Statistical problems related to irrational rotations

Annals of the Institute of Statistical Mathematics vol. 58 (2006) pp.573-593

Teturo Kamae* and Hayato Takahashi †

Abstract

Let $\xi_i := \lfloor i\alpha+\beta\rfloor - \lfloor (i-1)\alpha+\beta\rfloor$ $(i = 1, 2, \cdots, m)$ be random variables as functions of β in the probability space $[0, 1)$ with the Lebesgue measure, where $\alpha \in [0, 1]$ is considered as an unknown parameter which we want to estimate from the observation $\xi := \xi_1\xi_2\cdots\xi_m$. Let an observation ξ be given, which is a finite Sturmian sequence. We determine the likelihood function $P_\alpha(\xi)$, the probability of getting ξ when the unknown parameter is α as a function of α, and obtain the maximal likelihood estimator $\hat{\alpha}(\xi)$ as the ratio of 1 in a minimal cycle of ξ, where a factor η of ξ is called a minimal cycle if ξ is a factor of η^∞ and η has the minimum length among them. We also obtain a minimum sufficient statistics. The sample mean $(\xi_1 + \xi_2 + \cdots + \xi_m)/m$ which is an unbiased estimator of α is not admissible if $m = 6$ or $m \geq 8$ since it is not based on the minimum sufficient statistics.

1 Introduction

Let $\xi = \xi_1\xi_2\cdots\xi_m$ be a finite 0-1-sequence. We denote the length m of ξ by $|\xi|$ and the number of 1 in ξ by $|\xi|_1$. We also denote $\rho(\xi) := |\xi|_1/|\xi|$, the ratio of 1 in ξ. Let $\xi = \xi_1\xi_2\cdots\xi_m$ and $\eta = \eta_1\eta_2\cdots\eta_n$ be finite 0-1-sequences. We say that η is a *factor* of ξ if there exists an integer i with $0 \leq i \leq m-n$ such that $\eta_j = \xi_{i+j}$ $(j = 1, 2, \cdots, n)$. In this case, we denote $\eta \prec \xi$. We say that η is a *prefix* of ξ if the above holds with $i = 0$.

A finite 0-1-sequence is called *nontrivial* if it contains both 0 and 1.

For a finite 0-1-sequence $\xi = \xi_1\xi_2\cdots\xi_m$, we denote by $\underline{\Omega}_\xi$ the set of $(\alpha, \beta) \in [0, 1] \times [0, 1)$ satisfying

$$\xi_i = \lfloor i\alpha+\beta\rfloor - \lfloor (i-1)\alpha+\beta\rfloor \quad (i = 1, 2, \cdots, m), \tag{1.1}$$

where $\lfloor\ \rfloor$ denotes the floor function, while $\lceil\ \rceil$ denotes the ceiling function. We also denote by $\overline{\Omega}_\xi$ the set of $(\alpha, \beta) \in [0, 1] \times (0, 1]$ satisfying

$$\xi_i = \lceil i\alpha+\beta\rceil - \lceil (i-1)\alpha+\beta\rceil \quad (i = 1, 2, \cdots, m). \tag{1.2}$$

Then, it holds for any finite 0-1-sequence ξ that $\underline{\Omega}_\xi \neq \emptyset$ if and only if $\overline{\Omega}_\xi \neq \emptyset$ (Lemma 1). We call ξ a (finite) *Sturmian* sequence if $\underline{\Omega}_\xi \neq \emptyset$. We denote by St_m the set of Sturmian sequences of length m.

*Department of Mathematics, Osaka City University, Osaka, 558-8585 Japan (kamae@sci.osaka-cu.ac.jp)

†Department of Statistical Science, The Graduate University for Advanced Studies, Tokyo, 106-8569 Japan (takahasi@ism.ac.jp)

An infinite sequence $\xi_1\xi_2\xi_3\cdots$ such that

$$\xi_i = \lfloor i\alpha + \beta \rfloor - \lfloor (i-1)\alpha + \beta \rfloor \ (\forall i) \quad (\text{or} \ \ \xi_i = \lceil i\alpha + \beta \rceil - \lceil (i-1)\alpha + \beta \rceil \ (\forall i))$$

for a fixed pair $(\alpha, \beta) \in [0,1] \times [0,1)$ with irrational α is called an (infinite) *Sturmian* sequence which is a symbolic representation of the rotation $\mathcal{R}_\alpha \theta = \theta + \alpha$ $(mod\ 1)$. It determines α as the relative frequency of 1 in ξ. Moreover, such sequences are characterized as the least complex sequences other than the eventually periodic sequences. Since the finite sequence ξ does not determine α, it becomes a problem of statistical inference how to estimate α from ξ under a suitable statistical model. Sturmian sequences appear in biological neuron model. In [3], Hata showed that aperiodic spike sequences generated by a single neuron model (Nagumo-Sato model [2]) are Sturmian.

The partition of $[0,1] \times [0,1)$ by Ω_ξ's for $m = 3$ is as follows, where the set Ω_ξ is denoted simply by ξ:

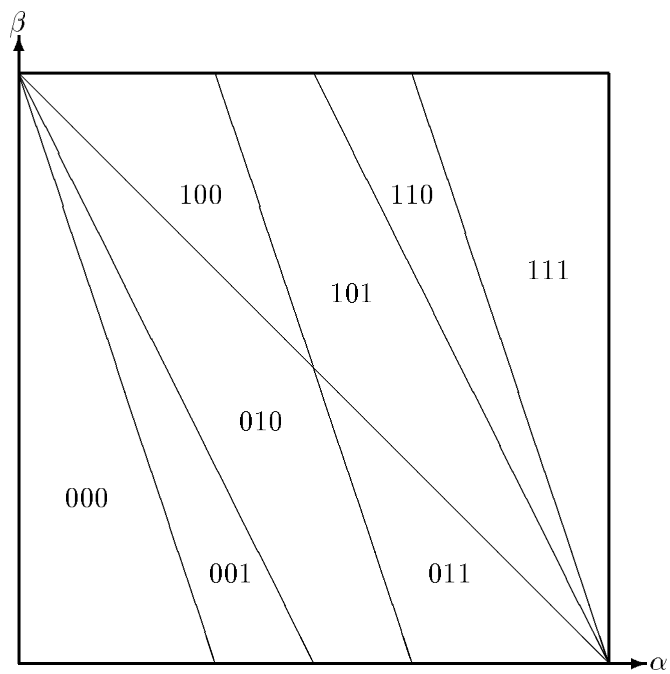

Figure 1: Partition by $\underline{\Omega}_\xi$'s for $m = 3$

We may consider $\xi = \xi_1\xi_2\cdots\xi_m$ as a random variable defined by (1.1) with random element β in the Lebesgue measure space $[0,1)$ and unknown parameter α in $[0,1]$. The sample space is St_m. As usual the probability (expectation, variance) under the parameter α is denoted by P_α (E_α, V_α, respectively). Thus, we have a statistical model $(\mathrm{St}_m, P_\alpha, \alpha \in [0,1])$.

By (1.1),

$$|\xi|_1 = \lfloor m\alpha + \beta \rfloor = \begin{cases} \lfloor m\alpha \rfloor & (\beta < 1 - \{m\alpha\}) \\ \lfloor m\alpha \rfloor + 1 & (\beta \geq 1 - \{m\alpha\}), \end{cases}$$

where { } denotes the fractional part. Hence, we have

$$
\begin{aligned}
&E_\alpha(\rho(\xi)) = (1/m)E_\alpha(|\xi|_1)\\
&= (1/m)(\lfloor m\alpha\rfloor(1-\{m\alpha\}) + (\lfloor m\alpha\rfloor+1)\{m\alpha\})\\
&= (1/m)(\lfloor m\alpha\rfloor + \{m\alpha\})\\
&= (1/m)(m\alpha) = \alpha \qquad (1.3)\\
&V_\alpha(\rho(\xi)) = (1/m^2)E_\alpha((|\xi|_1 - m\alpha)^2)\\
&= (1/m^2)((\lfloor m\alpha\rfloor - m\alpha)^2(1-\{m\alpha\}) + (\lfloor m\alpha\rfloor + 1 - m\alpha)^2\{m\alpha\})\\
&= (1/m^2)(\{m\alpha\}^2(1-\{m\alpha\}) + (1-\{m\alpha\})^2\{m\alpha\})\\
&= (1/m^2)\{m\alpha\}(1-\{m\alpha\}). \qquad (1.4)
\end{aligned}
$$

Therefore, the sample mean $\rho(\xi)$ is an unbiased estimator of α having the variance given by (1.4). It is not admissible if $m = 6$ or $m \geq 8$ under the quadratic loss function since it is not based on a minimum sufficient statistics (Theorem 1).

The following theorem is well known.

Theorem 1. (M. Morse, G.A. Hedlund [1]) *For any finite* 0-1*-sequence* ξ, ξ *is Sturmian if and only if it is balanced.*

Let m be a positive integer. Then, we have the partition (see Figure 1)

$$[0,1]\times[0,1) = \bigcup_{\xi\in\mathrm{St}_m} \underline{\Omega}_\xi \quad \text{(disjoint)}. \qquad (1.5)$$

This partition is discussed in Yasutomi [5] and Berstel & Pocchiola [4].

In [4], it is proved that for any finite Sturmian sequence ξ, the domain $\underline{\Omega}_\xi$ is surrounded by at most 4 pieces of line segments, at most 2 from above and at most 2 from below. Thus, there are only 3 cases as in Figure 2 for the shape of $\underline{\Omega}_\xi$.

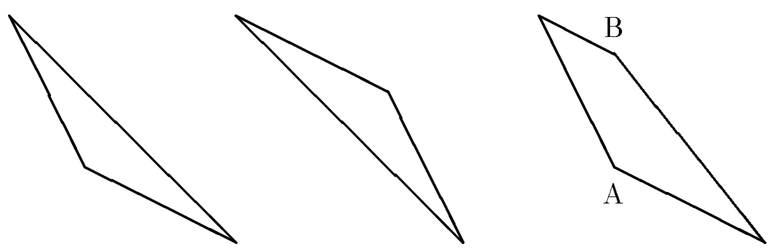

Figure 2: Shape of $\underline{\Omega}_\xi$

In this paper, we prove that in the 3rd case in Figure 2, the horizontal positions of A and B coincide (Lemma 7), which implies that the graph of $P_\alpha(\xi)$ with respect to α given ξ is of triangular shape as in Figure 3. The value $\hat{\alpha} = \hat{\alpha}(\xi)$ which maximize $P_\alpha(\xi)$ is the *maximal likelihood estimator*. That is,

$$P_{\hat{\alpha}}(\xi) = \max_{\alpha\in[0,1]} P_\alpha(\xi). \qquad (1.6)$$

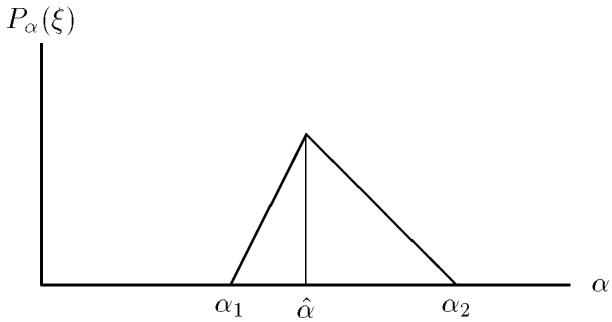

Figure 3: Likelihood function

In [6, 7], it is proved that $|\alpha_1 - \alpha_2| \le \frac{2m-1}{m(m-1)}$ for any $m = 2, 3, \cdots$ and $\xi \in \mathrm{St}_m$, where α_1 and α_2 are such that $[\alpha_1, \alpha_2]$ is the support of the likelihood function $P_\alpha(\xi)$ (Figure 3). This implies that the maximal likelihood estimator $\hat{\alpha}$ as a function of $(\alpha, \beta) \in [0,1] \times [0,1)$ and m through $\xi \in \mathrm{St}_m$ defined by (1.1) converge to α as $m \to \infty$ for any $\beta \in [0,1)$, which is a stronger property than the strong consistency, i.e. $\lim_{m\to\infty} \hat{\alpha} = \alpha$ ($P_\alpha - a.e.$).

Let $\xi = \xi_1 \xi_2 \cdots \xi_m$ be a 0-1-sequence. A positive integer p is called a *period* of ξ if

$$\xi_i = \xi_{i+p} \quad (i = 1, 2, \cdots, m-p). \tag{1.7}$$

This is equivalent to say that for any factor η of ξ with $|\eta| = p$, $\xi \prec \eta^\infty$ holds, where η^∞ implies the infinite time concatenation of η.

The minimum positive integer p as (1.7) is denoted by $\mathrm{per}(\xi)$. Note that $\mathrm{per}(\xi)$ exists always since $|\xi|$ is clearly a period of ξ. A factor η of ξ is called a *minimal cycle* of ξ if $|\eta| = \mathrm{per}(\xi)$. We define $\hat{\rho}(\xi) := \rho(\eta)$, where η is any minimal cycle of ξ.

Recall that a statistics $T = T(\xi)$ is called *sufficient* if for any $\xi \in \mathrm{St}_m$ and t, the conditional distribution $P_\alpha(\xi \mid T = t)$ does not depend on $\alpha \in [0,1]$ as long as $P_\alpha(T = t) > 0$.

This condition of sufficiency is equivalent to that for any ξ, $\xi' \in \mathrm{St}_m$, $T(\xi) = T(\xi')$ holds if and only if $\hat{\alpha}(\xi) = \hat{\alpha}(\xi')$, $\alpha_1(\xi) = \alpha_1(\xi')$ and $\alpha_2(\xi) = \alpha_2(\xi')$ holds (Figure 3). A sufficient statistics T is called a *minimum* sufficient statistics if for any sufficient statistics T', the partition on St_m induced by T' is finer than that induced by T. Note that a minimum sufficient statistics is unique in the sense of the partition induced on St_m. Clearly, the triple $(\hat{\alpha}, \alpha_1, \alpha_2)$ is a minimum sufficient statistics.

For $\xi \in \mathrm{St}_m$, we define

$$\begin{aligned} \underline{I}(\xi) &:= \left\{ i \in \{0, 1, \cdots, m\};\ \Xi_i - i\hat{\rho}(\xi) = \min_{0 \le j \le m} (\Xi_j - j\hat{\rho}(\xi)) \right\} \\ \overline{I}(\xi) &:= \left\{ i \in \{0, 1, \cdots, m\};\ \Xi_i - i\hat{\rho}(\xi) = \max_{0 \le j \le m} (\Xi_j - j\hat{\rho}(\xi)) \right\}, \end{aligned}$$

where $\Xi_i := \xi_1 + \cdots + \xi_i$ $(i = 0, 1, \cdots, n)$. The maximum or minimum value in $\overline{I}(\xi)$ or $\underline{I}(\xi)$ considered as a function of ξ is denoted by $\max \overline{I}(\xi)$, $\min \overline{I}(\xi)$, $\max \underline{I}(\xi)$ or $\min \underline{I}(\xi)$. We prove that $\max \overline{I}(\xi) - \min \underline{I}(\xi)$ is the slope of the left line segment and $\max \underline{I}(\xi) - \min \overline{I}(\xi)$ is minus of the slope of the right line segment in Figure 3.

In this paper, we prove the following theorem.

Theorem 2. *For the statistical model* $(\mathrm{St}_m, P_\alpha, \alpha \in [0,1])$ *with the quadratic loss function, we have*
(1) *The maximal likelihood estimator* $\hat{\alpha}$ *satisfies that* $\hat{\alpha}(\xi) = \hat{\rho}(\xi)$ *and the likelihood at* $\hat{\alpha}$ *satisfies that* $P_{\hat{\alpha}}(\xi) = 1/\mathrm{per}(\xi)$.
(2) *As for* α_1 *and* α_2 *in* Figure 3, *it holds for any nontrivial* $\xi \in \mathrm{St}_m$ *that*

$$\alpha_1 = \hat{\alpha} - \frac{1}{(\max \overline{I}(\xi) - \min \underline{I}(\xi))\mathrm{per}(\xi)}$$
$$\alpha_2 = \hat{\alpha} + \frac{1}{(\max \underline{I}(\xi) - \min \overline{I}(\xi))\mathrm{per}(\xi)}$$

(3) *The statistics* $(\hat{\rho},\ \max \overline{I} - \min \underline{I},\ \max \underline{I} - \min \overline{I})$ *is a minimum sufficient statistics.*
(4) *The sample mean* $\rho = \rho(\xi)$ *is not based on the minimum sufficient statistics and is not admissible if* $m = 6$ *or* $m \geq 8$.
(5) *The Bayes estimate* α_3 *with respect to the uniform prior distribution on* $\alpha \in [0,1]$ *is determined by*

$$\alpha_3 = \frac{\hat{\alpha} + \alpha_1 + \alpha_2}{3}$$

(6) *There is no UMVUE (Uniformly Minimum Variance Unbiased Estimator) for* α *if* $m \geq 3$.

Remark 1. For $m = 1, 2, 3, 4, 5$ and 7, the sample mean ρ is based on the above minimum sufficient statistics. But we do not know whether it is admissible or not except for rather trivial cases $m = 1, 2$ where ρ is admissible.

2 Prime segments

Let $\xi = \xi_1\xi_2 \cdots \xi_m$ be a 0-1-sequence. We denote

$$\Xi_0 := 0 \quad \text{and} \quad \Xi_i := \sum_{j=1}^{i} \xi_j \quad (i = 1, 2, \cdots, m). \tag{2.1}$$

Lemma 1. *Let* $\xi = \xi_1\xi_2 \cdots \xi_m$ *be a* 0-1-*sequence.*
(1) $\underline{\Omega}_\xi \neq \emptyset$ *if and only if* $\overline{\Omega}_\xi \neq \emptyset$.
(2) *For* $(\alpha, \beta) \in [0,1] \times [0,1)$, *the condition* (1.1) *is equivalent to the following condition:*

$$\Xi_i \leq i\alpha + \beta < \Xi_i + 1 \quad (i = 0, 1, \cdots, m). \tag{2.2}$$

(3) *For* $(\alpha, \beta) \in [0,1] \times (0,1]$, *the condition* (1.2) *is equivalent to the following condition:*

$$\Xi_i < i\alpha + \beta \leq \Xi_i + 1 \quad (i = 0, 1, \cdots, m). \tag{2.3}$$

(4) *If* ξ *is nontrivial, then* $\underline{\Omega}_\xi$ *is the set of* (α, β) *satisfying* (2.2).
(5) *If* ξ *is nontrivial, then* $\overline{\Omega}_\xi$ *is the set of* (α, β) *satisfying* (2.3).

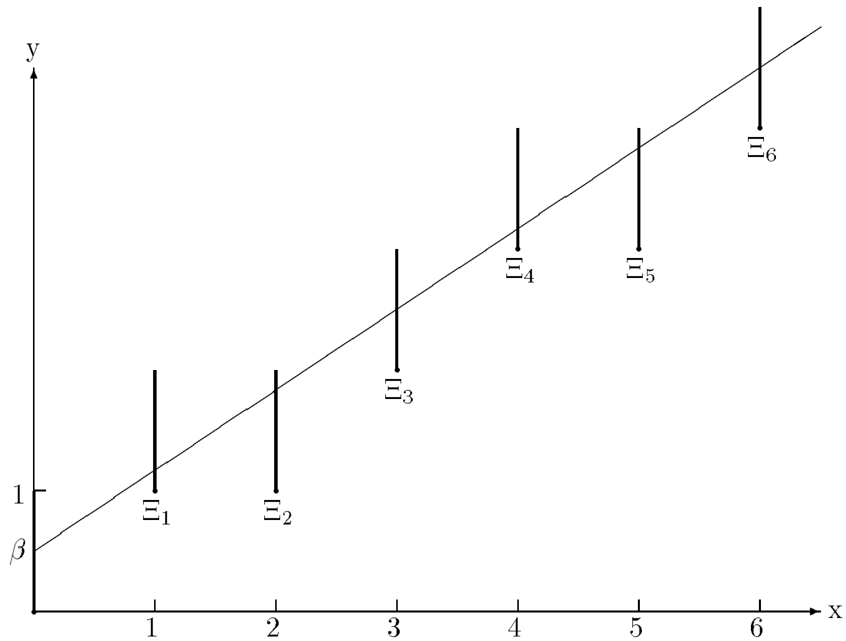

Figure 4: The graph of $y = \alpha x + \beta$ with $(\alpha, \beta) \in \underline{\Omega}_\xi$

Proof. (1) If $\underline{\Omega}_\xi \neq \emptyset$, then there exists $(\alpha, \beta) \in [0,1] \times [0,1)$ satisfying (1.1). Then, there exists β' with $\beta \leq \beta' < 1$ satisfying (1.1) such that $i\alpha + \beta'$ is not an integer for any $i = 0, 1, \cdots, m$. Then, we have

$$\xi_i = \lfloor i\alpha + \beta' \rfloor - \lfloor (i-1)\alpha + \beta' \rfloor = \lceil i\alpha + \beta' \rceil - \lceil (i-1)\alpha + \beta' \rceil.$$

Thus, $(\alpha, \beta') \in \overline{\Omega}_\xi$ and $\overline{\Omega}_\xi \neq \emptyset$. The converse is proved similarly.

(2) If $(\alpha, \beta) \in [0,1] \times [0,1)$ satisfies (1.1), then for $i = 1, 2, \cdots, m$, we have

$$\begin{aligned} \Xi_i &= \sum_{j=1}^{i} \xi_j \\ &= \sum_{j=1}^{i} (\lfloor j\alpha + \beta \rfloor - \lfloor (j-1)\alpha + \beta \rfloor) \\ &= \lfloor i\alpha + \beta \rfloor. \end{aligned}$$

Hence, we have (2.2).

Conversely, if $(\alpha, \beta) \in [0,1] \times [0,1)$ satisfies (2.2), then for $i = 1, 2, \cdots, m$, we have

$$\lfloor i\alpha + \beta \rfloor - \lfloor (i-1)\alpha + \beta \rfloor = \Xi_i - \Xi_{i-1} = \xi_i.$$

(3) The proof is similar to (2).

(4) Assume that (2.2) holds for (α, β). Then, we have $0 = \Xi_0 \leq \beta < \Xi_0 + 1 = 1$. Moreover, since ξ is nontrivial, we have $1 \leq \Xi_m \leq m-1$. Hence, $m\alpha \geq \Xi_m - \beta > 0$ and $m\alpha < \Xi_m + 1 - \beta \leq m$, so that $\alpha \in (0,1)$. Thus, $(\alpha, \beta) \in [0,1] \times [0,1)$. Then by (1), $(\alpha, \beta) \in \underline{\Omega}_\xi$. The converse follows from (1).

(5) The proof is similar to (4). $\square$

Let $\xi = \xi_1 \xi_2 \cdots \xi_m$ be a Sturmian sequence, which is fixed throughout this section. Let $\underline{\Gamma}_\xi$ and $\overline{\Gamma}_\xi$ be the minimal concave function and the maximal convex function, respectively, defined on the interval $[0, m]$ satisfying that

$$\underline{\Gamma}_\xi(i) \geq \Xi_i \quad (i = 0, 1, \cdots, m) \tag{2.4}$$

and

$$\overline{\Gamma}_\xi(i) \leq \Xi_i + 1 \quad (i = 0, 1, \cdots, m), \tag{2.5}$$

respectively. Clearly, they are piecewise linear functions such that $\underline{\Gamma}_\xi(0) = 0$, $\underline{\Gamma}_\xi(m) = \Xi_m$, $\overline{\Gamma}_\xi(0) = 1$, $\overline{\Gamma}_\xi(m) = \Xi_m + 1$ (Figure 5).

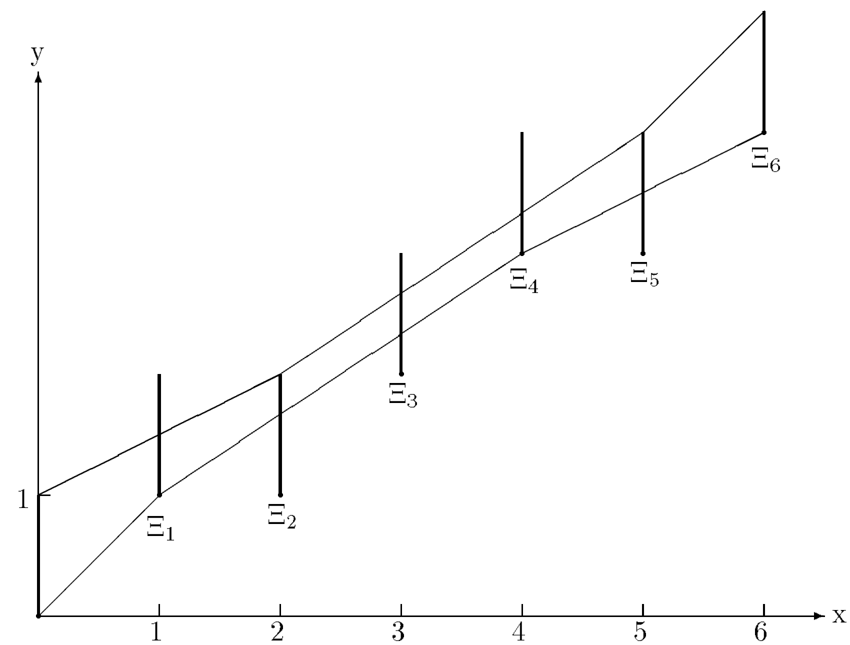

Figure 5: The graph of $\underline{\Gamma}_\xi$ and $\overline{\Gamma}_\xi$

A point (i, j) in $[0, m] \times [0, \infty)$ is called an *integer point* if both i and j are integers. A closed line segment AB with $A = (A_x, A_y), B = (B_x, B_y)$ and $A_x < B_x$ contained in the graph of $\overline{\Gamma}_\xi$ such that the set of integer points on AB is $\{A, B\}$ is called a *prime segment* of $\overline{\Gamma}_\xi$. A maximal closed line segment AB with $A_x < B_x$ contained in the graph of $\overline{\Gamma}_\xi$ is called a *maximal segment* of $\overline{\Gamma}_\xi$. In the same way, we define a *prime segment* of $\underline{\Gamma}_\xi$ and a *maximal segment* of $\underline{\Gamma}_\xi$.

Lemma 2. (1) *For any* $(\alpha,\beta)\in[0,1]\times[0,1)$, $(\alpha,\beta)\in\underline{\Omega}_\xi$ *holds if and only if* $\underline{\Gamma}_\xi(x)\le x\alpha+\beta<\overline{\Gamma}_\xi(x)$ *for any* $x\in[0,m]$. *In particular,* $\underline{\Gamma}_\xi(x)<\overline{\Gamma}_\xi(x)$ *holds for any* $x\in[0,m]$.
(2) *For any* $(\alpha,\beta)\in[0,1]\times(0,1]$, $(\alpha,\beta)\in\overline{\Omega}_\xi$ *holds if and only if* $\underline{\Gamma}_\xi(x)<x\alpha+\beta\le\overline{\Gamma}_\xi(x)$ *for any* $x\in[0,m]$.
(3) *Any integer point* A *on the graph of* $\underline{\Gamma}_\xi$ *(or* $\overline{\Gamma}_\xi$*) satisfies that* $A=(i,\Xi_i)$ *(or* $A=(i,\Xi_i+1)$*, respectively) for some* $i=0,1,\cdots,m$.
(4) *There is no integer point in the domain*

$$\{(x,y);\ 0\le x\le m,\ \underline{\Gamma}_\xi(x)<y<\overline{\Gamma}_\xi(x)\}.$$

(5) *For a maximal segment* CD *of* $\underline{\Gamma}_\xi$ *(or* $\overline{\Gamma}_\xi$*), both* C *and* D *are integer points. Moreover, there exists a positive integer* k *such that for any prime segment* AB *of* $\underline{\Gamma}_\xi$ *(or* $\overline{\Gamma}_\xi$*) contained in* CD*, we have* $\vec{CD}=k\vec{AB}$*, where* $\vec{AB}$ *implies the vector from* A *to* B.

Proof. (1) Let $(\alpha,\beta)\in\underline{\Omega}_\xi$. Then by Lemma 1, we have (2.2). Therefore, there exists $\beta'>\beta$ such that

$$\Xi_i\le i\alpha+\beta<i\alpha+\beta'\le\Xi_i+1\quad(i=0,1,\cdots,m).\tag{2.6}$$

Since the functions $x\alpha+\beta$ and $x\alpha+\beta'$ of $x\in[0,m]$ are concave and convex at the same time satisfying (2.6), we have

$$\underline{\Gamma}_\xi(x)\le x\alpha+\beta<x\alpha+\beta'\le\overline{\Gamma}_\xi(x)\tag{2.7}$$

for any $x\in[0,m]$ by the minimality and the maximality of $\underline{\Gamma}_\xi$ or $\overline{\Gamma}_\xi$, respectively.

Conversely, if $\underline{\Gamma}_\xi(x)\le x\alpha+\beta<\overline{\Gamma}_\xi(x)$ holds for any $x\in[0,m]$. Then, we have (2.2) since

$$\Xi_i\le\underline{\Gamma}_\xi(i)\le i\alpha+\beta<\overline{\Gamma}_\xi(i)\le\Xi_i+1\quad(i=0,1,\cdots,m).$$

Then by (2) of Lemma 1, we have $(\alpha,\beta)\in\underline{\Omega}_\xi$.

Finally, since there exists $(\alpha,\beta)\in\underline{\Omega}_\xi$, we have $\underline{\Gamma}_\xi(x)<\overline{\Gamma}_\xi(x)$ for any $x\in[0,m]$ by (2.7).

(2) The proof is same as (1).

(3) Let (i,j) be an integer point on $\underline{\Gamma}_\xi$. Then, by (1), we have

$$\Xi_i\le\underline{\Gamma}_\xi(i)=j<\overline{\Gamma}_\xi(i)\le\Xi_i+1,$$

which implies $j=\Xi_i$.

The proof is similar for $\overline{\Gamma}_\xi$.

(4) Suppose that there exists an integer point (i,j) satisfying that $0\le i\le m$ and $\underline{\Gamma}_\xi(i)<j<\overline{\Gamma}_\xi(i)$. Then, we have

$$\Xi_i\le\underline{\Gamma}_\xi(i)<j<\overline{\Gamma}_\xi(i)\le\Xi_i+1,$$

which is a contradiction since both Ξ_i and j are integers.

(5) If either C or D is not an integer point, then we can decrease the function $\underline{\Gamma}_\xi$ near the point C or D, respectively, keeping the concavity and the inequality (2.4), which contradicts with the minimality. Thus, C and D are integer points.

Let $C=(C_x,C_y)$, $D=(D_x,D_y)$ be their coordinates and let k be the greatest common divisor of D_x-C_x and D_y-C_y. Then, any prime segment AB contained in CD satisfies that $\vec{CD}=k\vec{AB}$ since B is the nearest integer point on CD to the right of A.

The proof is similsr for $\overline{\Gamma}_\xi$. $\square$

Let CD be a maximal segment of $\underline{\Gamma}_\xi$. We call CD *central* if there exists $(\alpha, \beta) \in \underline{\Omega}_\xi$ such that CD is on the graph $y = x\alpha + \beta$. Let CD be a maximal segment of $\overline{\Gamma}_\xi$. We call CD *central* if there exists $(\alpha, \beta) \in \overline{\Omega}_\xi$ such that CD is on the graph $y = x\alpha + \beta$. A prime segment CD of $\underline{\Gamma}_\xi$ (or $\overline{\Gamma}_\xi$) is called *central* if it is contained in a central maximal segment of $\underline{\Gamma}_\xi$ (or $\overline{\Gamma}_\xi$, respectively).

Lemma 3. (1) *Let AB be a central maximal segment of $\underline{\Gamma}_\xi$ (or $\overline{\Gamma}_\xi$), then we have $B_x > 2A_x$ and $2B_x - A_x > m$.*
(2) *A central maximal segment of $\underline{\Gamma}_\xi$ (or $\overline{\Gamma}_\xi$) is unique if it exists.*

Proof. (1) Suppose to the contrary that either $B_x \leq 2A_x$ or $2B_x - A_x \leq m$ holds. Without loss of generality, we assume $2B_x - A_x \leq m$. Define B' by $\vec{AB'} = 2\vec{AB}$. Then, we have $B'_x \leq m$.
Since by (5) of Lemma 2, both A and B are integer points, B' is also an integer points. Since AB is central, there exist $(\alpha, \beta) \in \underline{\Omega}_\xi$ such that AB is on the graph $y = x\alpha + \beta$. Since B' is also on this graph, we have $\Xi_{B'_x} \leq B'_y < \Xi_{B'_x} + 1$. Since B' is above the graph $y = \underline{\Gamma}_\xi(x)$, we have $B'_y > \underline{\Gamma}_\xi(B'_x) \geq \Xi_{B'_x}$. Thus, $\Xi_{B'_x} < B'_y < \Xi_{B'_x} + 1$. This is a contradiction since both $\Xi_{B'_x}$ and B'_y are integers.
(2) Suppose that there exist 2 distinct central maximal segments AB and CD of $\underline{\Gamma}_\xi$. Assume that $B_x \leq C_x$. Since $0 \leq A_x < B_x \leq C_x < D_x \leq m$, either $B_x \leq m/2$ or $C_x \geq m/2$. This implies that either $2B_x - A_x \leq m$ or $D_x \leq 2C_x$, which contradicts with (1). $\square$

Lemma 4. *At least one of the central maximal segment of $\overline{\Gamma}_\xi$ or the central maximal segment of $\underline{\Gamma}_\xi$ exists.*

Proof. Assume that the central maximal segment of $\overline{\Gamma}_\xi$ does not exist. By (1) of Lemma 2, there exists $(\alpha, \beta) \in \underline{\Omega}_\xi$ such that $\underline{\Gamma}_\xi(x) \leq x\alpha + \beta < \overline{\Gamma}_\xi(x)$ for any $x \in [0, m]$. Fixing α, we increase $\beta \in [0, 1)$ until $\overline{\Gamma}_\xi(x) = x\alpha + \beta$ holds for some $x \in [0, m]$ for the first time. If the equality holds for more than 1 point, then the equality holds for a maximal segment of $\overline{\Gamma}_\xi$, which is central since it is on the graph $y = x\alpha + \beta$ with $(\alpha, \beta) \in \overline{\Omega}_\xi$, which contradicts with our assumption.
Hence,

$$\underline{\Gamma}_\xi(x) < x\alpha + \beta \leq \overline{\Gamma}_\xi(x) \tag{2.8}$$

holds for any $x \in [0, m]$ with the equality $x\alpha + \beta = \overline{\Gamma}_\xi(x)$ for just one point $x = x_0$. Then, $A = (x_0, \overline{\Gamma}_\xi(x_0)) = (x_0, x_0\alpha + \beta)$ is an integer point since it must be a broken point of the piecewise linear graph $y = \overline{\Gamma}_\xi(x)$ by the uniqueness.
Let α increase and β decrease keeping the value $x_0\alpha + \beta$ invariant until the graph $y = x\alpha + \beta$ touch the graph $y = \overline{\Gamma}_\xi(x)$ for the first time. By the above argument with our assumption, the graph $y = x\alpha + \beta$ does not touch the graph $y = \overline{\Gamma}_\xi(x)$ at a different point from A before it touch the graph $y = \underline{\Gamma}_\xi(x)$ for the first time. Let the value (α, β) when it touch the graph $y = \underline{\Gamma}_\xi(x)$ for the first time be (α_1, β_1). Then we have

$$\underline{\Gamma}_\xi(x) \leq x\alpha_1 + \beta_1 \leq \overline{\Gamma}_\xi(x) \tag{2.9}$$

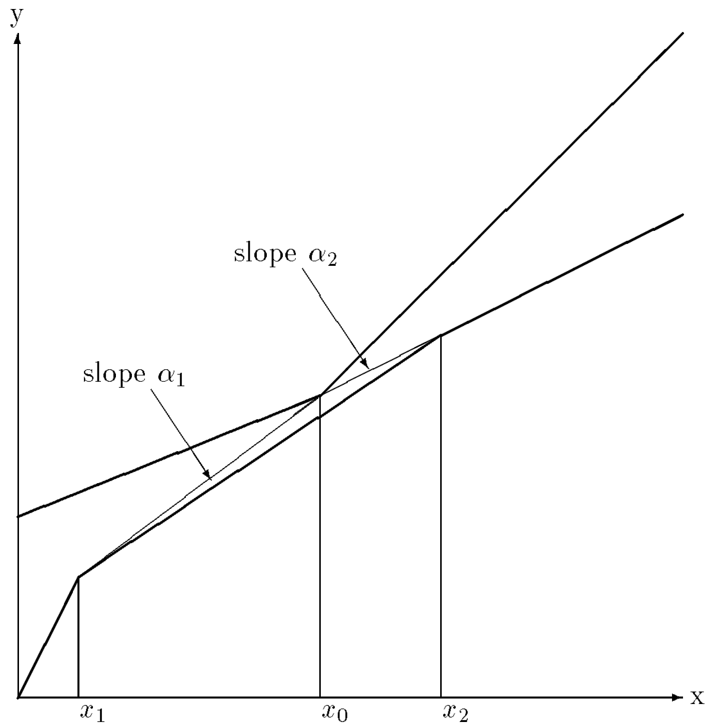

Figure 6:

for any $x \in [0, m]$, and $\underline{\Gamma}_\xi(x) = x\alpha_1 + \beta_1$ holds for some $x \in [0, m]$, say x_1, while $x\alpha_1 + \beta_1 = \overline{\Gamma}_\xi(x)$ holds if and only if $x = x_0$ (Figure 6).

Starting again from (3.5), let α decrease and β increase keeping the value $x_0\alpha + \beta$ invariant until the graph $y = x\alpha + \beta$ touch the graph $y = \underline{\Gamma}_\xi(x)$ for the first time and the value (α, β) when it touch the graph $y = \underline{\Gamma}_\xi(x)$ for the first time be (α_2, β_2). Then we have

$$\underline{\Gamma}_\xi(x) \leq x\alpha_2 + \beta_2 \leq \overline{\Gamma}_\xi(x) \tag{2.10}$$

for any $x \in [0, m]$, and $\underline{\Gamma}_\xi(x) = x\alpha_2 + \beta_2$ holds for some $x \in [0, m]$, say x_2, while $x\alpha_2 + \beta_2 = \overline{\Gamma}_\xi(x)$ holds if and only if $x = x_0$.

Then, we have $x_1 < x_0 < x_2$ and

$$\underline{\Gamma}_\xi(x_0) < \overline{\Gamma}_\xi(x_0) = x_0\alpha_1 + \beta_1 = x_0\alpha_2 + \beta_2.$$

Since

$$\frac{\underline{\Gamma}_\xi(x_0) - \underline{\Gamma}_\xi(x_1)}{x_0 - x_1} < \frac{\overline{\Gamma}_\xi(x_0) - \underline{\Gamma}_\xi(x_1)}{x_0 - x_1} = \alpha_1,$$

there exists $x_3 \in (x_1, x_0)$ such that $\underline{\Gamma}_\xi{}'(x_3) < \alpha_1$. In the same way, there exists $x_4 \in (x_0, x_2)$ such that $\underline{\Gamma}_\xi{}'(x_4) > \alpha_2$. Therefore, we have $x_1 < x_3 < x_4 < x_2$ and

$$\underline{\Gamma}_\xi{}'(x_1 - 0) \geq \alpha_1 > \underline{\Gamma}_\xi{}'(x_3) \geq \underline{\Gamma}_\xi{}'(x_4) > \alpha_2 \geq \underline{\Gamma}_\xi{}'(x_2 + 0).$$

Hence, there exists a maximal segment BC of $\underline{\Gamma}_\xi$ with the slope $\alpha_3 := \underline{\Gamma}_\xi{}'(x_3)$ satisfying that $\alpha_1 > \alpha_3 > \alpha_2$ and $x_1 \leq B_x < C_x \leq x_2$.

We prove that BC is central. Let $y = x\alpha_3 + \beta_3$ be the graph which contains BC. Then, by (1) of Lemma 2, it is sufficient to prove that

$$\underline{\Gamma}_\xi(x) \le x\alpha_3 + \beta_3 < \overline{\Gamma}_\xi(x) \quad (\forall x \in [0, m]).$$

Since $\underline{\Gamma}_\xi(x) \le x\alpha_3 + \beta_3$ holds for any $x \in [0, m]$ by the concavity of $\underline{\Gamma}_\xi$, it is sufficient to prove that

$$x\alpha_3 + \beta_3 < \overline{\Gamma}_\xi(x) \quad (\forall x \in [0, m]). \tag{2.11}$$

Since BC is below the graph $y = x\alpha_1 + \beta_1$ except possibly for B and $\alpha_3 < \alpha_1$, we have

$$x\alpha_3 + \beta_3 < x\alpha_1 + \beta_1 \le \overline{\Gamma}_\xi(x)$$

for any $x \in (B_x, m]$. In the same way, we have

$$x\alpha_3 + \beta_3 < x\alpha_2 + \beta_2 \le \overline{\Gamma}_\xi(x)$$

for any $x \in [0, C_x)$. Thus, we have (2.11). □

Lemma 5. *Assume that a central prime segment of* $\underline{\Gamma}_\xi$ *and a central prime segment of* $\overline{\Gamma}_\xi$ *exist at the same time. Then, their lengths and slopes coincide.*

Proof. Let AB and CD be the central prime segments of $\underline{\Gamma}_\xi$ and of $\overline{\Gamma}_\xi$, respectively. Suppose that either their slops or lengths do not coincide.

Let $y = x\alpha + \beta$ and $y = x\alpha' + \beta'$ be the graphs containing AB or CD, respectively. Let $x_1 = A_x \wedge C_x$ and $x_2 = B_x \vee D_x$. Denote

$$\Lambda := \{(x, y);\ x_1 \le x \le x_2,\ x\alpha + \beta \le y \le x\alpha' + \beta'\}.$$

Since both AB and CD are central, $\Lambda \setminus (AB \cup CD)$ is contained in the domain $\{(x, y);\ 0 \le x \le m,\ \underline{\Gamma}_\xi(x) < y < \overline{\Gamma}_\xi(x)\}$ and has no integer point by (4) of Lemma 2.

Assume that the slopes of AB and CD do not coincide. Without loss of generality, assume that the slope of AB is less than the slope of CD. Let F be the point such that $\vec{DF} = \vec{CA}$ if $A_x \le C_x$, and $\vec{BF} = \vec{AC}$ if $A_x > C_x$. Then, F is an integer point since A, B, C, D are integer points. Moreover, we have $F \in \Lambda \setminus (AB \cup CD)$, which contradicts with that $\Lambda \setminus (AB \cup CD)$ has no integer point.

Assume that AB and CD have the same slope but different lengths, say $|\vec{AB}| < |\vec{CD}|$. Let F be the point such that $\vec{BF} = \vec{AC}$. Then, F is an integer point since A, B, C, D are integer points. Moreover, F is in the interior of the line segment CD, which contradicts with that CD is a prime segment of $\overline{\Gamma}_\xi$.

Thus, we have $\vec{AB} = \vec{CD}$, which implies that a central prime segment of $\underline{\Gamma}_\xi$ and a central prime segments of $\overline{\Gamma}_\xi$ have the same length and slope. □

3 Shape of $\underline{\Omega}_\xi$

Let $\xi = \xi_1\xi_2\cdots\xi_m$ be a Sturmian sequence, which we fix throughout this section. In [4], the duality between the domains

$$\begin{aligned} \underline{\Sigma}_\xi &:= \{(x,y);\ 0 \le x \le m,\ \underline{\Gamma}_\xi(x) \le y < \overline{\Gamma}_\xi(x)\} \\ \overline{\Sigma}_\xi &:= \{(x,y);\ 0 \le x \le m,\ \underline{\Gamma}_\xi(x) < y \le \overline{\Gamma}_\xi(x)\} \end{aligned} \tag{3.1}$$

and the domains $\underline{\Omega}_\xi$ and $\overline{\Omega}_\xi$ is discussed, although the notations there are slightly different from ours. We restate the results there in our framework.

For $(x,y) \in [0,m]\times[0,\infty)$, denote

$$\begin{aligned} (x,y)^* &= \{(\alpha,\beta) \in [0,1]\times[0,1];\ x\alpha+\beta = y\} \\ (x,y)^+ &= \{(\alpha,\beta) \in [0,1]\times[0,1];\ x\alpha+\beta < y\} \\ (x,y)^- &= \{(\alpha,\beta) \in [0,1]\times[0,1];\ x\alpha+\beta > y\} \\ (x,y)^{*+} &= (x,y)^+ \cup (x,y)^* \\ (x,y)^{*-} &= (x,y)^- \cup (x,y)^*. \end{aligned}$$

Thus, $(x,y)^*$ is a straight line in the domain $[0,1]\times[0,1]$. Conversely, for $(\alpha,\beta) \in [0,1]\times[0,1]$, denote

$$(\alpha,\beta)^* = \{(x,y) \in [0,m]\times[0,\infty);\quad x\alpha+\beta = y\},$$

so that $(\alpha,\beta)^*$ is a straight line in $[0,m]\times[0,\infty)$. For a subset S of $[0,m]\times[0,\infty)$ or a subset T of $[0,1]\times[0,1]$, we denote

$$S^* = \bigcap_{(x,y)\in S} (x,y)^*$$

and

$$T^* = \bigcap_{(\alpha,\beta)\in T} (\alpha,\beta)^*.$$

Then we have the duality that $((x,y)^*)^* = (x,y)$ and $((\alpha,\beta)^*)^* = (\alpha,\beta)$ for any $(x,y) \in [0,m]\times[0,\infty)$ and $(\alpha,\beta) \in [0,1]\times[0,1]$.

Lemma 6. *We have*

$$\underline{\Omega}_\xi = \bigcap_{i=0}^{m}(i,\Xi_i)^{*+} \cap \bigcap_{i=0}^{m}(i,\Xi_i+1)^{-}.$$

and

$$\overline{\Omega}_\xi = \bigcap_{i=0}^{m}(i,\Xi_i)^{+} \cap \bigcap_{i=0}^{m}(i,\Xi_i+1)^{*-}.$$

Hence, $\underline{\Omega}_\xi$ *(or* $\overline{\Omega}_\xi$*) is a convex domain surrounded by a finite number of line segments with nonpositive slopes.*

Proof. Clear from Lemma 1. □

Let $\underline{\Omega_\xi}^{cl}$ be the closure of $\underline{\Omega_\xi}$, which is a compact convex set surrounded by line segments with nonpositive slopes. Let $ex(\underline{\Omega_\xi}^{cl})$ be the set of extremal points of $\underline{\Omega_\xi}^{cl}$. Take a point on a boundary of $\underline{\Omega_\xi}^{cl}$ and move it around the boundary counterclockwise. Then, the direction of the movement changed when it arrives at extremal points. There is a unique extremal point such that the horizonal component of the direction changes from negative to positive at this point, which is called the *left vertex*. Also, there is a unique extremal point such that the horizonal component of the direction changes from positive to negative at this point, which is called the *right vertex*. The other extremal points are either upper or lower as defined below.

An *upper edge* of $\underline{\Omega_\xi}$ is defined as $\overline{\Omega}_\xi \cap (i, \Xi_i + 1)^*$ for some $i = 0, 1, \cdots, m$ if it contains at least 2 points. A *lower edge* of $\underline{\Omega_\xi}$ is defined as $\underline{\Omega_\xi} \cap (i, \Xi_i)^*$ for some $i = 0, 1, \cdots, m$ if it contains at least 2 points. An *upper vertex* is an intersection of 2 distinct upper edges belonging to $\overline{\Omega}_\xi$. A *lower vertex* is an intersection of 2 distinct upper edges belonging to $\underline{\Omega_\xi}$.

Lemma 7. (1) *For $Q \in [0,1] \times (0,1]$, Q is an upper vertex of $\underline{\Omega_\xi}$ if and only if Q^* contains the central maximal segment of $\overline{\Gamma}_\xi$. In this case, the slope of the central maximal segment of $\overline{\Gamma}_\xi$ is Q_α, where $Q = (Q_\alpha, Q_\beta)$. Moreover, an upper vertex is unique if it exists.*
(2) *For $Q \in [0,1] \times [0,1)$, Q is a lower vertex of $\underline{\Omega_\xi}$ if and only if Q^* contains the central maximal segment of $\underline{\Gamma_\xi}$. In this case, the slope of the central maximal segment of $\underline{\Gamma_\xi}$ is Q_α. Moreover, a lower vertex is unique if it exists.*
(3) *Either the upper vertex or the lower vertex of $\underline{\Omega_\xi}$ exists. Moreover, if both of the upper vertex P and the lower vertex Q exist, then we have $P_\alpha = Q_\alpha$.*
(4) *If the central maximal segment of $\underline{\Gamma_\xi}$ or the central maximal segment of $\overline{\Gamma}_\xi$ exists, then their slopes coincide with the maximal likelihood estimator $\hat{\alpha}(\xi)$* (Figure 3).

Proof. (1) Let $Q \in \overline{\Omega}_\xi$ be an upper vertex of $\underline{\Omega_\xi}$. It is the intersection of 2 distinct upper edges, say PQ and QR, where $P_\alpha < Q_\alpha < R_\alpha$. Since PQ is an upper edge, there exists $i = 0, 1, \cdots, m$ such that the graph $y = x\alpha + \Xi_i + 1 - i\alpha$ in $[0, m] \times [0, \infty)$ passing $(i, \Xi_i + 1)$ is contained in the domain $\overline{\Sigma}_\xi$ for any α with $P_\alpha \leq \alpha \leq Q_\alpha$. Also, there there exists $j = 0, 1, \cdots, m$ such that the graph $y = x\alpha + \Xi_j + 1 - j\alpha$ in $[0, m] \times [0, \infty)$ passing $(j, \Xi_j + 1)$ is contained in the domain $\overline{\Sigma}_\xi$ for any α with $Q_\alpha \leq \alpha \leq R_\alpha$.

For $\alpha = Q_\alpha$, the graphs $y = x\alpha + \Xi_i + 1 - i\alpha$ and $y = x\alpha + \Xi_j + 1 - j\alpha$ should coincide, since otherwise, one of them is above the point $(i, \Xi_i + 1)$ or $(j, \Xi_j + 1)$ and is not in the domain $\overline{\Sigma}_\xi$, which is a contradiction. Therefore, those graphs coincide and pass both points $(i, \Xi_i + 1)$ and $(j, \Xi_j + 1)$. Since this graph is contained in $\overline{\Sigma}_\xi$, these points are on the central maximal segment of $\overline{\Gamma}_\xi(x)$ with the slope Q_α.

Conversely, if Q^* contains the central maximal segment of $\overline{\Gamma}_\xi(x)$, say AB. Then, the graph $y = xQ_\alpha + Q_\beta$ contains AB and is contained in the domain $\overline{\Sigma}_\xi$. Starting from $\alpha = Q_\alpha$ and $\beta = Q_\beta$, we can decrease α and increase β keeping the graph in the domain $\overline{\Sigma}_\xi$ and keeping the equation $A_x\alpha + \beta = A_y$. Hence, $\overline{\Omega}_\xi \cap A^*$ contains at least 2 points including Q. The same thing holds for $\overline{\Omega}_\xi \cap B^*$. Hence, they are distinct upper edges whose intersection is Q. Thus, Q is an upper edge.

(2) The proof is similar to (1).

(3) It follows from (1), (2) and Lemma 5.

(4) Clear from (1), (2) and the shape of the domain $\underline{\Omega}_\xi$ (Figure 2). $\square$

4 Main results

Let $\xi = \xi_1\xi_2\cdots\xi_m$ be a Sturmian sequence, which we fix throughout this section.

For a pair of positive integers (u, v) with $v \le u$, we denote by $\underline{\lambda}(u, v)$ the 0-1-sequence of length u such that

$$\underline{\lambda}(u, v)_i = \lfloor iv/u \rfloor - \lfloor (i-1)v/u \rfloor \quad (i = 1, 2, \cdots, u),$$

and by $\overline{\lambda}(u, v)$ the 0-1-sequence of length u such that

$$\overline{\lambda}(u, v)_i = \lceil iv/u \rceil - \lceil (i-1)v/u \rceil \quad (i = 1, 2, \cdots, u).$$

For integers i, j with $0 \le i < j \le m$, we denote

$$\xi[i, j] := \xi_{i+1}\xi_{i+2}\cdots\xi_j.$$

Lemma 8. (1) *Let AB be a prime segment of $\underline{\Gamma}_\xi$. Then, we have $A_y = \Xi_{A_x}$, $B_y = \Xi_{B_x}$ and $\xi[A_x, B_x] = \underline{\lambda}(B_x - A_x, B_y - A_y)$. Moreover, $B_x - A_x$ and $B_y - A_y$ are coprime.*
(2) *Let AB be a prime segment of $\overline{\Gamma}_\xi$. Then, we have $A_y = \Xi_{A_x} + 1$, $B_y = \Xi_{B_x} + 1$ and $\xi[A_x, B_x] = \overline{\lambda}(B_x - A_x, B_y - A_y)$. Moreover, $B_x - A_x$ and $B_y - A_y$ are coprime.*
(3) *Let AB be a prime segment of $\underline{\Gamma}_\xi$ ($\overline{\Gamma}_\xi$). It is central if and only if $B_x - A_x$ is a period of ξ. Moreover, $B_x - A_x = \mathrm{per}(\xi)$ holds in this case.*

Proof. (1) Let AB be a prime segment of $\underline{\Gamma}_\xi$. Since A and B are integer points on the graph $y = \underline{\Gamma}_\xi(x)$, we have $A_y = \Xi_{A_x}$ and $B_y = \Xi_{B_x}$ by (3) of Lemma 2. Since AB is in the domain $\underline{\Sigma}_\xi$, (i, Ξ_i) is on or below AB and $(i, \Xi_i + 1)$ is above AB for any i with $A_x \le i \le B_x$. Therefore, we have

$$\Xi_i \le (i - A_x)(B_y - A_y)/(B_x - A_x) + A_y < \Xi_i + 1, \tag{4.1}$$

or equivalently,

$$\Xi_i = \lfloor (i - A_x)(B_y - A_y)/(B_x - A_x) \rfloor + A_y \tag{4.2}$$

for any i with $A_x \le i \le B_x$. Hence, we have

$$\begin{aligned} \xi_{A_x+i} &= \Xi_{A_x+i} - \Xi_{A_x+i-1} \\ &= \lfloor i(B_y - A_y)/(B_x - A_x) \rfloor - \lfloor (i-1)(B_y - A_y)/(B_x - A_x) \rfloor \\ &= \underline{\lambda}(B_x - A_x, B_y - A_y)_i \end{aligned}$$

for $i = 1, 2, \cdots, B_x - A_x$, and hence $\xi[A_x, B_x] = \underline{\lambda}(B_x - A_x, B_y - A_y)$.

That $B_x - A_x$ and $B_y - A_y$ are coprime follows from the fact that AB contains no integer point other than the end points A and B.

(2) The proof is similar with the proof in (1).

(3) Let AB be a central prime segment of $\underline{\Gamma}_\xi$. Then, the graph $y = x\alpha + \beta$ $(x \in [0, m])$ which contains AB is in $\underline{\Sigma}_\xi$, so that for any i with $0 \le i \le m$, (i, Ξ_i) is on or below the graph

and $(i, \Xi_i + 1)$ is above the graph. Therefore, we have (4.1), and hence, (4.2) for any i with $0 \leq i \leq m$. This implies that

$$\Xi_{i+B_x-A_x} = \Xi_i + B_y - A_y \quad (0 \leq i \leq m - (B_x - A_x)),$$

and hence,

$$\xi_{i+B_x-A_x} = \xi_i \quad (1 \leq i \leq m - (B_x - A_x)).$$

Thus, $B_x - A_x$ is a period of ξ.

Conversely, let AB be a prime segment of $\underline{\Gamma}_\xi$ such that $B_x - A_x$ is a period of ξ. Then, we have

$$\xi_{i+B_x-A_x} = \xi_i \quad (1 \leq i \leq m - (B_x - A_x)).$$

Since $\Xi_{A_x} = A_y$ and $\Xi_{B_x} = B_y$, it follows that

$$\sum_{i=1}^{B_x-A_x} \xi_{\ell+i} = B_y - A_y$$

for any ℓ with $1 \leq \ell \leq m - (B_x - A_x)$. Moreover since $\xi[A_x, B_x] = \underline{\lambda}(B_x - A_x, B_y - A_y)$, we have

$$\begin{aligned}
\Xi_i &= A_y + \sum_{h=1}^{j} \underline{\lambda}(B_x - A_x, B_y - A_y)_h + k(B_y - A_y) \\
&= A_y + \lfloor j(B_y - A_y)/(B_x - A_x) \rfloor + k(B_y - A_y) \\
&= \lfloor (j + k(B_x - A_x))(B_y - A_y)/(B_x - A_x) + A_y \rfloor \\
&= \lfloor (i - A_x)(B_y - A_y)/(B_x - A_x) + A_y \rfloor
\end{aligned}$$

for any i with $0 \leq i \leq m$, where j and k are integers with $1 \leq j \leq B_x - A_x$ and $i = A_x + j + k(B_x - A_x)$. Therefore, the graph $y = (x - A_x)(B_y - A_y)/(B_x - A_x) + A_y$ $(x \in [0, m])$ which contains AB is in the domain $\underline{\Sigma}_\xi$. Thus, AB is central.

Let AB be a prime segment of $\underline{\Gamma}_\xi$. Suppose that $p := \mathrm{per}(\xi) < B_x - A_x$. Let $B_x - A_x = p + r$ with a positive integer r. Since (i, Ξ_i) is below the line segment AB for any i with $A_x < i < B_x$, $(A_x + p, \Xi_{A_x+p})$ and $(A_x + r, \Xi_{A_x+r})$ are below AB. Hence, we have

$$\frac{\Xi_{A_x+p} - A_y}{p} < \frac{B_y - A_y}{B_x - A_x} \tag{4.3}$$

$$\frac{\Xi_{A_x+r} - A_y}{r} < \frac{B_y - A_y}{B_x - A_x}. \tag{4.4}$$

Moreover, since p is a period of ξ, we have

$$\begin{aligned}
B_y - \Xi_{A_x+p} &= \Xi_{B_x} - \Xi_{A_x+p} \\
&= \Xi_{A_x+p+r} - \Xi_{A_x+p} \\
&= \Xi_{A_x+r} - \Xi_{A_x}.
\end{aligned}$$

Then by (4.3) and (4.4), we have a contradiction that

$$\begin{aligned} B_y - A_y &= (B_y - \Xi_{A_x+p}) + (\Xi_{A_x+p} - A_y) \\ &= (\Xi_{A_x+r} - \Xi_{A_x}) + (\Xi_{A_x+p} - A_y) \\ &< \frac{B_y - A_y}{B_x - A_x}\, r + \frac{B_y - A_y}{B_x - A_x}\, p \\ &= \frac{(B_y - A_y)(r+p)}{B_x - A_x} = B_y - A_y. \end{aligned}$$

Thus, $B_x - A_x \le \mathrm{per}(\xi)$ holds for any prime segment of $\underline{\Gamma}_\xi$, and $B_x - A_x = \mathrm{per}(\xi)$ holds for any central prime segment of $\underline{\Gamma}_\xi$ since in this case, $B_x - A_x$ is a period of ξ as is proved in the above.

The proof is similar for a prime segment in $\overline{\Gamma}_\xi$. □

Let AB be a central prime segment of $\underline{\Gamma}_\xi$. Then, the factor $\xi[A_x, B_x]$ of ξ is called a *convex kernel* of ξ. Also, for a central prime segment AB of $\overline{\Gamma}_\xi$, the factor $\xi[A_x, B_x]$ of ξ is called a *concave kernel* of ξ.

Lemma 9. (1) *For a convex kernel η of ξ, we have $\eta = \underline{\lambda}(u, v)$ with $u := |\eta| = \mathrm{per}(\xi)$ and $v := |\eta|_1 = \mathrm{per}(\xi)\hat{\rho}(\xi)$. Hence, a convex kernel of ξ is unique if it exists.*
(2) *For a concave kernel ζ of ξ, we have $\eta = \overline{\lambda}(u, v)$ with $u := |\zeta| = \mathrm{per}(\xi)$ and $v := |\zeta|_1 = \mathrm{per}(\xi)\hat{\rho}(\xi)$. Hence, a concave kernel of ξ is unique if it exists.*
(3) *Either the convex kernel of ξ or the concave kernel of ξ exists.*
(4) *For the maximal likelihood estimator, we have $\hat{\alpha}(\xi) = \hat{\rho}(\xi)$.*
(5) *For the convex kernel η of ξ, we have*

$$\min\{|\theta|_1 - |\theta|\rho(\eta)\} = -1 + \frac{1}{|\eta|}, \tag{4.5}$$

and for the concave kernel ζ of ξ, we have

$$\max\{|\theta|_1 - |\theta|\rho(\zeta)\} = 1 - \frac{1}{|\zeta|},$$

where "min" and "max" are for all prefixes θ of η or ζ, respectively.

Proof. (1) follows from Lemma 8 and the fact that

$$|\eta|_1 = \mathrm{per}(\xi)\rho(\eta) = \mathrm{per}(\xi)\hat{\rho}(\xi).$$

(2) Similar to (1).
(3) follows from Lemma 4.
(4) follows from (4) of Lemma 7.
(5) Let $u := |\eta|$ and $v := |\eta|_1$. Then, u and v coprime by Lemma 8.
Since $\eta = \underline{\lambda}(u, v)$, we have

$$|\eta_1\eta_2\cdots\eta_i|_1 - |\eta_1\eta_2\cdots\eta_i|\rho(\eta) = \lfloor iv/u \rfloor - iv/u = -j/u$$

for any $i = 0, 1, \cdots, u$, where $j \equiv iv \ (mod\ u)$ with $0 \le j \le u-1$. Since u and v are coprime, there exists $i = 0, 1, \cdots, u$ such that $j = u-1$ holds in the above. Thus, we have (4.5).

The other part is proved in the same way. □

Proof of Theorem 2

(1) We have already proved that $\hat{\alpha}(\xi) = \hat{\rho}(\xi)$ in Lemma 9.

To prove that $P_{\hat{\alpha}}(\xi) = 1/\mathrm{per}(\xi)$, without loss of generality, we may assume that there exists the convex kernel η of ξ such that $\eta = \xi[i, i+u]$ with $u = \mathrm{per}(\xi)$. Then, AB is a central prime segment of $\underline{\Gamma}_\xi$, where $A = (i, \Xi_i)$, $B = (i+u, \Xi_{i+u})$. Let AB be on the graph $y = x\alpha + \beta_1$ $(x \in [0, m])$. Then, by Lemma 7, $\alpha = \hat{\alpha} = \hat{\alpha}(\xi)$. Let β_2 be the maximum value of β such that the graph $y = x\alpha + \beta$ $(x \in [0, m])$ is in the domain $\overline{\Sigma}_\xi$. Then, $y = x\hat{\alpha} + \beta$ $(x \in [0, m])$ is in the domain $\underline{\Sigma}_\xi$ if and only if $\beta_1 \le \beta < \beta_2$. Hence by (1) of Lemma 2, $P_{\hat{\alpha}}(\xi) = \beta_2 - \beta_1$.

Since the graph $y = x\hat{\alpha} + \beta_2$ $(x \in [0, m])$ is in the domain $\overline{\Sigma}_\xi$, any point $(h, \Xi_h + 1)$ for $h = 0, 1, \cdots, m$ is on or above the graph. Moreover, by the maximality of β_2, there exists an integer h with $h \in [0, m]$ such that $(h, \Xi_h + 1)$ is on the graph $y = x\hat{\alpha} + \beta_2$. Since $u = \mathrm{per}(\xi)$ and $\hat{\alpha} = \rho(\eta)$, such an h can be found in the interval $[i, i+u]$. Therefore, we have

$$\beta_2 = \Xi_h + 1 - h\hat{\alpha} = \min_{i \le j \le i+u} \{\Xi_j + 1 - j\hat{\alpha}\}.$$

Since $\beta_1 = \Xi_i - i\hat{\alpha}$, we have

$$\begin{aligned} \beta_2 - \beta_1 &= \min_{i \le j \le i+u} \{(\Xi_j + 1 - j\hat{\alpha}) - (\Xi_i - i\hat{\alpha})\} \\ &= 1 + \min_{i \le j \le i+u} \{|\xi_{i+1}\xi_{i+2}\cdots\xi_j|_1 - |\xi_{i+1}\xi_{i+2}\cdots\xi_j|\rho(\eta)\}. \end{aligned}$$

Hence, by (5.6),

$$P_{\hat{\alpha}}(\xi) = \beta_2 - \beta_1 = \frac{1}{|\eta|} = \frac{1}{\mathrm{per}(\xi)},$$

which completes the proof of (1).

(2) Let $\xi \in \mathrm{St}_m$ be nontrivial. Let $\alpha = \hat{\alpha} - \varepsilon$ for a sufficiently small $\varepsilon > 0$. Let the graph $y = x\alpha + \beta_1$ $(0 \le x \le m)$ be in $\underline{\Sigma}_\xi$ but not in $\overline{\Sigma}_\xi$ and let the graph $y = x\alpha + \beta_2$ $(0 \le x \le m)$ be in $\overline{\Sigma}_\xi$ but not in $\underline{\Sigma}_\xi$. Then, we have $P_\alpha(\xi) = \beta_2 - \beta_1$.

Let $i := \max \overline{I}(\xi)$ and $j := \min \underline{I}(\xi)$. Then, the point (i, Ξ_i) is on $y = x\alpha + \beta_1$ and the point $(j, \Xi_j + 1)$ is on $y = x\alpha + \beta_2$. Therefore,

$$\begin{aligned} P_\alpha(\xi) &= \beta_2 - \beta_1 \\ &= (\Xi_j + 1 - j\alpha) - (\Xi_i - i\alpha) \\ &= (\Xi_j + 1 - j(\hat{\alpha} - \varepsilon)) - (\Xi_i - i(\hat{\alpha} - \varepsilon)) \\ &= (\Xi_j + 1 - j\hat{\alpha}) - (\Xi_i - i\hat{\alpha}) - (i - j)\varepsilon \\ &= P_{\hat{\alpha}}(\xi) - (\max \overline{I}(\xi) - \min \underline{I}(\xi))\varepsilon, \end{aligned}$$

which implies that the slope of the graph between α_1 and $\hat{\alpha}$ in Figure 3 is $\max \overline{I}(\xi) - \min \underline{I}(\xi)$. Since $P_{\hat{\alpha}}(\xi) = 1/\mathrm{per}(\xi)$, this proved the formula for α_1. Similarly, we can prove the formula for α_2.

(3) The statistics $(\hat{\alpha},\ \max \overline{I} - \min \underline{I},\ \max \underline{I} - \min \overline{I})$ is the minimum sufficient statistics since it induces the same partition as $(\hat{\alpha}, \alpha_1, \alpha_2)$ by the above formulas. Note that $\mathrm{per}(\xi)$ is a function of $\hat{\alpha}$ since it is the denominator of the irreducible rational fraction equal to $\hat{\alpha}$.

(4) Let $T := (\hat{\alpha},\ \max \overline{I} - \min \underline{I},\ \max \underline{I} - \min \overline{I})$.

Let m be an even number with $m \geq 6$. Let

$$\begin{aligned} \xi &= \overline{\lambda}(m-1,\ 2)\, 1 \\ \eta &= \underline{\lambda}(m-1,\ 2)\, 0. \end{aligned}$$

Then, we have $|\xi| = |\eta| = m$, $\mathrm{per}(\xi) = \mathrm{per}(\eta) = m-1$ and $\hat{\rho}(\xi) = \hat{\rho}(\eta) = 2/(m-1)$ while $\rho(\xi) = 3/m$ and $\rho(\eta) = 2/m$. Moreover, it is easily seen that

$$\begin{aligned} \overline{I}(\xi) &= \{m/2\} \\ \overline{I}(\eta) &= \{0,\ m-1\} \\ \underline{I}(\xi) &= \{0,\ m-1\} \\ \underline{I}(\eta) &= \{(m-2)/2\}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \max \overline{I}(\xi) - \min \underline{I}(\xi) &= (m/2) - 0 = m/2 \\ \max \overline{I}(\eta) - \min \underline{I}(\eta) &= (m-1) - (m-2)/2 = m/2 \\ \max \underline{I}(\xi) - \min \overline{I}(\xi) &= (m-1) - (m/2) = (m-2)/2 \\ \max \underline{I}(\eta) - \min \overline{I}(\eta) &= (m-2)/2 - 0 = (m-2)/2. \end{aligned}$$

Hence, we have $T(\xi) = T(\eta)$ while $\rho(\xi) \neq \rho(\eta)$, which implies that ρ is not based on T.

Let m be an odd number with $m \geq 9$. Let

$$\begin{aligned} \xi &= \overline{\lambda}(m-2,\ 2)\, 10 \\ \eta &= \underline{\lambda}(m-2,\ 2)\, 00. \end{aligned}$$

Then, we have $|\xi| = |\eta| = m$, $\mathrm{per}(\xi) = \mathrm{per}(\eta) = m-2$ and $\hat{\rho}(\xi) = \hat{\rho}(\eta) = 2/(m-2)$ while $\rho(\xi) = 3/m$ and $\rho(\eta) = 2/m$. Moreover, it is easily seen that

$$\begin{aligned} \overline{I}(\xi) &= \{(m-1)/2\} \\ \overline{I}(\eta) &= \{0,\ m-2\} \\ \underline{I}(\xi) &= \{0,\ m-2\} \\ \underline{I}(\eta) &= \{(m-3)/2\}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \max \overline{I}(\xi) - \min \underline{I}(\xi) &= (m-1)/2 - 0 = (m-1)/2 \\ \max \overline{I}(\eta) - \min \underline{I}(\eta) &= (m-2) - (m-3)/2 = (m-1)/2 \\ \max \underline{I}(\xi) - \min \overline{I}(\xi) &= (m-2) - (m-1)/2 = (m-3)/2 \\ \max \underline{I}(\eta) - \min \overline{I}(\eta) &= (m-3)/2 - 0 = (m-3)/2. \end{aligned}$$

Hence, we have $T(\xi) = T(\eta)$ while $\rho(\xi) \neq \rho(\eta)$, which implies that ρ is not based on T.

(5) Since the Bayes estimate α_3 for the observation ξ with respect to the uniform prior distribution on $\alpha \in [0,1]$ is the mean of α measured by the normailized likelihood function for ξ and the graph of the likelihood function is as in Figure 3, we have

$$\alpha_3 = \frac{\hat{\alpha} + \alpha_1 + \alpha_2}{3}$$

(6) Suppose that there exists a UMVUE T for α. Consider 2 unbiased estimators

$$\begin{aligned} \rho(\xi) &= \frac{\xi_1 + \xi_2 + \cdots + \xi_m}{m} \\ \rho'(\xi) &:= \frac{\xi_1 + \xi_2 + \cdots + \xi_{m-1}}{m-1} \end{aligned}$$

of α.

At first, assume that m is odd. Then for $\alpha = (m+1)/(2m)$, we have $V_\alpha(\rho) = 0$, since P_α is supported by the following m sample points:

$$\begin{gathered} 1(01)(01)\cdots(01)\ ,\ (01)1(01)\cdots(01)\ ,\ \cdots\ ,\ (01)(01)\cdots 1(01)\ ,\ (01)(01)\cdots(01)1 \\ 1(10)(10)\cdots(10)\ ,\ (10)1(10)\cdots(10)\ ,\ \cdots\ ,\ (10)(10)\cdots 1(10) \end{gathered}$$

each point of which has the same weight $1/m$. On the other hand, if $\alpha = 1/2$, then we have $V_\alpha(\rho') = 0$, since P_α is supported by the following 2 sampe points:

$$0101\cdots010\ ,\ 1010\cdots101$$

each point of which has the same weight $1/2$. Note that the sample point $1010\cdots101$ belongs to the both sets. Since T is UMVUE, we have

$$V_{(m-1)/(2m)}(T) \leq V_{(m-1)/(2m)}(\rho) = 0 \tag{4.6}$$

$$V_{1/2}(T) \leq V_{1/2}(\rho') = 0. \tag{4.7}$$

It follows from (4.6) that $T(1010\cdots101) = (m-1)/(2m)$, while by (4.7), we have $T(1010\cdots101) = 1/2$, which is a contradiction. Thus, UMVUE does not exist.

For the case that m is even, we can do the same argument for $\alpha = 1/2$ and $m/(2m-2)$ to lead a contradiction. □

Example 1. St_6 consists of the following 36 elements:

000000	100000	010000	001000	000100	100100
010100	000010	100010	010010	001010	101010
011010	010110	110110	101110	011110	111110
000001	100001	010001	001001	101001	100101
010101	110101	101101	011101	111101	101011
011011	111011	110111	101111	011111	111111

where 010100, 001010, 101001 and 100101 have the same $(\hat{\alpha}, \alpha_1, \alpha_2) = (2/5, 1/3, 1/2)$ and 101011, 110101, 010110 and 011010 have the same $(\hat{\alpha}, \alpha_1, \alpha_2) = (3/5, 1/2, 2/3)$. In the other cases, the sample mean coincides if the minimum sufficient statistics coincides. Since $\rho(010100) = \rho(001010) = 1/3$, $\rho(101001) = \rho(100101) = 1/2$, $\rho(101011) = \rho(110101) = 2/3$ and $\rho(010110) = \rho(011010) = 1/2$, we have

$$U(\xi) := E(\rho|\hat{\alpha}, \alpha_1, \alpha_2)(\xi) = \begin{cases} 5/12 & \xi \in \{010100, 001010, 101001, 100101\} \\ 7/12 & \xi \in \{101011, 110101, 010110, 011010\} \\ \rho(\xi) & \text{otherwise.} \end{cases}$$

Clearly, U is an unbiased estimator of α. For any α with $1/3 < \alpha \leq 2/5$, since we have

$$P_\alpha(010100) = P_\alpha(001010) = P_\alpha(101001) = P_\alpha(100101) = 3\alpha - 1,$$

it holds that

$$\begin{aligned} V_\alpha(U) &= V_\alpha(\rho) - 4(3\alpha - 1)(\frac{1}{2}(\frac{1}{3} - \frac{5}{12})^2 + \frac{1}{2}(\frac{1}{2} - \frac{5}{12})^2) \\ &= V_\alpha(\rho) - 4(1/2)\{6\alpha\}/144 \\ &= \{6\alpha\}(1 - \{6\alpha\})/36 - \{6\alpha\}/72 \\ &= \{6\alpha\}(1/2 - \{6\alpha\})/36 \end{aligned}$$

In the same way, we have

$$V_\alpha(U) = \begin{cases} \{6\alpha\}(1/2 - \{6\alpha\})/36 & 1/3 < \alpha \leq 2/5 \\ (\{6\alpha\} - 1/3)(1 - \{6\alpha\})/36 & 2/5 \leq \alpha < 1/2 \\ \{6\alpha\}(2/3 - \{6\alpha\})/36 & 1/2 < \alpha \leq 3/5 \\ (\{6\alpha\} - 1/2)(1 - \{6\alpha\})/36 & 3/5 \leq \alpha \leq 2/3 \\ \{6\alpha\}(1 - \{6\alpha\})/36 & \text{otherwise.} \end{cases}$$

Example 2. St_7 consists of the following 36 elements:

000000	100000	010000	001000	000100	100100
010100	000010	100010	010010	001010	101010
011010	010110	110110	101110	011110	111110
000001	100001	010001	001001	101001	100101
010101	110101	101101	011101	111101	101011
011011	111011	110111	101111	011111	111111

where 010100, 001010, 101001 and 100101 have the same $(\hat{\alpha}, \alpha_1, \alpha_2) = (2/5, 1/3, 1/2)$ and 101011, 110101, 010110 and 011010 have the same $(\hat{\alpha}, \alpha_1, \alpha_2) = (3/5, 1/2, 2/3)$. In the other cases, the sample mean coincides if the minimum sufficient statistics coincides.

Acknowledement

The authors thank Prof. Shin-ichi Yasutomi (Suzuka), Prof. Valérie Berthé (Montpellier), Prof. Shunji Ito (Tsuda College), Prof. Kunio Tanabe, Prof. Takashi Tsuchiya and Prof. Satoshi Kuriki of The Graduate University for Advanced Studies (Tokyo) and Prof. Luca Zamboni (North Texas University) for their suggestions and useful discussions on the subject.

References

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

[2] J. Nagumo and S. Sato, On a response characteristic of a mathematical neuron model, Kybernetic, **10** (1972), 155–164.

[3] M. Hata, Dynamics of Caianiello's equation, J. Math. Kyoto Univ., **22** (1982), 155–173.

[4] Jean Berstel and Michel Pocchiola, Random generation of finite Sturmian words, Discrete Mathematics 153 (1996), 29-35.

[5] Shin-ichi Yasutomi, The continued fraction expansion of α with $\mu(\alpha) = 3$, Acta Arith. 84-4 (1998), 337-374.

[6] Hayato Takahashi, Parametric models and Kolmogorov complexity, Doctoral Thesis at The Graduate University for Advanced Studies (Tokyo), 2002.

[7] Hayato Takahashi and Kazuyuki Aihara, Algorithmic analysis of irrational rotations in a single neuron model, submitting.