figure 2, the orbits carrying the measures μ_{10}, μ_{1100}, μ_{110100} and μ_{110010} are illustrated. It is not hard to check that these are pairwise incomparable. Just consider the increasing functions $\chi_{[\overline{10},\overline{1}]}$, $\chi_{[\overline{0110},\overline{1}]}$, $\chi_{[\overline{011010},\overline{1}]}$, $\chi_{[\overline{010110},\overline{1}]}$, and compare their space averages with respect to each of the measures.

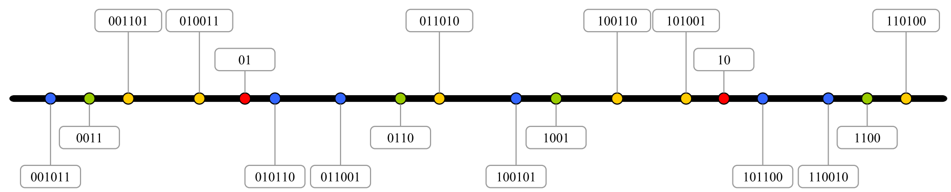

FIGURE 2. The pairwise incomparable shift-invariant probability measures μ_{10}, μ_{1100}, μ_{110100} and μ_{110010} of frequency $1/2$.

However, proving incomparability between measures directly from the definition of dominance (or its reformulations) becomes increasingly difficult as the period of the orbits carrying the corresponding measures gets larger. In this last example, one may observe that all of these measures have frequency (defined as the measure of the cylinder set $\langle 1 \rangle$) equal to $1/2$. In Lemma 5.1, it is shown that, in fact, any atomic measures with equal frequency are incomparable. Other results concerning incomparability between measures can be found in Section 5.

Using the various results on comparability and incomparability of measures (Sections 3 and 5) one can establish all relations between measures supported on periodic orbits of small period. For example, there are 71 measures supported on periodic orbits of period up to 8. The Sturmian measure $S_{1/2}$ is dominated by 16 of those, namely $S_{4/7}$, $S_{3/5}$, $S_{5/8}$, $S_{2/3}$, $S_{5/7}$, $S_{3/4}$, $S_{4/5}$, $S_{5/6}$, $S_{6/7}$, $S_{7/8}$, S_1, $\mu_{11101010}$, μ_{111010}, $\mu_{1111010}$, $\mu_{11111010}$ and $\mu_{11110110}$. By symmetry, it dominates $S_{3/7}$, $S_{2/5}$, $S_{3/8}$, $S_{1/3}$, $S_{2/7}$, $S_{1/4}$, $S_{1/5}$, $S_{1/6}$, $S_{1/7}$, $S_{1/8}$, S_0, $\mu_{10101000}$, μ_{101000}, $\mu_{1010000}$, $\mu_{10100000}$ and $\mu_{10010000}$, and is incomparable with the remaining 38 measures.

This complicated structure of the partial order of first-order stochastic dominance in $\mathcal{M}_X$ is illustrated in a Hasse diagram in Figure 4 where all measures supported on periodic orbits of period up to 7 are compared.

Overview. The article is organised as follows. Section 2 consists of preliminaries on symbolic dynamics, first-order stochastic dominance, and Sturmian measures. Sections 3 and 5 present a study on comparability and incomparability between measures. In Section 4, it is proved that Sturmian measures are totally ordered with respect to $\prec$. In Section 6, we take a closer look at the fine structure of the partially ordered set $(\mathcal{M}_X, \prec)$, give a Hasse diagram for dominance relations between atomic measures supported on orbits of small period, and prove that $(\mathcal{M}_X, \prec)$ is not a lattice.

2. PRELIMINARIES

Notation. Let X denote the full shift on two symbols, i.e. $\{0,1\}^{\mathbb{N}}$. An element of X will be denoted by $x = (x_n)_{n=1}^{\infty}$, where $x_n \in \{0,1\}$ for all $n \in \mathbb{N}$.

Definition 2.1 (*Lexicographic order, topology*)**.** For $x, x' \in X$, we write $x < x'$ if there exists an $N \in \mathbb{N}$ with $x_n = x'_n$ for $1 \leq n < N$ and $x_N < x'_N$; we write $x \leq x'$ if $x = x'$ or $x < x'$. This lexicographic order $\leq$ is a total order on X. Intervals on X are defined in the usual way. The shift X is equipped with the product topology; the (left) shift map $\sigma : X \to X$, defined by $(\sigma x)_n = x_{n+1}$ for $n \in \mathbb{N}$ is then continuous.

Definition 2.2 (*Words*)**.** Elements of X are called sequences (or infinite words). A (finite) word is any element of the set $\bigcup_{i=0}^{\infty}\{0,1\}^n$, $n \geq 1$ (by convention, the unique element of $\{0,1\}^n$ is the empty word) and will be written as $w = w_1 w_2 \ldots w_n$. The length of w denoted by $|w|$, is n. For each $n \in \mathbb{N}$, define $\pi_n : X \to \{0,1\}^n$ by $\pi_n(x) = x_1 x_2 \ldots x_n$. A sequence $x \in X$ is called periodic if there exists $q \in \mathbb{N}$ such that $x_n = x_{n+q}$ for all $n \in \mathbb{N}$. The smallest such q is called the period of x, and $\pi_q(x)$ is called the corresponding periodic word.

Definition 2.3 (*Smallest and largest points*)**.** Let $x \in X$ be a periodic sequence of period q. Define $x^+ := \max\{\sigma^n(x) : n \geq 0\}$, the lexicographically largest point in the orbit of x, and denote by w_x its corresponding periodic word $\pi_q(x^+)$. Similarly, let $x^- := \min\{\sigma^n(x) : n \geq 0\}$, the lexicographically smallest point in the orbit of x.

Definition 2.4 (*Cylinder set, invariant measures*)**.** Let w be any finite word. Define the cylinder set of w by $\langle w \rangle := \{x \in X : \pi_{|w|}(x) = w\}$, a closed interval in X. Let $\mathcal{M}_X$ denote the collection of shift-invariant Borel probability measures on X.

Notation. Let $x \in X$ be periodic. Denote by μ_x, the invariant probability measure in $\mathcal{M}_X$ whose support equals $\{\sigma^n(x), n \geq 0\}$. For simplicity, in examples, we will (almost always) write μ_{w_x} instead of $\mu_{\overline{w_x}} = \mu_x$. For example, the measure supported on the period-3 orbit $\{0010010\ldots, 0100100\ldots, 1001001\ldots\}$ will simply be represented by μ_{100}.

Definition 2.5 (*Measure frequency*)**.** Let $\mu \in \mathcal{M}_X$ be any shift-invariant Borel probability measure. Define its frequency to be the probability with which the digit 1 appears, i.e. $\varphi(\mu) := \mu(\langle 1 \rangle)$.

Definition 2.6 (*Conjugate*)**.** Define the *conjugate* of a sequence $x = (x_i)_{i=1}^{\infty} \in X$ by $\mathrm{conj}(x) = (y_i)_{i=1}^{\infty} \in X$, where $y_i := 1 - x_i$, $n \in \mathbb{N}$.

Remark 1. Note that, for $x \in X$, $\varphi(\mu_x) = 1 - \varphi(\mu_{\mathrm{conj}(x)})$ and for a Borel set $A \subset X$, $\mu_{\mathrm{conj}(x)}(A) = \mu_x(\mathrm{conj}(A))$.

Definition 2.7 (*Concatenation*)**.** Let x, $\hat{x} \in X$ be periodic sequences with $x^+ \leq \hat{x}^+$. Define the concatenation of x, $\hat{x}$, by $x * \hat{x} = \overline{w_{\hat{x}} w_x}$. If $\mu_x, \mu_{\hat{x}} \in \mathcal{M}_X$ are atomic measures supported on those orbits, denote the concatenation of the two measures by $\mu_x * \mu_{\hat{x}} = \mu_{x * \hat{x}}$.

Note that concatanation of orbits can be defined in different ways, for example by concatenating the corresponding periodic words of the lexicographically smallest point in each orbit.

2.1. **Stochastic dominance.**

Definition 2.8. Let $\mu(f) := \int f d\mu$ and $\mu, \nu \in \mathcal{M}_X$. We say that μ is *dominated* (first-order stochastically) by ν (or ν *dominates* μ), and write $\mu \prec \nu$, if $\mu(f) \leq \nu(f)$ for every increasing[1] function $f : X \to \mathbb{R}$.

[1] $f : X \to \mathbb{R}$ is *increasing* if $f(x) \leq f(x')$ whenever $x \leq x'$.

Some examples of measures ordered by $\prec$ are, $\mu_0 \prec \mu_1$, $\mu_{01} \prec \mu_1$, $\mu_{10} \prec \mu_{11010}$. Not all measures are comparable under this dominance relation, for example observe that $\int \chi_{[1100,\overline{1}]} d\mu_{10} = 0 < 1/4 = \int \chi_{[1100,\overline{1}]} d\mu_{1100}$ and $\int \chi_{[10,\overline{1}]} d\mu_{10} = 1/2 > 1/4 = \int \chi_{[10,\overline{1}]} d\mu_{1100}$, i.e. there is no dominance between μ_{10} and μ_{1100}, and thus $\prec$ defines a partial order on $\mathcal{M}_X$.

Definition 2.9 (*Incomparable measures*)**.** Let $\mu, \nu \in \mathcal{M}_X$. If μ is not dominated by ν, and ν is not dominated by μ, we say μ and ν are *incomparable* with respect to first-order stochastic dominance, and write $\mu \curlywedge \nu$.

The following is a well known characterisation of first-order stochastic dominance using so-called couplings (see Strassen's theorem [23], [18]):

Definition 2.10 (*Couplings and first-order stochastic dominance*)**.** Let $\mu, \nu \in \mathcal{M}_X$. A coupling of μ and ν is a probability measure γ on the product space $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ such that the marginals of γ coincide with μ and ν, i.e. $\gamma(A \times \{0,1\}^{\mathbb{N}}) = \mu(A)$ and $\gamma(\{0,1\}^{\mathbb{N}} \times A) = \nu(A)$ for all Borel sets $A \subseteq \{0,1\}^{\mathbb{N}}$.

Then, two measures $\mu, \nu \in M_X$ satisfy $\mu \prec \nu$ if and only if they admit a coupling (i.e. a probability measure on $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ with marginals μ, ν) whose support is contained in the half plane $\{(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \;:\; x \leq y\}$.

3. On comparability of measures

In this section, conditions that establish dominance between measures are investigated. These results are crucial in understanding the fine structure of the partially ordered set $(\mathcal{M}_X, \prec)$.

Firstly, observe that $\prec$ is a transitive relation. Simple approximation arguments (cf. e.g. [16]) give the following reformulations of dominance:

Lemma 3.1. *Let $\mu, \nu \in \mathcal{M}_X$, then the following are equivalent:*
(i) $\mu \prec \nu$,
(ii) $\mu(x, \overline{1}] \leq \nu(x, \overline{1}]$ *for all* $x \in X$,
(iii) $\mu[x, \overline{1}] \leq \nu[x, \overline{1}]$ *for all* $x \in X$.

It is also useful to simplify this lemma, for the case of atomic measures.

Lemma 3.2. *Let $\mu, \nu \in \mathcal{M}_X$, and in addition, let μ be an atomic measure. Then the following are equivalent:*
(i) $\mu \prec \nu$,
(ii) $\mu(x, \overline{1}] \leq \nu(x, \overline{1}]$ *for all* $x \in \operatorname{supp}(\mu)$,
(iii) $\mu[x, \overline{1}] \leq \nu[x, \overline{1}]$ *for all* $x \in \operatorname{supp}(\mu)$.

Proof. $(i) \Rightarrow (ii)$ and (iii) are obvious from the previous lemma. Now it will be shown that $(ii) \Rightarrow (i)$. For any $y \in X$, and let $\tilde{x} = \min\{x \in \operatorname{supp}(\mu) \;:\; x \geq y\}$. Then $\mu(y, \overline{1}] = \mu(\tilde{x}, \overline{1}] \leq \nu(\tilde{x}, \overline{1}] \leq \nu(y, \overline{1}]$. Similarly $(iii) \Rightarrow (i)$ is shown. □

Remark 2. It is easy to show that Lemma 3.2 also holds if measure ν is atomic, and the assumption that μ is atomic is removed. In this case, the equivalence holds if inequalities in (ii) and (iii) hold for all $x \in \operatorname{supp}(\nu)$.

Lemma 3.3. *Let $x \in X$ be periodic and w be its corresponding periodic word (note that it is not necessary that $w = w_x$, i.e. that x is lexicographically largest in its orbit). Then,*
(a) *If* $\overline{1w} = (\overline{1w})^+$, *the lexicographically largest point in its orbit, then* $\mu_{\overline{w}} \prec \mu_{\overline{1w}}$.

(b) *If* $\overline{0w} = (\overline{0w})^-$, *the lexicographically smallest point in its orbit, then* $\mu_{\overline{0w}} \prec \mu_{\overline{w}}$.

Proof. By Lemmas 3.1 and 3.2, it is enough to show

$$\mu_{\overline{w}}[\sigma^i(\overline{w}), \overline{1}] \leq \mu_{\overline{1w}}[\sigma^i(\overline{w}), \overline{1}] \tag{1}$$

for all $i \in \{0, \ldots, q-1\}$, where q is the period of x. If $w = 1$ then (1) is trivially true. Now let $w \neq 1$, then $ww \ldots < w1w1 \ldots$ and hence

$$\sigma^i(\overline{w}) < \sigma^i(\overline{w1}) < \overline{1w}, \quad \text{for all } i \in \{0, \ldots, q-1\} \tag{2}$$

Let $b_1 < \ldots < b_q$ be the atoms of $\mu_{\overline{w}}$, where $\{b_j\}_{j=1}^q$ is a permutation of $\{\sigma^i(\overline{w})\}_{i=0}^{q-1}$. Then (2) implies

$$\mu_{\overline{w}}[b_j, \overline{1}] = \frac{q-j+1}{q} \leq \frac{q-j+1+1}{q+1} \leq \mu_{\overline{1w}}[b_j, \overline{1}] \tag{3}$$

for all $j = 1, \ldots, q$, which proves (a). The proof of (b) is identical to that of (a) and is omitted. □

Example. Lemma 3.3 implies that the measures $\mu_{11001010}$, $\mu_{11010010}$ and $\mu_{11010100}$ are all dominated by $\mu_{1101010}$. Another dominance relation implied is $\mu_{10100} \prec \mu_{110010}$. Note that $\overline{10010}$ is not lexicographically largest in its orbit.

Remark 3. In Lemma 3.3, the condition that $\overline{1w}$ (respectively $\overline{0w}$) is lexicographically largest (respectively smallest) in its orbit, is necessary. For example, if $w = 100110$ then it is not implied that μ_{100110} is dominated by $\mu_{1100110}$. In fact it will be shown later that the two measures are incomparable.

The following theorem gives an insight to the structure and combinatorics of periodic orbits that are constructed via concatenations, in relation to the periodic orbits they were constructed from.

Theorem 3.4. *Consider periodic points* x, $\hat{x} \in X$, *with* $x^+ \leq \hat{x}^+$. *Then*

$$\mu_x \prec \mu_{\hat{x}} \iff \mu_x \prec \mu_{x*\hat{x}} \iff \mu_{x*\hat{x}} \prec \mu_{\hat{x}}$$

Proof. Assume without loss of generality that $x = x^+$ and $\hat{x} = \hat{x}^+$. If $x = \hat{x}$ then the theorem holds trivially. Now, let $x \neq \hat{x}$. If $x = \overline{0}$ or $\hat{x} = \overline{1}$, then all relations dominance relations hold (by Lemma 3.3 or trivially). Now suppose that $x \neq \overline{0}$ and $\hat{x} \neq \overline{1}$.

First, it will be shown that if $\mu_x \prec \mu_{\hat{x}}$, then $\mu_x \prec \mu_{x*\hat{x}} \prec \mu_{\hat{x}}$. By Lemma 3.1, it is enough to show

$$\mu_x[y, \overline{1}] \leq \mu_{x*\hat{x}}[y, \overline{1}] \leq \mu_{\hat{x}}[y, \overline{1}] \quad \text{for all } y \in X. \tag{4}$$

Inequality (4) will be proved in two stages. For the left inequality, by Lemma 3.2, it is enough to prove the following claim:

Claim 1. Let $x^{(1)} < \ldots < x^{(q)} = x^+$ be the atoms of μ_x where q is the period of x. Then,

$$\mu_x[x^{(i)}, \overline{1}] \leq \mu_{x*\hat{x}}[x^{(i)}, \overline{1}] \quad \text{for all } i = 1, \ldots, q. \tag{5}$$

Proof of Claim. We have that, $\mu_x[x^{(i)}, \overline{1}] = (q-i+1)/q$ and $\mu_{\hat{x}}[x^{(i)}, \overline{1}] = \delta_i/\hat{q}$ where $\hat{q}$ is the period of $\hat{x}$ and δ_i denotes the cardinality $|\cdot|$ of the set $[x^{(i)}, \overline{1}] \cap$

$\text{supp}(\mu_{\hat{x}})$, $i = 1, \dots, q$. Using Lemma 3.1, the hypothesis implies $(q-i+1)/q \leq \delta_i/\hat{q}$ for all $i = 1, \dots, q$, and consequently,

(6)
$$\mu_x[x^{(i)}, \overline{1}] = \frac{q-i+1}{q} \leq \frac{(q-i+1)+\delta_i}{q+\hat{q}} \leq \frac{\delta_i}{\hat{q}} = \mu_{\hat{x}}[x^{(i)}, \overline{1}] \quad \text{for all } i = 1, \dots, q.$$

It will be shown that, by construction of $x * \hat{x}$,

$$\mu_{x*\hat{x}}[x^{(i)}, \overline{1}] = \frac{(q-i+1)+\delta_i}{q+\hat{q}} \quad \text{for all } i = 1, \dots, q. \tag{7}$$

Let $x^+ = \overline{b_1 \dots b_q} = \overline{w_x}$, $\hat{x}^+ = \overline{c_1 \dots c_{\hat{q}}} = \overline{w_{\hat{x}}}$. Then $x * \hat{x} = \overline{c_1 \dots c_{\hat{q}} b_1 \dots b_q}$. Consider the intervals $J_i = [x^{(i)}, x^{(i+1)})$, $i \in \{1, \dots, q-1\}$, $J_q = [x^{(q)}, \overline{1}]$. Define $k_i \in \{2, \dots, q\}$, where $i \in \{1, \dots, q-1\}$, to be such that $x^{(i)} := \overline{b_{k_i} \dots b_q b_1 \dots b_{k_i - 1}}$ and $k_q := 1$. Then $\{k_1, \dots k_q\}$ is a permutation of $\{1, \dots, q\}$. Now, consider the point $\overline{b_{k_i} \dots b_q w_{\hat{x}} b_1 \dots b_{k_i-1}}$ in the support of $\mu_{x*\hat{x}}$. We claim that the point $\overline{b_{k_i} \dots b_q w_{\hat{x}} b_1 \dots b_{k_i-1}}$ belongs to J_i.

Observe that,

$$x^{(i)} = \overline{b_{k_i} \dots b_q b_1 \dots b_{k_i-1}} < \overline{b_{k_i} \dots b_q w_{\hat{x}} b_1 \dots b_{k_i-1}} \tag{8}$$

since $\overline{w_x} < \overline{w_{\hat{x}} w_x}$. Also, by the definition of the points $x^{(i)}$, it is clear that for $i < q$,

$$x^{(i)} = \overline{b_{k_i} \dots b_q b_1 \dots b_{k_i-1}} < \overline{b_{k_{i+1}} \dots b_q b_1 \dots b_{k_{i+1}-1}} = x^{(i+1)}. \tag{9}$$

Now, since $\sigma^{q-k_i+1}(x^{(i)}) = x^+ > \sigma^{q-k_i+1}(x^{(i+1)})$, it cannot be that $x^{(i)}$ and $x^{(i+1)}$ have the same first $q - k_i + 1$ digits (because it would contradict that fact that $x^{(i)} < x^{(i+1)}$). Therefore an inequality has to appear in those first $q - k_i + 1$ digits.

This implies

$$\overline{b_{k_i} \dots b_q w_{\hat{x}} b_1 \dots b_{k_i-1}} < \overline{b_{k_{i+1}} \dots b_q b_1 \dots b_{k_{i+1}-1}} = x^{(i+1)} < 1. \tag{10}$$

By inequalities (8) and (10), it is derived that there is at least one atom of $\mu_{x*\hat{x}}$ of the form $\overline{b_{k_i} \dots b_q w_{\hat{x}} b_1 \dots b_{k_i-1}}$, $k_i \in \{2, \dots, q\}$ in the interior of each interval J_i, $i \in \{1, \dots, q-1\}$, and clearly $\overline{b_1 \dots b_q w_{\hat{x}}}$ is in the interior of J_q. Since there is a total of q of these points that lie in q intervals, there is precisely *one* of those atoms of $\mu_{x*\hat{x}}$ in each J_i, $i \in \{1, \dots, q\}$.

Now let $\hat{x}^{(1)} < \dots < \hat{x}^{(\hat{q})} = \hat{x}^+$ be the atoms of $\mu_{\hat{x}}$ and let $i \in \{1, \dots, q\}$ be such that $\text{supp}(\mu_{\hat{x}}) \cap J_i \neq \emptyset$ and let $j \in \{1, \dots, \hat{q}\}$ be such that $\hat{x}^{(j)} \in \text{supp}(\mu_{\hat{x}}) \cap [x^{(i)}, x^{(i+1)}]$, where clearly, $j = j(i)$. Define $n_j \in \{2, \dots, \hat{q}\}$ to be such that $\hat{x}^{(j)} = \overline{c_{n_j} \dots c_{\hat{q}} c_1 \dots c_{n_j-1}}$ and $n_{\hat{q}} = 1$. Consider the point $\overline{c_{n_j} \dots c_{\hat{q}} w_x c_1 \dots c_{n_j-1}}$ in the support of $\mu_{x*\hat{x}}$. Then, we claim

$$x^{(i)} < \overline{c_{n_j} \dots c_{\hat{q}} w_x c_1 \dots c_{n_j-1}}. \tag{11}$$

By definition of $\hat{x}^{(j)} = \hat{x}^{(j(i))}$, $\overline{b_{k_i} \dots b_q b_1 \dots b_{k_i-1}} = x^{(i)} < \hat{x}^{(j)} = \overline{c_{n_j} \dots c_{\hat{q}} c_1 \dots c_{n_j-1}}$. Now, if $(x^{(i)})_\lambda < (\hat{x}^{(j)})_\lambda$ for some $\lambda \in \{1, \dots, \hat{q} - n_j + 1\}$, then inequality (11) holds. Otherwise, $(x^{(i)})_\lambda = (\hat{x}^{(j)})_\lambda$ for all $\lambda \in \{1, \dots, \hat{q} - n_j + 1\}$, but then (11) is implied by the fact that

(12)
$$\sigma^{\hat{q}-n_j+1}(x^{(i)}) < x^+ = \overline{w_x} < \overline{w_x w_{\hat{x}}} = w_x c_1 \dots c_{\hat{q}} = \sigma^{\hat{q}-n_j+1}(\overline{c_{n_j} \dots c_{\hat{q}} w_x c_1 \dots c_{n_j-1}})$$

Now, the inequality $\overline{w_x} < \overline{w_{\hat{x}}}$ implies

$$\overline{c_{n_j} \dots c_{\hat{q}} w_x c_1 \dots c_{n_j-1}} < \overline{c_{n_j} \dots c_{\hat{q}} c_1 \dots c_{n_j-1}} = \hat{x}^{(j)} < x^{(i+1)} < 1. \tag{13}$$

Combining (11) and (13) we see that for each point $\hat{x}^{(j)}$ that lies in J_i, $i \in \{1, \dots, q\}$, there is an atom of $\mu_{x*\hat{x}}$ that lies in $[x^{(i)}, \hat{x}^{(j)}] \subsetneq [x^{(i)}, x^{(i+1)})$. Since there is a total of $\hat{q}$ atoms of $\mu_{\hat{x}}$ (all of which lie in $\cup_{i=1}^{q} J_i$), and the same number of atoms of $\mu_{x*\hat{x}}$ of the form $\overline{c_n \dots c_{\hat{q}} w_x c_1 \dots c_{n-1}}$, $n \in \{1, \dots, \hat{q}\}$, then, using all the above work,

$$\left|\operatorname{supp}(\mu_{x*\hat{x}}) \cap J_i\right| = 1 + \left|\operatorname{supp}(\mu_{\hat{x}}) \cap J_i\right| \quad , \ i \in \{1, \dots, q\}.$$

Thus,

$$\mu_{x*\hat{x}}(J_i) = \frac{1 + \left|\operatorname{supp}(\mu_{\hat{x}}) \cap J_i\right|}{q + \hat{q}} \quad , \ i \in \{1, \dots, q\}$$

and so

$$\mu_{x*\hat{x}}[x^{(i)}, \overline{1}] = \sum_{l=i}^{q} \frac{1 + \left|\operatorname{supp}(\mu_{\hat{x}}) \cap J_l\right|}{q + \hat{q}} = \frac{(q - i + 1) + \delta_i}{q + \hat{q}} \quad , \ i \in \{1, \dots, q\}$$

which is the required equality in (7), hence Claim 1 is proved.

For the right inequality in (4), the argument is similar. By Lemma 3.2 and Remark 2, is it enough to prove the following claim:

Claim 2. Let $\hat{x}^{(1)} < \dots < \hat{x}^{(\hat{q})} = \hat{x}^+$ be the atoms of $\mu_{\hat{x}}$. Then,

$$\mu_{x*\hat{x}}[\hat{x}^{(j)}, \overline{1}] \leq \mu_{\hat{x}}[\hat{x}^{(j)}, \overline{1}], \quad \text{for all } j = 1, \dots, \hat{q}. \tag{14}$$

Proof of Claim. For each $j \in \{1, \dots, \hat{q}\}$, $\mu_{\hat{x}}[\hat{x}^{(j)}, \overline{1}] = (\hat{q} - j + 1)/\hat{q}$ and $\mu_x[\hat{x}^{(j)}, \overline{1}] = \hat{\delta}_j / q$ where $\hat{\delta}_j := \left|\operatorname{supp}(\mu_x) \cap (\hat{x}^{(j)}, \overline{1}]\right|$. By hypothesis, $\mu_x[\hat{x}^{(j)}, \overline{1}] \leq \mu_{\hat{x}}[\hat{x}^{(j)}, \overline{1}]$, which implies

$$\mu_x[\hat{x}^{(j)}, \overline{1}] = \frac{\hat{\delta}_j}{q} \leq \frac{\hat{\delta}_j + \hat{q} - j + 1}{q + \hat{q}} \leq \frac{\hat{q} - j + 1}{\hat{q}} = \mu_{\hat{x}}[\hat{x}^{(j)}, \overline{1}].$$

It will be shown that, by construction of $x * \hat{x}$,

$$\mu_{x*\hat{x}}(\hat{x}^{(j)}, \overline{1}] \leq \frac{\hat{\delta}_j + 1 + \hat{q} - j}{q + \hat{q}} \quad \text{for all } j = 1, \dots, \hat{q}, \tag{15}$$

from which inequality (14) follows. Consider the interval $[\hat{x}^{(j)}, \overline{1}]$, where $\hat{x}^{(j)} \in \operatorname{supp}(\mu_{\hat{x}})$, $j \in \{1, \dots, \hat{q}\}$.

In the proof of (7), it was shown that there is precisely one atom of $\mu_{x*\hat{x}}$ of the form, $\overline{b_k \dots b_q w_{\hat{x}} b_1 \dots b_{k-1}}$, $k \in \{1, \dots, q\}$ in each J_i. Since there are $\hat{\delta}_j$ atoms of μ_x in $[\hat{x}^{(j)}, \overline{1}]$, there are at least $\hat{\delta}_j$ and at most $\hat{\delta}_j + 1$ atoms of $\mu_{x*\hat{x}}$ of the form $\overline{b_k \dots b_q w_{\hat{x}} b_1 \dots b_{k-1}}$, $k \in \{1, \dots, q\}$, in $[\hat{x}^{(j)}, \overline{1}]$, $j \in \{1, \dots, \hat{q}\}$.

In addition, as seen in inequality (13) for any point $\hat{x}^{(j)} \in \operatorname{supp}(\mu_{\hat{x}})$, there is an atom of $\mu_{x*\hat{x}}$ of the form $\overline{c_{n_j} \dots c_q w_x c_1 \dots c_{n_j-1}} < \hat{x}^{(j)}$. This implies that there are at least $j = \mu_{\hat{x}}[\overline{0}, \hat{x}^{(j)}]$ atoms of $\mu_{x*\hat{x}}$ of the form $\overline{c_n \dots c_q w_x c_1 \dots c_{n-1}}$, $n \in \{1, \dots, \hat{q}\}$ which lie in $[\overline{0}, \hat{x}^{(j)})$, $j \in \{1, \dots, \hat{q}\}$ (out of a total of $\hat{q}$ points of this form).

Therefore, $\mu_{x*\hat{x}}[\hat{x}^{(j)}, \overline{1}] \leq \hat{\delta}_j + 1 + \hat{q} - j$ for all $j \in \{1, \dots, \hat{q}\}$, which proves (15), and thus Claim 2 is proved, which completes the first part of the proof.

Now it will be shown that if $\mu_x \prec \mu_{x*\hat{x}}$ or $\mu_{x*\hat{x}} \prec \mu_{\hat{x}}$, then $\mu_x \prec \mu_{\hat{x}}$. Assume that $\mu_x \curlywedge \mu_{\hat{x}}$. Now, observe that $x^+ < (x*\hat{x})^+ < \hat{x}^+$, hence:

$$\mu_x[(x*\hat{x})^+, \overline{1}] = 0 < \mu_{x*\hat{x}}[(x*\hat{x})^+, \overline{1}] \tag{16}$$

and

$$\mu_{x*\hat{x}}[\hat{x}^+, \overline{1}] = 0 < \mu_{\hat{x}}[\hat{x}^+, \overline{1}]. \tag{17}$$

Now, since $\mu_x \curlywedge \mu_{\hat{x}}$, by Lemma 3.2, there exists some $x^{(i)} \in \text{supp}(\mu_x)$, $i = 1, 2, \dots q$, such that $\mu_x[x^{(i)}, \overline{1}] > \mu_{\hat{x}}[x^{(i)}, \overline{1}]$. Carrying through the exact same argument as in Claim 1 (with the inequalities in (6) in the other direction), we derive that

$$\mu_x[x^{(i)}, \overline{1}] > \mu_{x*\hat{x}}[x^{(i)}, \overline{1}] > \mu_{\hat{x}}[x^{(i)}, \overline{1}] \tag{18}$$

for some $x^{(i)} \in \text{supp}(\mu_x)$. But this, together with (16) and (17), implies $\mu_x \curlywedge \mu_{x*\hat{x}}$ and $\mu_{x*\hat{x}} \curlywedge \mu_{\hat{x}}$, a contradiction. □

Remark 4. Note that Theorem 3.4 is easily shown to hold if we define the concatenation of orbits by concatenating the corresponding periodic words of the lexicographically smallest point in each orbit, i.e. if $x * \hat{x} := \pi_q(x^-)\pi_{\hat{q}}(\hat{x}^-)$, see Definition 2.7 for comparison.

Lemma 3.5. *Let $x, \hat{x} \in X$ be periodic of the same period, with $x \le \hat{x}$ and assume that their periodic words differ only by one digit. Then, $\mu_x \prec \mu_{\hat{x}}$.*

Proof. Without loss of generality, it can be assumed that the corresponding periodic words of $x, \hat{x}$ agree on the first $q-1$ digits and disagree on the q^{th} digit, i.e. $x_i = \hat{x}_i$ for all $i = \{1, \dots, q-1\}$ and $x_q = 0, \hat{x}_q = 1$.

Then, $x_{\lambda q} = 0 < 1 = \hat{x}_{\lambda q}$ for all $\lambda \in \mathbb{N}$ and $x_{n+\lambda q} = x_{n+\lambda q}$ for $\mathcal{N} \in \{1, \dots, q-1\}$, $\lambda \in \mathbb{N}_0$. This implies $\sigma^\kappa(x) < \sigma^\kappa(\hat{x})$ for all $\kappa \in \mathbb{N}$, and hence

$$\mu_x[\sigma^\kappa(x), \overline{1}] = \frac{\left|\,\text{supp}(\mu_x) \cap (\sigma^\kappa(x), \overline{1}]\,\right|}{q} \le \frac{\left|\,\text{supp}(\mu_{\hat{x}}) \cap (\sigma^\kappa(x), \overline{1}]\,\right|}{q} = \mu_{\hat{x}}[\sigma^\kappa(x), \overline{1}]$$

for all $\kappa \in \{0, \dots, q-1\}$. By Lemmas 3.1 and 3.2, the result follows. □

Example. By Lemma 3.5, $\mu_{1000} \prec \mu_{1100}$. Then Theorem 3.4 implies $\mu_{1000} \prec \mu_{11001000} \prec \mu_{1100}$.

Lemma 3.6. *Let $x, \hat{x} \in X$ be periodic. If $\mu_x \prec \mu_{\hat{x}}$ then $\mu_{\text{conj}(\hat{x})} \prec \mu_{\text{conj}(x)}$.*

Proof. We will argue by contradiction. Assume that $\mu_{\text{conj}(\hat{x})}$ is not dominated by $\mu_{\text{conj}(x)}$. Then, by Lemma 3.1, there exists some $y \in X$ such that $\mu_{\text{conj}(\hat{x})}(y, \overline{1}] > \mu_{\text{conj}(x)}(y, \overline{1}]$. However, by Remark 1 this implies that $\mu_{\hat{x}}[\overline{0}, \overline{1} - y) > \mu_x[\overline{0}, \overline{1} - y)$, hence $1 - \mu_{\hat{x}}[\overline{1} - y, \overline{1}] > 1 - \mu_x[\overline{1} - y, \overline{1}]$, and so $\mu_{\hat{x}}[\overline{1} - y, \overline{1}] < \mu_x[\overline{1} - y, \overline{1}]$ for some $y \in X$. This contradicts our assumption that $\mu_x \prec \mu_{\hat{x}}$, by Lemma 3.1. □

4. Sturmian sequences and measures

The focus of this section will be the remarkable one parameter family of Sturmian sequences and measures. Their definition and several of their properties will be recalled here whilst as a corollary of Theorem 3.4, it will be proved that Sturmian measures are totally ordered with respect to first-order stochastic dominance.

First, consider the semi-conjugacy $\psi : X \to [0,1]$, $\psi(x_i)_{i=1}^\infty = \sum_{i=1}^\infty \frac{x_i}{2^i} \in [0,1]$, between the doubling map $T : [0,1] \to [0,1]$, $T(x) = 2x \pmod 1$ and the (left) shift map $\sigma : X \to X$, $\sigma(x_n)_{n=1}^\infty = (x_{n+1})_{n=1}^\infty$.

Lemma 4.1. *There exists precisely one shift-invariant probability measure on $\mathcal{M}_X$ whose support is contained in an interval $[x, \hat{x}]$ such that $m(\psi[x, \hat{x}]) = 1/2$ where m denotes Lebesgue measure on $[0, 1]$.*

Proof. See [7, 8]. □

Definition 4.2. Any shift-invariant Borel probability measure in $\mathcal{M}_X$ is called Sturmian if its support lies in an interval $[x, \hat{x}]$, $x, \hat{x} \in X$ such that $m(\psi[x, \hat{x}]) = 1/2$.

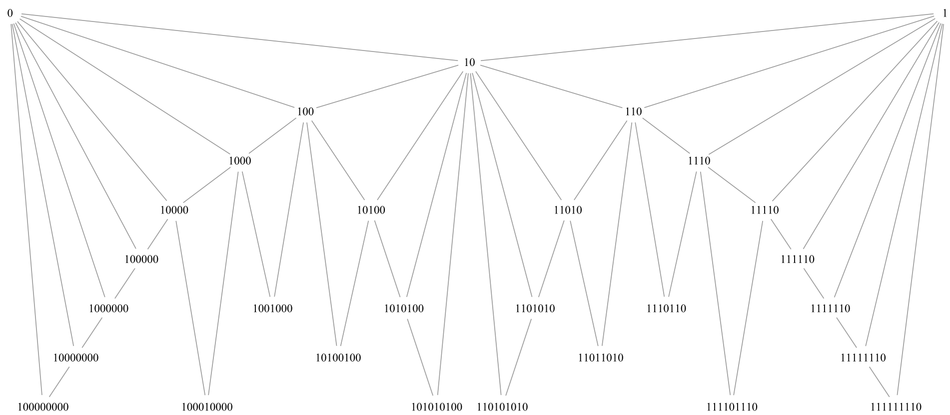

Figure 3. A concatenation procedure that generates the largest point in the support of a Sturmian measure $S_{p/q}$

The following are well known properties of Sturmian sequences and measures.

Proposition 4.3.
(i) *For every $\varrho \in [0, 1]$ there exists a unique Sturmian measure whose frequency equals ϱ. This measure will be denoted by S_ϱ.*
(ii) *For $\varrho \in (0, 1)$ if $R_\varrho(t) = t + \varrho \pmod 1$ and*

$$x_n^{(\varrho)}(t) := \begin{cases} 0 & \textit{if } R_\varrho^{n-1}(t) \in [0, 1-\varrho) \\ 1 & \textit{otherwise}, \end{cases}$$

then S_ϱ is the push forward of Lebesgue measure on $[0, 1)$ under the map $t \mapsto (x_n^{(\varrho)}(t))_{n=1}^\infty$.
(iii) *If ϱ is irrational, the support of S_ϱ is a uniquely ergodic Cantor set. If ϱ is rational then S_ϱ is supported on a single periodic orbit; if $\varrho = p/q$, $p, q \in \mathbb{N}$, $p \in [0, q-1]$, and $\gcd(p, q) = 1$, and the points in this orbit are $s_1 < \ldots < s_q$, then the shift σ acts as a cyclic permutation:*

$$\sigma(s_i) = \begin{cases} s_{i+p} & \textit{if } i \in [0, q-p] \\ s_{i+p-q} & \textit{if } i \in [q-p+1, q], \end{cases}$$

i.e. it is combinatorially equivalent to the action of $R_{\pi/q}(t) := t + p/q \pmod 1$ on any of its periodic orbits.
(iv) *The largest (or smallest) point in the support of a Sturmian measure can be generated by a concatenation procedure analogous to the Farey construction of rational numbers (see Figure 3 and e.g. [9, Ch. III]). For $n \geq 1$, let $\mathfrak{F}_n$ denote the*

order-n Farey sequence*, i.e. the increasing finite sequence consisting of rationals in* $[0,1]$ *whose denominator is at most* n *(*0 *and* 1 *are included as* $\frac{0}{1}$ *and* $\frac{1}{1}$ *respectively). If* $p_1/q_1 < p_2/q_2$ *(with* $p_1, p_2, q_1, q_2 \in \mathbb{N}$, $\gcd(p_1, q_1) = 1 = \gcd(p_2, q_2)$*) are Farey-consecutive (i.e. they appear as successive terms in some* $\mathfrak{F}_n$*), and* $s^+_{p/q} = s^+_\varrho$ *is the largest point in the support of* $S_{p/q}$*, (with* $p, q \in \mathbb{N}, \gcd(p, q) = 1$*), then*

$$s^+_{\frac{p_1+p_2}{q_1+q_2}} = s^+_{\frac{p_2}{q_2}} s^+_{\frac{p_1}{q_1}} .$$

If $\varrho \in [0,1]$ *is irrational, and* $\{\varrho_i\}_{i=1}^\infty$ *is any sequence of rational numbers in* $[0,1]$ *converging to* ϱ*, then* s^+_ϱ *(the largest point in the support of* S_ϱ*)) is the limit of* $s^+_{\varrho_i}$ *as* $i \to \infty$*.*

Proof. Property (i) follows from existence and uniqueness of S_ϱ for $\varrho \in [0,1)$ (See [7, 8]). For property (ii), see for example [2, 5, 8, 14, 19]. Property (iii) follows from (ii), also see [8]. For (iv), the proof is easily adapted from one in [20], see also [11]. □

Corollary 4.4. *Sturmian measures are totally ordered with respect to first-order stochastic dominance, i.e. if* $\varrho < \varrho'$, $\varrho, \varrho' \in [0,1]$*, then* $S_\varrho \prec S_{\varrho'}$*.*

Proof. First, it will be shown that Sturmian measures with rational rotation number are totally ordered. It is clear that $S_0 \prec S_1$ (the Dirac measures at $\overline{0}$ and $\overline{1}$ respectively) but by Proposition 4.3, the largest point in the support of a every Sturmian measure with rational rotation number can be generated using concatenations analogous to the Farey construction of rational numbers. Hence, by Theorem 3.4, if $\frac{p_1}{q_1} < \frac{p_2}{q_2}$ (with $p_1, p_2, q_1, q_2 \in \mathbb{N}$, $\gcd(p_1, q_1) = 1 = \gcd(p_2, q_2)$) are Farey-consecutive, then

$$S_{p_1/q_1} \prec S_{p_1/q_1} * S_{p_2/q_2} = S_{\frac{p_1+p_2}{q_1+q_2}} \prec S_{p_2/q_2} .$$

Since this construction generates all Sturmian sequences, by Theorem 3.4, $S_\varrho \prec S'_\varrho$ for all $\varrho < \varrho'$, $\varrho, \varrho' \in \mathbb{Q}$.

It is now enough to show that $S_{\tilde\varrho} \prec S_{\hat\varrho}$ for all $\tilde\varrho < \hat\varrho$, $\tilde\varrho \in \mathbb{Q}$, $\hat\varrho \notin \mathbb{Q}$. Consider a sequence $\{\varrho_i\}_{i=1}^\infty$ of rational numbers in $[0,1]$ such that $\tilde\varrho < \varrho_1 < \varrho_2 < \ldots < \varrho_i$, converging to $\hat\varrho \notin \mathbb{Q}$ as $i \to \infty$.

Sturmian measures with rational rotation number are totally ordered with respect to $\prec$, hence we have that $S_{\tilde\varrho} \prec S_{\varrho_1} \prec S_{\varrho_2} \prec \ldots \prec S_{\varrho_i}$. Then, since $\varrho \mapsto S_\varrho$ is continuous (in the vague topology, see [7]),

$$S_{\varrho_i}(x, \overline{1}] = \int_X \chi_{(x,\overline{1}]} dS_{\varrho_i} \xrightarrow{\varrho \to \hat\varrho} \int_X \chi_{(x,\overline{1}]} dS_{\hat\varrho} = S_{\hat\varrho}(x, \overline{1}]$$

for all $x \in X$. Thus, by Lemma 3.2, $S_{\tilde\varrho} \prec S_{\varrho_i} \prec S_{\hat\varrho}$, which completes the proof. □

Remark 5. This corollary can also be proved by utilising the characterisation of stochastic dominance by couplings. More precisely, one can consider the partial order $\leq_1$ on the shift $\{0,1\}^{\mathbb{N}}$, defined by

$$x_1 x_2 \ldots \leq_1 y_1 y_2 \ldots \iff \forall\, n \geq 0 \quad x_1 + x_2 + \ldots x_n \leq y_1 + y_2 + \ldots y_n$$

for sequences $x = x_1 x_2 \ldots$, $y = y_1 y_2 \ldots \in \{0,1\}^{\mathbb{N}}$ which, in particular, implies that $x \leq y$ for the lexicographic order. Using the couplings characterisation, it can be shown that $S_\varrho(f) \leq S_{\varrho'}(f)$ for $\varrho \leq \varrho'$, for all bounded measurable functions $f : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ which are increasing with respect to $\leq_1$ (which then implies $S_\varrho \prec S'_\varrho$).

The order $\leq_1$ is also studied in [6], where it is shown that the support of each Sturmian measure S_ϱ is totally ordered with respect to $\leq_1$.

5. On incomparability of measures

In this section some incomparability conditions will be presented which can be used to easily identify incomparable pairs of measures.

Proposition 5.1. *Consider atomic measures $\mu, \nu \in \mathcal{M}_X$ with $\mu \neq \nu$. If μ and ν have equal frequency, i.e. if $\varphi(\mu) = \varphi(\nu)$, then $\mu \curlywedge \nu$.*

First, a more general result will be proven.

Lemma 5.2. *Suppose $\hat{\mu}, \hat{\nu} \in \mathcal{M}_X$ are atomic measures, with $\hat{\mu} \neq \hat{\nu}$, $\hat{\mu} \prec \hat{\nu}$. Let $\hat{x}^{(1)} < \ldots < \hat{x}^{(q)}$ be the atoms of $\hat{\mu}$ where q is the period of the orbit, and let $J_i = [\hat{x}^{(i)}, \hat{x}^{(i+1)})$, $i \in \{1, \ldots, q-1\}$ and $J_q = [\hat{x}^{(q)}, \bar{1}]$. If $f : X \to \mathbb{R}$ is an increasing function such that its restriction to $\big(\mathrm{supp}(\hat{\mu}) \cup \mathrm{supp}(\hat{\nu})\big) \cap J_i$ is not a constant value for at least one value of $i \in \{1, \ldots, q\}$ for which $\mathrm{supp}(\hat{\nu}) \cap J_i \neq \emptyset$, then $\hat{\mu}(f) < \hat{\nu}(f)$.*

Proof. Let $\hat{y}^{(1)} < \ldots < \hat{y}^{(\hat{q})}$ be the atoms of $\hat{\nu}$, where $\hat{q}$ is the period of the orbit. We have the following

$$\begin{aligned}
\hat{\mu}(f) &= \int_X f(x) d\hat{\mu}(x) \\
&= f(\hat{x}^{(1)})\hat{\mu}[\hat{x}^{(1)}, \hat{x}^{(2)}) + \ldots + f(\hat{x}^{(q-1)})\hat{\mu}[\hat{x}^{(q-1)}, \hat{x}^{(q)}) + f(\hat{x}^{(q)})\hat{\mu}[\hat{x}^{(q)}, \bar{1}] \\
&= \sum_{i=1}^{q-1} f(\hat{x}^{(i)})\hat{\mu}[\hat{x}^{(i)}, \hat{x}^{(i+1)}) + f(\hat{x}^{(q)})\hat{\mu}[\hat{x}^{(q)}, \bar{1}] \\
&= \sum_{i=1}^{q-1} f(\hat{x}^{(i)})(\hat{\mu}[\hat{x}^{(i)}, \bar{1}] - \hat{\mu}[\hat{x}^{(i+1)}, \bar{1}]) + f(\hat{x}^{(q)})\hat{\mu}[\hat{x}^{(q)}, \bar{1}] \\
&= f(\hat{x}^{(1)})\hat{\mu}[\hat{x}^{(1)}, \bar{1}] + \sum_{i=1}^{q-1}(f(\hat{x}^{(i+1)}) - f(\hat{x}^{(i)}))\hat{\mu}[\hat{x}^{(i+1)}, \bar{1}] \\
&\leq f(\hat{x}^{(1)})\hat{\nu}[\hat{x}^{(1)}, \bar{1}] + \sum_{i=1}^{q-1}(f(\hat{x}^{(i+1)}) - f(\hat{x}^{(i)}))\hat{\nu}[\hat{x}^{(i+1)}, \bar{1}] \quad \text{by Lemma 3.2} \\
&= \sum_{i=1}^{q-1} f(\hat{x}^{(i)})\hat{\nu}[\hat{x}^{(i)}, \hat{x}^{(i+1)}) + f(\hat{x}^{(q)})\hat{\nu}[\hat{x}^{(q)}, \bar{1}] \, . \qquad (19)
\end{aligned}$$

It will now be shown that this last expression is strictly smaller than $\int f(x)d\hat{\nu}(x)$. Let $\alpha_i := \big| J_i \cap \mathrm{supp}(\hat{\nu}) \big|$, $i \in \{1, \ldots, q\}$. If i is such that $\alpha_i = 0$, then J_i does not contain any atoms of $\hat{\nu}$, hence $f(\hat{x}^{(i)})\hat{\nu}(J_i) = 0$. If i is such that $\alpha_i \neq 0$, denote the atoms of $\hat{(\nu)}$ that lie in J_i by $\hat{y}^{(k_i)} < \ldots < \hat{y}^{(k_i+\alpha_i-1)}$.

Then, since f is increasing,

$$f(\hat{x}^{(i)})\hat{\nu}(J_i) \leq f(\hat{y}^{(k_i)})\hat{\nu}(\{\hat{y}^{(k_i)}\}) + \ldots + f(\hat{y}^{(k_i+\alpha_i-1)})\hat{\nu}(\{\hat{y}^{(k_i+\alpha_i-1)}\}). \qquad (20)$$

By hypothesis f is not constant on $\big(\mathrm{supp}(\hat{\mu}) \cup \mathrm{supp}(\hat{\nu})\big) \cap J_i$ for at least one value of $i \in \{1, \ldots, q\}$, so for that value of i, the inequality (20) is strict. Thus by inequalities (19) and (20) and since all atoms of $\hat{\nu}$ lie in $[\hat{x}^{(1)}, \overline{1}]$,

$$\hat{\mu}(f) \leq \sum_{i=1}^{q} f(\hat{x}^{(i)})\hat{\nu}(J_i) < \sum_{j=1}^{\hat{q}} f(\hat{y}^{(j)})\hat{\nu}\{\hat{y}^{(j)}\} = \int_X f(x)d\hat{\nu}(x) = \hat{\nu}(f)$$

which completes the proof. □

Proof of Lemma 5.1. This will be proved by contradiction. Assume μ and ν are comparable, say $\mu \prec \nu$, without loss of generality. Define $f : X \to \mathbb{R}$ by $f(x) = \sum_{n=1}^{\infty} x_n 2^{-n}$ and observe that f is increasing. For any atomic measure $m \in \mathcal{M}_X$ with atoms $x^{(1)} < \ldots < x^{(\tilde{q})}$, $x^{(i)} = \{x_n^{(i)}\}_{n=1}^{\tilde{q}}$, where $\tilde{q}$ is the period of the orbit, the following holds:

$$\begin{aligned}
\int f(x)dm(x) &= \sum_{i=1}^{\tilde{q}} f(x^{(i)})m(\{x^{(i)}\}) = \frac{1}{\tilde{q}}\sum_{i=1}^{\tilde{q}} f(x^{(i)}) \\
&= \frac{1}{\tilde{q}}\sum_{i=1}^{\tilde{q}}\Big(\sum_{n=1}^{\infty} x_n^{(i)}2^{-n}\Big) \\
&= \frac{1}{\tilde{q}}\sum_{i=1}^{\tilde{q}}\Big(\sum_{n=1}^{\tilde{q}} x_n^{(i)}2^{-n}\big(\frac{2^{\tilde{q}}}{2^{\tilde{q}}-1}\big)\Big) \\
&= \frac{2^{\tilde{q}}}{\tilde{q}(2^{\tilde{q}}-1)}\sum_{n=1}^{\tilde{q}}\Big(\sum_{i=1}^{\tilde{q}} x_n^{(i)}2^{-n}\Big) \\
&= m\langle 1\rangle \frac{2^{\tilde{q}}}{2^{\tilde{q}}-1}\sum_{n=1}^{\tilde{q}} 2^{-n} = m\langle 1\rangle
\end{aligned}$$

Therefore, $\mu(f) = \mu\langle 1\rangle = \nu\langle 1\rangle = \nu(f)$. But f is increasing and clearly, its restriction to $\big(\mathrm{supp}(\hat{\mu}) \cup \mathrm{supp}(\hat{\nu})\big) \cap J_i$ is not a constant value for *any* value of $i \in \{1, \ldots, q\}$ for which $\mathrm{supp}(\hat{\nu}) \cap J_i \neq \emptyset$, so Lemma 5.2 implies $\mu(f) < \nu(f)$, which contradicts the assumption. Thus, μ and ν are incomparable. □

Lemma 5.3. *The following hold:*

(i) *Let $x, \hat{x}$ be periodic. If $x^+ < \hat{x}^+$ and $\hat{x}^- < x^-$, then $\mu_x \curlywedge \mu_{\hat{x}}$.*
(ii) *Let $x \in X$ be periodic, with $x \neq \overline{0}, \overline{1}$. Then $\mu_x \curlywedge \mu_{\overline{1w_x0}}$.*
(iii) *Let $x, \hat{x} \in X$ be periodic. If $\varphi(\mu_x) < \varphi(\mu_{\hat{x}})$ and $\hat{x}^+ < x^+$ or $\hat{x}^- < x^-$, then $\mu_x \curlywedge \mu_{\hat{x}}$.*
(iv) *Let $x, \hat{x} \in X$ be periodic of period $q, \hat{q}$ respectively, with $q < \hat{q}$.*
  (a) *If $x^+ < \hat{x}^+$ and $\sigma^i(\hat{x}^+) < x^+$ for all $i = 1, \ldots, \hat{q}-1$, then $\mu_x \curlywedge \mu_{\hat{x}}$.*
  (b) *If $\hat{x}^- < x^-$ and $x^- < \sigma^i(\hat{x}^-)$ for all $i = 1, \ldots, \hat{q}-1$, then $\mu_x \curlywedge \mu_{\hat{x}}$.*

Proof.
Proof of part (i). By hypothesis, $\mu_x[\hat{x}^+, \overline{1}] = 0 < \mu_{\hat{x}}[\hat{x}^+, \overline{1}]$ and $\mu_x[x^-, \overline{1}] = 1 > \mu_{\hat{x}}[x^-, \overline{1}]$. By Lemma 3.1, this implies $\mu_x \curlywedge \mu_{\hat{x}}$.

Proof of part (ii). We have that $\sigma^i(\overline{w_x01}) < \sigma^i(\overline{w_x}) \leq \overline{w_x} < \overline{1w_x0}$ for all $i = 0, \ldots, q$, where q is the period of x. This implies $\mu_x[\overline{1w_x0}, \overline{1}] = 0 < \mu_{\overline{1w_x0}}[\overline{1w_x0}, \overline{1}]$ and $\mu_x[x^+, \overline{1}] = \frac{1}{q} > \frac{1}{q+2} = \mu_{\overline{1w_x0}}[x^+, \overline{1}]$ which yields incomparability.

Proof of part (iii). The inequality $\varphi(\mu_x) < \varphi(\mu_{\hat{x}})$ implies $\mu_x(\langle 1\rangle) < \mu_{\hat{x}}(\langle 1\rangle)$ i.e.

$$\mu_x[1\overline{0}, \overline{1}] < \mu_{\hat{x}}[1\overline{0}, \overline{1}] \tag{21}$$

Now $\hat{x}^+ < x^+$ implies

$$\mu_x[x^+, \overline{1}] > 0 = \mu_{\hat{x}}[x^+, \overline{1}] \tag{22}$$

and $\hat{x}^- < x^-$ implies

$$\mu_x[x^-, \overline{1}] = 1 > \mu_{\hat{x}}[x^-, \overline{1}] \tag{23}$$

So (21) together with (22) or (23) prove the claim.

Proof of part (iv). The hypothesis implies $\mu_x[x^+, \overline{1}] = \frac{1}{q} > \frac{1}{\hat{q}} = \mu_{\hat{x}}[x^+, \overline{1}]$ and $\mu_x[\hat{x}^+, \overline{1}] = 0 < \frac{1}{\hat{q}} = \mu_{\hat{x}}[\hat{x}^+, \overline{1}]$, so by Lemma 3.1 the result follows. The proof of (b) is similar.

□

6. Fine structure

The preceding results in Sections 3 and 5 make it easy to establish dominance relations between atomic measures supported on periodic orbits of small period.

Proposition 6.1. *The dominance relations of measures on $\mathcal{M}_X$ supported on periodic orbits of period up to 7 are as shown in the Hasse diagram in Figure 4.*

Proof. There are 41 measures supported on periodic orbits of period up to 7, thus there are 820 dominance relations between them, out of which 189 are incomparability relations.

Colours are used to indicate which lemma or proposition was used to show each dominance relation. So, more precisely, green arrows were used to represent Lemma 3.3, orange for Theorem 3.4, blue for Lemma 3.5, black for Lemma 3.6 and red for Lemma 3.2.

Incomparability relations are all proved easily using Lemmas 5.1, 5.3, and 3.2. Due to the transitivity of first-order stochastic dominance, the Hasse diagram in Figure 4 represents *all* relations between measures supported on periodic orbit of up to period 7.

□

It is natural to ask if the partially ordered set $(\mathcal{M}_X, \prec)$ is a lattice, i.e. if any two measures have a first-order stochastically least upper bound and greatest lower bound. The following theorem shows that this is not the case.

Theorem 6.2. *The partially ordered set $(\mathcal{M}_X, \prec)$ is not a lattice.*

Proof. Consider the measures μ_{10} and μ_{1100}. We will show that these do not have a least upper bound. Let $\nu \in \mathcal{M}_X$ be such that $\mu_{10} \prec \nu$ and $\mu_{1100} \prec \nu$. Then, $0 = \mu_{10}\langle 00\rangle \geq \nu\langle 00\rangle$ hence $\nu\langle 00\rangle = 0$. Now $\varphi(\nu) = 1 - \nu\langle 0\rangle = 1 - \nu(\sigma^{-1}\langle 0\rangle) = 1 - \nu(\langle 00\rangle \cup \langle 10\rangle) = 1 - \nu\langle 10\rangle = 1 - (\nu\langle 1\rangle - \nu\langle 11\rangle) = 1 - \varphi(\nu) + \nu\langle 11\rangle$.

Hence,

$$\varphi(\nu) = \frac{1 + \nu\langle 11\rangle}{2} \geq \frac{1 + \mu_{1100}\langle 11\rangle}{2} = 5/8.$$

But by Lemma 3.2, the incomparable measures of frequency 5/8, $\mu_{11101010}$ and $\mu_{11011010}$, dominate both the measures μ_{10}, μ_{1100}. Consequently, μ_{10} and μ_{1100} do not have a least upper bound. □

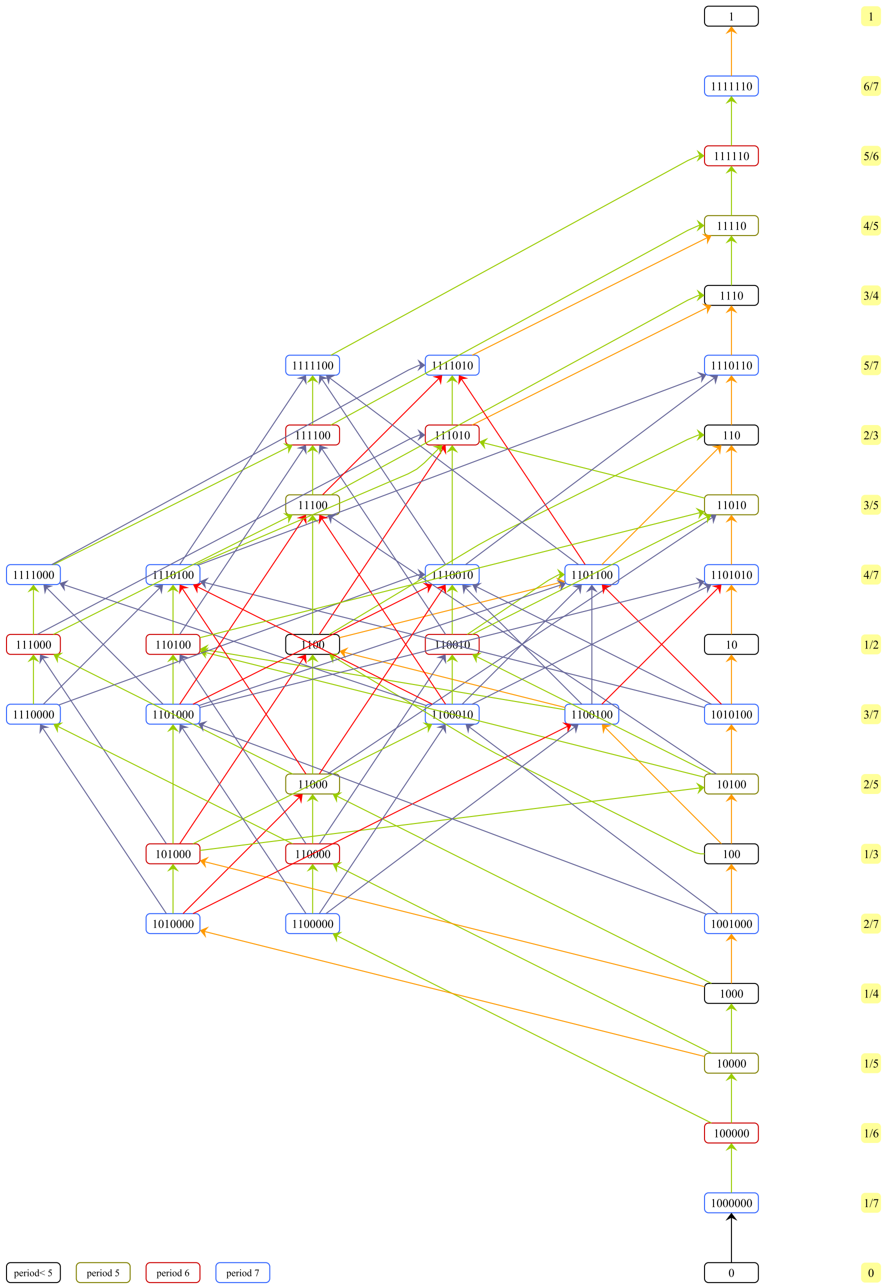

Figure 4. Hasse diagram of first-order stochastic dominance for measures supported on periodic orbits of period up to 7. Orbits that carry measures with equal frequency are displayed on the same horizontal line, and frequencies decrease from top to bottom.

References

[1] J.-P. Allouche and J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.

[2] V. Anagnostopoulou, Sturmian Measures and Stochastic Dominance in Ergodic Optimization, Ph.D. thesis, Queen Mary University of London, 2009.

[3] V. Anagnostopoulou and O. Jenkinson, Which beta-shifts have a largest invariant measure?, Jour. Lon. Math. Soc., **7**9 (2009), 445–464.

[4] J. Berstel and P. Séébold, Sturmian Words, in Algebraic Combinatorics on Words (Encyclopaedia of Mathematics and its Applications 90), (M. Lothaire), Cambridge University Press, (2002), 45–110.

[5] T. Bousch, Le poisson n'a pas d'arêtes, Ann. Inst. Henri Poincaré (Proba. et Stat.), **3**6 (2000), 489–508.

[6] T. Bousch, Une propriété de domination convexe pour les orbites sturmiennes, Can. Jour. Math.,**67** (2015), 90-106.

[7] T. Bousch and J. Mairesse, Asymptotic height optimization for topical IFS, tetris heaps, and the finiteness conjecture, Jour. Amer. Math. Soc., **15** (2002), 77–111.

[8] S. Bullett and P. Sentenac, Ordered orbits of the shift, square roots, and the devil's staircase, *Math. Proc. Camb. Phil. Soc.*, **115** (1994), 451–481.

[9] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 5^{th} edition, Oxford University Press, 1979.

[10] O. Jenkinson, Conjugacy Rigidity, Cohomological Triviality, and Barycentres of Invariant Measures, Ph.D. thesis, University of Warwick, 1996.

[11] O. Jenkinson, Frequency locking on the boundary of the barycentre set, Exp. Math., **9** (2000), 309–317.

[12] O. Jenkinson, Maximum hitting frequency and fastest mean return time, Nonlinearity, **1**8 (2005), 2305–2321.

[13] O. Jenkinson, Ergodic optimization, Discrete & Cont. Dyn. Sys., **15** (2006), 197–224.

[14] O. Jenkinson, Optimization and majorization of invariant measures, Electron. Res. Announc. Amer. Math. Soc., **13** (2007), 1–12.

[15] O. Jenkinson, A partial order on ×2-invariant measures, Math. Res. Lett., **15** (2008), 893–900.

[16] T. Kamae, U. Krengel and G. L. O'Brien, Stochastic inequalities on partially ordered spaces, Ann. Prob., **5** (1977), 899–912.

[17] H. Levy, Stochastic Dominance: Investment Decision Making under Uncertainty, 3rd edition, Springer, 2015.

[18] T. Lindvall, On Strassen's theorem on stochastic domination, Electron. Commun. Probab., **4** (1999), 51–59.

[19] M. Morse and G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, Amer. J. Math., **62** (1940), 1–42.

[20] K. Petersen, Some Sturmian symbolic dynamics, Available from: http://petersen.web.unc.edu/some-slides-from-talks/

[21] N. Pytheas Fogg, Substitutions in Dynamics, Arithmetics and Combinatorics, Springer Lecture Notes in Mathematics **1794**, 2002.

[22] A. Rényi, Representations of real numbers and their ergodic properties, Acta. Math. Acad. Sci. Hungar., **8** (1957), 477–493.

[23] V. Strassen, The existence of probability measures with given marginals, Ann. Math. Statist., **36** (1965), 423-439.

[24] P. Veerman, Symbolic dynamics of order-preserving orbits, Physica D, **2**9 (1987), 191–201.

E-mail address: v.anagnostopoulou@qmul.ac.uk