Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XIX (2019), 757-780

Central limit theorem for probability measures defined by sum-of-digits function in base 2

JORDAN EMME AND PASCAL HUBERT

Abstract. In this paper we prove a central limit theorem for some probability measures defined as asymptotic densities of integer sets defined via sum-of-digit-function. To any non-negative integer a we can associate a measure on $\mathbb{Z}$ called μ_a such that, for any d, $\mu_a(d)$ is the asymptotic density of the set of non-negative integers n such that $s_2(n+a)-s_2(n) = d$ where $s_2(n)$ is the number of digits "1" in the binary expansion of n. We express this probability measure as a product of matrices whose coefficients are operators of $l^1(\mathbb{Z})$. Then we take a sequence of integers $(a_X(n))_{n\in\mathbb{N}}$ defined via a balanced Bernoulli sequence X. We prove that, for almost every sequence, and after renormalization by the typical variance, we have a central limit theorem by computing all the moments and proving that they converge towards the moments of the normal law $\mathcal{N}(0, 1)$.

Mathematics Subject Classification (2010): 37A45 (primary); 11P99, 60F05 (secondary).

1. Introduction

1.1. Background

In this paper we study some properties of sets defined via sum-of-digit function in base 2. Namely, for a given integer a, we study the asymptotic density of set of integers such that the difference of 1 in their binary expansion before and after addition with a is a given integer. More precisely, we define

$$\forall n \in \mathbb{N},\ s_2(n) = \sum_{k=0}^{m} n_k,$$

where

$$n = \sum_{k=0}^{m} n_k 2^k,$$

Received September 18, 2017; accepted in revised form September 18, 2017.
Published online June 2019.

and

$$\forall a \in \mathbb{N},\ \forall d \in \mathbb{Z},\ \mu_a(d) = \lim_{N\to+\infty} \frac{1}{N}\#\{n < N \mid s_2(n+a) - s_2(n) = d\}.$$

This can be linked with the correlation functions that are studied, for instance, in [1] or with the properties of the sum-of-digit functions which have been extensively studied, for instance [5] in, or, more recently, in [8]. We can also quote [10] for the links between Thue-Morse sequence and the sum-of-digits function in base 2. The type of questions answered in this article also share some similarity with [13] and [12].

More precisely, we are interested in normality properties of such sets, and in this paper we give a central limit theorem for a random a. This kind of properties have raised a considerable interest in number theory. We can quote [4, 6, 11] for some of these normality properties for q-additive functions.

In [9], we were interested in those densities of sets and more precisely in their asymptotic properties as a goes to infinity. The methods for computing those densities were essentially combinatorial. In this paper, we are closer to dynamical systems, as we study a random product of matrices.

1.2. Results

Definition 1.1. Let $n \in \mathbb{N}$. There exists a unique smallest $m \in \mathbb{N}$ and a unique sequence $\{n_0, \dots, n_m\} \in \{0, 1\}^m$ such that

$$n = \sum_{k=0}^{m} n_k \cdot 2^k.$$

Set $\underline{n}_2 = n_m \dots n_0$. The word $(\underline{n}_2)$ is an element of the free monoid $\{0, 1\}^*$.

Definition 1.2. Define the sum-of-digits function in base 2 s_2, as:

$$\begin{aligned} s_2 : \mathbb{N} &\to \mathbb{N} \\ n &\mapsto \sum_{k=0}^{m} n_k, \end{aligned}$$

where $\underline{n}_2 = n_m \dots n_0$.

We are interested in the following equation with parameters $a \in \mathbb{N}$, $d \in \mathbb{Z}$ and unknown $n \in \mathbb{N}$:

$$s_2(n+a) - s_2(n) = d.$$

More precisely, we wish to understand the asymptotic densities of the following sets:

$$\mathcal{E}_{a,d} := \{n \in \mathbb{N} \mid s_2(n+a) - s_2(n) = d\}.$$

In [9], we prove the following:

Proposition 1.3. *For any $a \in \mathbb{N}$, $d \in \mathbb{Z}$, there exists a finite set of words $\mathcal{P}_{a,d} := \{w_1, \ldots, w_k\} \subset \{0,1\}^*$ such that:*

$$\mathcal{E}_{a,d} = \bigcup_{i \in \{1,\ldots,k\}} [w_i],$$

where $[w]$ is the set of integers n such that $\underline{n}_2$ ends with w.

Remark 1.4. Remark that $\mathcal{P}_{a,d}$ is finite and, possibly, empty (whenever $d > s_2(a)$ actually).

From Proposition 1.3, it is clear that the densities of the sets $\mathcal{E}_{a,d}$ exist. This was known since [1]. In this way we define the main object of our study:

Definition 1.5. Let us define, for any $a \in \mathbb{N}$, the probability measure μ_a by:

$$\forall d \in \mathbb{Z},\ \mu_a(d) := \lim_{N \to +\infty} \frac{\#\{n \le N \mid s_2(n+a) - s_2(n) = d\}}{N}.$$

Remark 1.6. Remark that in order to get a probability measure, we cannot exchange the roles of d and a. Indeed, one can check that for any d, the sequence $(\mu_a(d))_{a \in \mathbb{N}}$ is not even in $l^1(\mathbb{N})$.

We are interested in asymptotic properties of the measures μ_a as a goes to infinity in a certain sense. Namely, we define the following quantity:

Definition 1.7. Let $a \in \mathbb{N}$. Define the quantity:

$$l(a) := \#\left\{\text{occurences of "01" in } \underline{a}_2\right\}.$$

We recall two theorems from [9]:

Theorem 1.8. *There exists a constant $C > 0$ such that, for any $a \in \mathbb{N}$:*

$$\|\mu_a\|_2 \le C \cdot l(a)^{-1/4}.$$

Theorem 1.9. *For any $a \in \mathbb{N}$, the probability measure μ_a has mean 0 and its variance is bounded by:*

$$l(a) - 1 \le \mathrm{Var}(\mu_a) \le 4l(a) + 2.$$

This last theorem raises the question of whether the ratio $\frac{\mathrm{Var}(\mu_a)}{l(a)}$ converges as $l(a)$ goes to infinity. We do this in the generic case for the balanced Bernoulli measure. More precisely, we have the following central limit theorem:

Theorem 1.10. *Let $X = (X_n)_{n\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ be a generic sequence for the balanced Bernoulli measure. Define the sequence $(a_X(n))_{n\in\mathbb{N}}$ in the following way:*

$$a_X(n) = \sum_{k=0}^{n} X_k \cdot 2^k.$$

For any $n \in \mathbb{N}$, let $\widetilde{\mu}_{a_X(n)} \in l^1\left(\sqrt{\frac{2}{n}}\mathbb{Z}\right)$ defined by

$$\forall d \in \sqrt{\frac{2}{n}}\mathbb{Z},\ \widetilde{\mu}_{a_X(n)} = \mu_{a_X(n)}\left(\sqrt{\frac{n}{2}}d\right).$$

Then

$$\widetilde{\mu}_{a_X(n)} \xrightarrow[n\to+\infty]{\text{weak}} \varphi$$

where

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}t^2}.$$

Remark 1.11. An equivalent formulation, with the same notation is:

$$\forall x \in \mathbb{R},\ \lim_{n\to+\infty} \lim_{N\to+\infty} \frac{1}{N} \# \left\{ m \leq N \mid \frac{s_2(m + a_X(n)) - s_2(m)}{\sqrt{\frac{n}{2}}} \leq x \right\} = \Phi(x)$$

where Φ is the repartition function of the normal law $\mathcal{N}(0,1)$.

Finally we wish to state Cusick's conjecture.

Define the following quantity:

$$\forall a \in \mathbb{N},\ c_a := \lim_{N\to+\infty} \frac{1}{N} \#\{n \leq N \mid s_2(n+a) \geq s_2(n)\}.$$

The conjecture consists of two parts:

$$\forall a \in \mathbb{N}, c_a \geq \frac{1}{2}$$

and

$$\liminf_{a\to+\infty} c_a = \frac{1}{2}.$$

This question arose as Cusick was working on a similar combinatorial conjecture in [3]. For recent advances on this, one can look at [15] and [7]. In particular, the main theorem of this paper has for immediate corollary that $\frac{1}{2}$ is an accumulation point for the sequence $(c_a)_{a\in\mathbb{N}}$.

Moreover, the proof of our theorem answers a question left open in [7] since it states that the difference $s_2(n+a) - s(n)$ is "usually normally distributed with mean zero and variance $\frac{|a|_2}{2}$" where $|a|_2$ denotes the length of a in base 2.

1.3. Outline of the paper

The goal of this paper is to demonstrate Theorem 1.10 by a moments method. Namely, given a sequence of probability measures, we prove the weak convergence towards the normal law $\mathcal{N}(0, 1)$ by proving that all the moments of this sequence converge towards the moments of the normal law.

This article is organised as follows:

Section 2 deals with the measures μ_a, that we already studied in [9]. We recall some of their properties (and most importantly a recurrence relation between them) and write them as a finite product of matrices whose coefficients are operators on $l^1(\mathbb{Z})$. This is a convenient form for our study since it allows to compute the Fourier transforms of these measures explicitly.

Section 3 is devoted to the proof of Theorem 1.10. It is divided in the following way.

In Subsection 3.1, we explicit this Fourier transform and give its Taylor expansion around 0 at order 2. This is what we need in order to compute the variance of μ_a. We also mention that the characteristic function can be written as a product of matrices, which is the form that will be studied throughout the article.

Subsection 3.2 is devoted to the computation of the variance of μ_a. First we do this in the general case and give an explicit formula depending only on the binary decomposition of a. We remark that this expression depends on some correlations of sequences in $\{0, 1\}^{\mathbb{N}}$.

Then, in Subsection 3.3, we want to compute the "generic" behaviour of μ_a (meaning for a a whose binary expansion is given by a balanced Bernoulli sequence). For this, we use a result in [2] to estimate the correlation terms. It appears that in the generic case, the variance is approximately $\frac{|a|_2}{2}$ (where $|a|_2$ is the length of $\underline{a}_2$. So we know that in order to get a central limit theorem, we have to renormalize μ_a by the squareroot of its variance namely $\sqrt{\frac{|a|_2}{2}}$.

In Subsection 3.5, since we have to compute all the moments, we need to know all the coefficients in the Taylor expansion of the characteristic function but it seems difficult to give their expression in the general case. Hence we wish to understand how "big" the different terms are in order to know which one will be killed by the renormalization and which one will contribute. To that end, we classify the terms in the Taylor expansion and we bound them using some algebraic properties of the matrices involved in this computation. This in turn gives bounds on the moments.

Finally, in Subsection 3.6, we show that the moments converge towards the moments of the normal law. Thanks to the study of the contributions from the previous section, we are limited to actually computing the terms that have a chance to contribute. Some elementary linear algebra and the study of the correlations appearing in Section 3.2 are the essential tools for this.

We would like to underline the fact that the way we compute the moments limits gives a bound on the speed of convergence of the moments. However, this speed is dependant on the order of the moment and thus gives no clue as to the speed of convergence of the measures towards the normal law. This could be further studied along with a local limit theorem.

ACKNOWLEDGEMENTS. We wish to thank Christian Mauduit and Joël Rivat for their interest in this problem and for sharing their knowledge in the historical and scientific background of this question. Of course we have to thank Alexander Bufetov for his precious help regarding the moments method, especially for giving the reference needed. We also would like to thank Thomas Stoll for mentionning Cusick's conjecture. We thank Julien Cassaigne for his useful remarks on the variance properties. Finally, we thank Lukas Spiegelhofer for making us aware of the works in [7].

2. Measures μ_a on $\mathbb{Z}$

Let us start by remarking the following:

Remark 2.1. Let $a \in \mathbb{N}$ and $d \in \mathbb{Z}$.

$$\mathcal{P}_{2a,d} = \left\{ w0 \mid w \in \mathcal{P}_{a,d} \right\} \cup \left\{ w1 \mid w \in \mathcal{P}_{a,d} \right\}$$

and

$$\mathcal{P}_{2a+1,d} = \left\{ w0 \mid w \in \mathcal{P}_{a,d-1} \right\} \cup \left\{ w1 \mid w \in \mathcal{P}_{a+1,d+1} \right\}.$$

For more details about this remark we refer the reader to [9].

From this we deduce the following:

Proposition 2.2. *For any $a \in \mathbb{N}$:*

$$\mu_{2a} = \mu_a$$

and

$$\mu_{2a+1}(d) = \frac{1}{2}\mu_a(d-1) + \frac{1}{2}\mu_{a+1}(d+1).$$

Remark 2.3. Notice that a probability measure on $\mathbb{Z}$ is, in particular, an element of $l^1(\mathbb{Z})$. In all that follows, for simplicity of writing, we will always identify a measure on $\mathbb{Z}$ with its associated sequence in $l^1(\mathbb{Z})$. In particular, we see μ_a both as a measure and as an element of $l^1(\mathbb{Z})$. Let us define the shift S on $l^1(\mathbb{Z})$.

$$\begin{array}{rccc} S: & l^1(\mathbb{Z}) & \to & l^1(\mathbb{Z}) \\ & (x_n)_{n\in\mathbb{Z}} & \mapsto & (x_{n+1})_{n\in\mathbb{Z}}. \end{array}$$

Then, the identities of Proposition 2.2 can be written:

$$\mu_{2a} = \mu_a$$

and

$$\mu_{2a+1} = \frac{1}{2}S^{-1}(\mu_a) + \frac{1}{2}S(\mu_{a+1}).$$

Example 2.4. It is easy to see that $\mu_0 = \delta_0$. Then, either by standard computation or by using Proposition 2.2, one obtains:

$$\mu_1 = \frac{1}{4}\sum_{\substack{n\in\mathbb{Z}\\ n\leq 1}} \delta_n \cdot 2^n.$$

Proposition 2.5. *For any $a \in \mathbb{N}$,*

$$\mu_a = \begin{pmatrix} Id & 0 \end{pmatrix} \; A_{a_0} \cdots A_{a_{n-1}} A_{a_n} \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix},$$

where $\underline{a}_2 = a_n \dots a_0$, and

$$A_0 = \begin{pmatrix} Id & 0 \\ \frac{1}{2}S^{-1} & \frac{1}{2}S \end{pmatrix}, \qquad A_1 = \begin{pmatrix} \frac{1}{2}S^{-1} & \frac{1}{2}S \\ 0 & Id \end{pmatrix}.$$

Proof. It suffices to remark that, for any $a \in \mathbb{N}$,

$$A_0 \begin{pmatrix} \mu_a \\ \mu_{a+1} \end{pmatrix} = \begin{pmatrix} \mu_{2a} \\ \mu_{2a+1} \end{pmatrix}$$

and

$$A_1 \begin{pmatrix} \mu_a \\ \mu_{a+1} \end{pmatrix} = \begin{pmatrix} \mu_{2a+1} \\ \mu_{2a+2} \end{pmatrix}$$

with Proposition 2.2. □

Notice that this proposition is a clearer version of [9, Theorem 1.2.1].

3. Central limit theorem

3.1. Characteristic function

Let $a \in \mathbb{N}$ with $\underline{a}_2 = a_n \dots a_0$. The characteristic function of μ_a, denoted by $\widehat{\mu_a}$, is defined in the standard way:

$$\forall\theta \in [0, 2\pi),\; \widehat{\mu_a}(\theta) = \sum_{d\in\mathbb{Z}} e^{id\theta}\mu_a(d).$$

By Proposition 2.2, the characteristic function $\widehat{\mu_a} : [0, 2\pi) \to \mathbb{C}$ is given by:

$$\widehat{\mu_a}(\theta) = \begin{pmatrix} 1 & 0 \end{pmatrix} \; \hat{A}_{a_0} \cdots \hat{A}_{a_{n-1}} \hat{A}_{a_n} \begin{pmatrix} \widehat{\mu_0}(\theta) \\ \widehat{\mu_1}(\theta) \end{pmatrix},$$

where

$$\hat{A}_0(\theta) := \begin{pmatrix} 1 & 0 \\ \frac{1}{2}e^{i\theta} & \frac{1}{2}e^{-i\theta} \end{pmatrix}, \quad \hat{A}_1(\theta) := \begin{pmatrix} \frac{1}{2}e^{i\theta} & \frac{1}{2}e^{-i\theta} \\ 0 & 1 \end{pmatrix}.$$

Indeed, from the recurrence relations in Proposition 2.2, one has

$$\widehat{\mu_{2a}}(\theta) = \widehat{\mu_a}(\theta)$$

and

$$\widehat{\mu_{2a+1}}(\theta) = \frac{1}{2}e^{i\theta}\widehat{\mu_a}(\theta) + \frac{1}{2}e^{-i\theta}\widehat{\mu_{a+1}}(\theta).$$

These recurrence relations on the characteristic function justify the fact, that we write it as a product of matrices.

A quick computation yields

$$\widehat{\mu_0}(\theta) = 1, \quad \widehat{\mu_1}(\theta) = \frac{e^{i\theta}}{2 - e^{-i\theta}},$$

and so,

$$\widehat{\mu_a}(\theta) = \begin{pmatrix} 1 & 0 \end{pmatrix} \hat{A}_{a_0} \cdots \hat{A}_{a_{n-1}} \hat{A}_{a_n} \begin{pmatrix} 1 \\ \dfrac{e^{i\theta}}{2 - e^{-i\theta}} \end{pmatrix}.$$

Now let us define the matrices playing a role in the Taylor expansion of $\widehat{\mu_a}$ near 0:

$$I_0 = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \alpha_0 = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}, \quad \beta_0 = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$
$$I_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}, \quad \alpha_1 = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad \beta_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Indeed, we have:

$$\hat{A}_j(\theta) = I_j + i\theta\alpha_j - \frac{1}{2}\theta^2\beta_j + O\big(\theta^3\big),$$

with $j \in \{0, 1\}$.

Notice that the Taylor expansion near 0 of $\theta \mapsto \frac{e^{-i\theta}}{2-e^{i\theta}}$ is:

$$\frac{e^{i\theta}}{2 - e^{-i\theta}} = 1 - \theta^2 + O\big(\theta^3\big).$$

3.2. Computation of the variance

Define the variance of μ_a:

$$\mathrm{Var}(\mu_a) = \sum_{d \in \mathbb{Z}} \mu_a(d) d^2.$$

Theorem 3.1. *For any $a \in \mathbb{N}$ with $\underline{a}_2 = a_n \dots a_0$, set, for any $j \in \{0, \dots, n\}$, $b_j = (-1)^{a_j+1}$. The variance of μ_a is given by the following:*

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=0}^{n-i} \frac{b_{k+i} b_k}{2^i} + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}}.$$

Proof. Notice that the variance is given by:

$$\mathrm{Var}(\mu_a) = \begin{pmatrix} 1 & 0 \end{pmatrix} \left(\beta_{a_0} + I_{a_0} \beta_{a_1} + \dots + I_{a_0} \cdots I_{a_{n-1}} \beta_{a_n} \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \end{pmatrix} I_{a_0} \cdots I_{a_n} \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

since $\alpha_j \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$ and $I_j \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and since the variance is given by the quadratic coefficient in the Taylor expansion of the characteristic function.

We now apply a change of basis to simultanously trigonalize the matrices I_0 and I_1.

Let us note that

$$P := \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

and compute

$$\forall j \in \{0, 1\}, \widetilde{I}_j := P I_j P^{-1} = \begin{pmatrix} 1 & \frac{(-1)^{j+1}}{2} \\ 0 & \frac{1}{2} \end{pmatrix}, \quad \widetilde{\beta}_j := P \beta_j P^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{(-1)^{j+1}}{2} & 0 \end{pmatrix},$$

and

$$P \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad P \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

With this change of basis, the variance becomes

$$\mathrm{Var}(\mu_a) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \left(\widetilde{\beta}_{a_0} + \widetilde{I}_{a_0} \widetilde{\beta}_{a_1} + \dots + \widetilde{I}_{a_0} \cdots \widetilde{I}_{a_{n-1}} \widetilde{\beta}_{a_n} \right) \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \widetilde{I}_{a_0} \cdots \widetilde{I}_{a_n} \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Notice now that for any $k \in \{0, \dots, n\}$,

$$\widetilde{I}_{a_0} \cdots \widetilde{I}_{a_k} = \begin{pmatrix} 1 & \sum_{i=0}^{k} \frac{b_i}{2^{k+1-i}} \\ 0 & \frac{1}{2^{k+1}} \end{pmatrix}$$

so, for any $k \in \{0, \dots, n-1\}$,

$$\widetilde{I}_{a_0} \cdots \widetilde{I}_{a_k} \widetilde{\beta}_{a_{k+1}} = \begin{pmatrix} \frac{1}{2} - \frac{b_{k+1}}{2} \sum_{i=0}^{k} \frac{b_i}{2^{k+1-i}} & 0 \\ -\frac{b_{k+1}}{2^{k+2}} & 0 \end{pmatrix}.$$

Hence we get

$$\begin{aligned} \mathrm{Var}(\mu_a) = {} & \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{b_0}{2} & 0 \end{pmatrix} + \sum_{k=0}^{n-1} \begin{pmatrix} \frac{1}{2} - \frac{b_{k+1}}{2} \sum_{i=0}^{k} \frac{b_i}{2^{k+1-i}} & 0 \\ -\frac{b_{k+1}}{2^{k+2}} & 0 \end{pmatrix} \right) \begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ & + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & \sum_{i=0}^{n} \frac{b_i}{2^{n+1-i}} \\ 0 & \frac{1}{2^{n+1}} \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix}. \end{aligned}$$

So

$$\begin{aligned} \mathrm{Var}(\mu_a) = {} & \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} 1 \\ -b_0 \end{pmatrix} + \sum_{k=0}^{n-1} \begin{pmatrix} 1 - b_{k+1} \sum_{i=0}^{k} \frac{b_i}{2^{k+1-i}} \\ -\frac{b_{k+1}}{2^{k+2}} \end{pmatrix} \right) \\ & + 1 + \sum_{i=0}^{n} \frac{b_i}{2^{n+1-i}} - \frac{1}{2^{n+1}}, \end{aligned}$$

which yields

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} + \frac{b_0}{2} - \frac{1}{2} \sum_{0 \le i \le k \le n-1} \frac{b_{k+1} b_i}{2^{k+1-i}} + \sum_{k=0}^{n-1} \frac{b_{k+1}}{2^{k+2}} + \sum_{i=0}^{n} \frac{b_i}{2^{n+1-i}},$$

and so

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2} \sum_{0 \le i \le k \le n-1} \frac{b_{k+1} b_i}{2^{k+1-i}} + \sum_{k=0}^{n} \frac{b_k}{2^{k+1}} + \sum_{i=0}^{n} \frac{b_i}{2^{n+1-i}},$$

which can be written

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2} \sum_{0 \le i \le k \le n-1} \frac{b_{k+1} b_i}{2^{k+1-i}} + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}},$$

or even

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=0}^{n-i} \frac{b_{k+i} b_k}{2^i} + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}}. \qquad \square$$

3.3. Generic case of the variance

In all that follows, we use the following notation:

- X denotes a sequence in $\{0, 1\}^{\mathbb{N}}$ (and we endow the set $\{0, 1\}^{\mathbb{N}}$ with the balanced Bernoulli probability measure;
- For any sequence X, define $a_X(n) = \sum_{k=0}^{n} X_k \cdot 2^k$;
- As in Theorem 3.1, for any $X \in \{0, 1\}^{\mathbb{N}}$, define the sequence $(b_j)_{j \in \mathbb{N}}$ by $b_j = (-1)^{X_j + 1}$.

We wish to prove the following:

Proposition 3.2. *For almost every $X \in \{0, 1\}^{\mathbb{N}}$*

$$\mathrm{Var}(\mu_{a_X(n)}) \underset{n\to\infty}{\sim} \frac{n}{2}.$$

In order to prove this proposition, we first need a technical lemma:

Lemma 3.3. *Let $X \in \{0, 1\}^{\mathbb{N}}$ and define the quantity:*

$$C_{2,n} = \max_{M,D} \left| \sum_{k=1}^{M} b_{k+d_1} b_{k+d_2} \right|,$$

where the maximum is taken on all $D = (d_1, d_2)$ and M such that $M + d_2 \le n$.

For almost every $X \in \{0, 1\}^{\mathbb{N}}$ and for every ε there exists $n_{\varepsilon,X}$ such that:

$$\forall n \ge n_{\varepsilon,X}, \ \left|C_{2,n}\right| < n^{\frac{1}{2}+\varepsilon}.$$

Proof. In [2], the following quantity is studied:

$$C_m\left((b_i)_{i\in\{0,\dots,n\}}\right) := \max_{M,D} \left| \sum_{k=1}^{M} b_{k+d_1} \times \dots \times b_{k+d_m} \right|,$$

where the maximum is taken on all $D = (d_1, \dots, d_m)$ and M such that $M + d_m \le n$.

So we have

$$C_{2,n} = C_2\left((b_i)_{i\in\{0,\dots,n\}}\right).$$

From (2.32) and (2.33) in [2], we know that, for any $l \geq 1$

$$\mathbb{E}\left(C_2\left((b_i)_{i\in\{0,\dots,n\}}\right)^{2l}\right) \leq 5n^{4+l}(4l)^l.$$

Let $\varepsilon > 0$; then

$$\mathbb{E}\left(\frac{C_{2,n}^{2l}}{n^{(\frac{1}{2}+\varepsilon)2l}}\right) \leq \frac{5n^{4+l}(4l)^l}{n^{(\frac{1}{2}+\varepsilon)2l}}$$

and

$$\frac{5n^{4+l}(4l)^l}{n^{(\frac{1}{2}+\varepsilon)2l}} = \frac{5(4l)^l}{n^{2\varepsilon l-4}}.$$

Now, if l is big enough, then $2\varepsilon l - 4 > 2$, and thus the series

$$\sum_{n=1}^{+\infty}\mathbb{E}\left(\frac{C_{2,n}^{2l}}{n^{(\frac{1}{2}+\varepsilon)2l}}\right)$$

converges.

By Borel-Cantelli lemma, $\frac{C_{2,n}^{2l}}{n^{(\frac{1}{2}+\varepsilon)2l}} \xrightarrow[n\to+\infty]{\text{a.s.}} 0$ and thus $\frac{C_{2,n}}{n^{(\frac{1}{2}+\varepsilon)}} \xrightarrow[n\to+\infty]{\text{a.s.}} 0$. Hence:

$$\text{a.e. } X,\ \exists n_{\varepsilon,X},\ \forall n \geq n_{\varepsilon,X},\ \frac{C_{2,n}}{n^{(\frac{1}{2}+\varepsilon)}} < 1$$

and thus

$$\left|C_{2,n}\right| < n^{\frac{1}{2}+\varepsilon}$$

for n big enough. □

Now let us prove Proposition 3.2.

Proof of Proposition 3.2. Note that, with Theorem 3.1, for any n,

$$\mathrm{Var}(\mu_a) = \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2}\sum_{i=1}^{n}\sum_{k=0}^{n-i}\frac{b_{k+i}b_k}{2^i} + \sum_{k=0}^{n}\frac{b_k - b_{n-k}}{2^{k+1}},$$

where the b_i are random variables which can take value in $\{-1, 1\}$ with probability $\frac{1}{2}$. The only thing to prove in order to get the result is that:

$$\lim_{n\to\infty}\frac{1}{n}\left(\sum_{i=1}^{n}\sum_{k=0}^{n-i}\frac{b_{k+i}b_k}{2^i}\right) = 0,$$

since it is obvious that

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{3}{2} - \frac{1}{2^{n+1}} + \sum_{k=0}^{n} \frac{b_k - b_{n-k}}{2^{k+1}}\right) = 0.$$

Let us estimate $\sum_{i=1}^{n}\sum_{k=0}^{n-i} \frac{b_{k+i}b_k}{2^i}$.

By using Lemma 3.3 and for any $\varepsilon > 0$ and for any n big enough, we get

$$\begin{aligned}\left|\sum_{i=1}^{n}\sum_{k=0}^{n-i} \frac{b_{k+i}b_k}{2^i}\right| &\leq \sum_{i=1}^{n} \frac{|C_{2,n}|}{2^i} \\ &\leq \sum_{i=1}^{n} \frac{n^{\frac{1}{2}+\varepsilon}}{2^i} \\ &\leq n^{\frac{1}{2}+\varepsilon},\end{aligned}$$

which ends the proof. □

3.4. Another proof for the typical variance

Notice that we could also do things differently in order to compute the typical variance without Lemma 3.3. Another way to write the variance, for any integer a, is the following:

$$\begin{aligned}\operatorname{Var}(\mu_a) &= \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2}\sum_{i=1}^{n}\sum_{k=0}^{n-i} \frac{b_{k+i}b_k}{2^i} + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}} \\ &= \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2}\left(\sum_{i=1}^{n}\sum_{k=0}^{n-i} \frac{1}{2^i} - 2\sum_{i=1}^{n} \frac{\sigma_i(a)}{2^i}\right) + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}},\end{aligned}$$

where

$$\begin{aligned}\sigma_i(a) = \#&\left\{\text{occurrences of } 0w1 \text{ in } \underline{a}_2 \mid |w| = i-1\right\} \\ &+ \#\left\{\text{occurrences of } 1w0 \text{ in } \underline{a}_2 \mid |w| = i-1\right\}.\end{aligned}$$

Hence

$$\begin{aligned}\operatorname{Var}(\mu_a) &= \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2}\left(\sum_{i=1}^{n} \frac{n-i+1}{2^i} - 2\sum_{i=1}^{n} \frac{\sigma_i(a)}{2^i}\right) + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}} \\ &= \frac{n+3}{2} - \frac{1}{2^{n+1}} - \frac{1}{2}\left((n+1)\left(1 - \frac{1}{2^n}\right) + \sum_{i=1}^{n} \frac{-i}{2^i} - 2\sum_{i=1}^{n} \frac{\sigma_i(a)}{2^i}\right) \\ &\quad + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}} = 1 + \frac{n}{2^{n+1}} + \frac{1}{2}\sum_{i=1}^{n} \frac{i}{2^i} + \sum_{i=1}^{n} \frac{\sigma_i(a)}{2^i} + \sum_{k=0}^{n} \frac{b_k + b_{n-k}}{2^{k+1}}.\end{aligned}$$

Now, notice that for any $(b_n)_{n\in\mathbb{N}} \in \{-1,1\}^{\mathbb{N}}$,

$$\lim_{n\to+\infty} \frac{1}{n}\left(1+\frac{n}{2^{n+1}}+\frac{1}{2}\sum_{i=1}^{n}\frac{i}{2^i}+\sum_{k=0}^{n}\frac{b_k+b_{n-k}}{2^{k+1}}\right)=0,$$

and that, having $(X_n)_{n\in\mathbb{N}}$ a sequence of independant variables indentically distributed with the balanced Bernoulli measure $\mathbb{P}$, the law of large number yields:

$$\forall i \in \mathbb{N}\backslash\{0\},\ \exists\, \mathcal{U}_i \subset \{0,1\}^{\mathbb{N}} \text{ such that } \begin{cases} \mathbb{P}(\mathcal{U}_i)=1 \\ \forall X \in \mathcal{U}_i,\ \displaystyle\lim_{n\to+\infty}\frac{\sigma_i(a_X(n))}{n}=\frac{1}{2}. \end{cases}$$

Define

$$\mathcal{U}=\bigcap_{i\in\mathbb{N}\backslash\{0\}}\mathcal{U}_i.$$

Now let us prove that for every $X \in \mathcal{U}$,

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i(a_X(n))}{2^i}=\frac{1}{2}.$$

It is easy to see that this limit exists (for any $i \le n, \sigma_i(a_X(n)) \le n$) so let us denote it by l. Let us write the following equality for any $n \in \mathbb{N}$ and any $N < n$:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i(a_X(n))}{2^i}=\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}+\frac{1}{n}\sum_{i=N+1}^{n}\frac{\sigma_i(a_X(n))}{2^i}.$$

So, for any $n \in \mathbb{N}$ and any $N < n$, we have:

$$\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}+\frac{1}{n}\sum_{i=N+1}^{n}\frac{\sigma_i(a_X(n))}{2^i}$$

$$\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}+\sum_{i=N+1}^{n}\frac{1}{2^i}$$

$$\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i(a_X(n))}{2^i}\le\frac{1}{n}\sum_{i=1}^{N}\frac{\sigma_i(a_X(n))}{2^i}+\frac{1}{2^N}.$$

Taking the limit as $n \to +\infty$ yields:

$$\frac{1}{2}\sum_{i=1}^{N}\frac{1}{2^i}\le l\le\frac{1}{2}\sum_{i=1}^{N}\frac{1}{2^i}+\frac{1}{2^N}.$$

Since this is true for all N, we have that:

$$\forall X \in \mathcal{U},\ \lim_{n\to+\infty}\frac{\mathrm{Var}(\mu_{a_X(n)})}{n}=\frac{1}{2}.$$

Remark 3.4. One can also notice, even if it is not the goal of our paper, that by doing exactly the same proof for a non balanced Bernoulli measure $(p, 1-p)$, one can prove that there exists a set $\widetilde{\mathcal{U}} \subset \{0,1\}^N$ of full measure such that:

$$\forall X \in \widetilde{\mathcal{U}}, \lim_{n\to+\infty} \frac{\mathrm{Var}(\mu_{a_X(n)})}{n} = 2p(1-p).$$

3.5. Upper bounds of the moments of $\mu_{a_X(n)}$

We now want to have bound on the moments of order $l \in \mathbb{N}$. Let us first remark that the only matrices appearing in the Taylor expansion of $\hat{A}_i(\theta)$ are I_i, α_i and β_i. Indeed:

$$\hat{A}_i(\theta)=\sum_{j=0}^{+\infty} \theta^j T_{i,j},$$

where

$$T_{i,0} = I_i, \quad T_{i,2j} = \frac{(-1)^j}{(2j)!}\beta_i, \quad T_{i,2j+1} = \frac{(-1)^j i}{(2j+1)!}\alpha_i.$$

Remark 3.5. Notice that the following relations hold:

$$I_0 \begin{pmatrix}1\\1\end{pmatrix} = I_1 \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}, \quad \alpha_0 \begin{pmatrix}1\\1\end{pmatrix} = \alpha_1 \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

and

$$\alpha_0 I_0 = \alpha_0 I_1 = \frac{1}{2}\alpha_0, \quad \alpha_1 I_0 = \alpha_1 I_1 = \frac{1}{2}\alpha_1.$$

Let us insist on the fact that these relations are crucial for our proof.

Now let us introduce the following norm on 2×2 matrices:

$$\|M\| = \max_{i\in\{1,2\}} \left(|M_{i,1}| + |M_{i,2}|\right),$$

which is induced by $\|\cdot\|_1$ on $\mathbb{R}^2$.

Notice that this defines a submultiplicative norm. Moreover,

$$\|I_0\| = \|I_1\| = \|\alpha_0\| = \|\alpha_1\| = \|\beta_0\| = \|\beta_1\| = 1.$$

Our goal is to compute all the moments of the probability measure $\mu_{a_X(n)}$ in the generic case as n goes to infinity. To that end, we arrange the terms appearing in the computation into different "types". A type is a couple (α^p, β^q) where p, q are non negative integers. They indicate the number of matrices of α and β appearing in the term. For instance, a term:

$$M = I_{a_0} \cdots I_{a_{i_0-1}} \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \alpha_{a_{i_1}} I_{a_{i_1+1}} \cdots I_{a_{i_2-1}} \beta_{a_{i_2}} I_{a_{i_2+1}} \cdots I_{a_n}$$

is of type (α^2, β^1).The order of appearance of the α and β does not have any influence on the type, so that

$$I_{a_0} \cdots I_{a_{i_0-1}} \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \beta_{a_{i_1}} I_{a_{i_1+1}} \cdots I_{a_{i_2-1}} \alpha_{a_{i_2}} I_{a_{i_2+1}} \cdots I_{a_n}$$

is also of type (α^2, β^1).

Let us denote by $\mathcal{F}^{(n)}_{(\alpha^p,\beta^q)}$ the set of all terms of type (α^p, β^q) in the expansion of $\widehat{\mu}_{a_X}(n)$.

This notation is introduced to ease the writing of the previous formulas as well as for handling terms with the same behaviour together. For instance, the formula for the variance becomes:

$$\operatorname{Var}\big(\mu_{a_X(n)}\big) = \begin{pmatrix}1 & 0\end{pmatrix} \sum_{M \in \mathcal{F}^{(n)}_{(\alpha^2,\beta^0)}} M \begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1 & 0\end{pmatrix} \sum_{M \in \mathcal{F}^{(n)}_{(\alpha^0,\beta^1)}} M \begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1 & 0\end{pmatrix} I_{a_0} \cdots I_{a_n} \begin{pmatrix}0\\2\end{pmatrix}.$$

Now notice that for any 2×2 matrix M

$$\left| \begin{pmatrix}1 & 0\end{pmatrix} M \begin{pmatrix}1\\1\end{pmatrix} \right| \le \|M\|,$$

so, in order to find an upper bound on terms of a given type, it suffices to understand

$$\sum_{M \in \mathcal{F}^{(n)}_{(\alpha^p,\beta^q)}} \|M\|.$$

Definition 3.6. We say that a type (α^p, β^q) contributes with weight at most k if

$$\sum_{M \in \mathcal{F}^{(n)}_{(\alpha^p,\beta^q)}} \|M\| = O\big(n^k\big).$$

Lemma 3.7. *For any pair of nonnegative integers (p, q), the type (α^p, β^q) contributes with weight at most q.*

Proof. We prove this lemma by induction on p.

First notice that $\#\mathcal{F}^{(n)}_{(\alpha^0,\beta^q)} = \binom{n+1}{q}$. Notice also that for any $M \in \mathcal{F}^{(n)}_{(\alpha^0,\beta^q)}$, $\|M\| \le 1$, since $\|\cdot\|$ is submultiplicative. This implies that the type $\big(\alpha^0, \beta^q\big)$ contributes with weight q.

Now let us assume that the type (α^p, β^q) contributes with weight q for a given p and let us prove that the type $\big(\alpha^{p+1}, \beta^q\big)$ has same weight. Now let us partition $\mathcal{F}^{(n)}_{(\alpha^{p+1},\beta^q)}$. Fix $k \le q$ and let us estimate terms that can be written

$M\alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \beta_{a_{i_1}} I_{a_{i_1+1}} \cdots I_{a_{i_2-1}} \beta_{a_{i_2}} \cdots I_{a_{i_k-1}} \beta_{a_{i_k}}$ where $M \in \mathcal{F}^{(i_0-1)}_{(\alpha^p,\beta^{q-k})}$. The sum of the norms of these terms is equal to:

$$
\begin{aligned}
&\sum_{p+q-k \le i_0 < \ldots < i_k \le n} \sum_{M \in \mathcal{F}^{(i_0-1)}_{(\alpha^p,\beta^{q-k})}} \left\| M\alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \beta_{a_{i_1}} \cdots \beta_{a_{i_k}} \right\| \\
&\le \sum_{p+q-k \le i_0 < \ldots < i_k \le n} \sum_{M \in \mathcal{F}^{(i_0-1)}_{(\alpha^p,\beta^{q-k})}} \left\| M\alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\| \\
&\le n^{k-1} \sum_{p+q-k \le i_0 < i_1 \le n} \sum_{M \in \mathcal{F}^{(i_0-1)}_{(\alpha^p,\beta^{q-k})}} \left\| M\alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\| \\
&\le n^{k-1} \sum_{p+q-k \le i_0 < i_1 \le n} \left\| \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\| \sum_{M \in \mathcal{F}^{(i_0-1)}_{(\alpha^p,\beta^{q-k})}} \|M\| \\
&\underset{\text{by induction}}{\le} n^{k-1} \sum_{p+q-k \le i_0 \le i_1 \le n} \left\| \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\| C i_0^{q-k} \\
&\le Cn^{q-1} \sum_{p+q-k \le i_0 < i_1 \le n} \left\| \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\|,
\end{aligned}
$$

and, thanks to Remark 3.5,

$$
\begin{aligned}
&\le Cn^{q-1} \sum_{p+q-k \le i_0 < i_1 \le n} \left\| \alpha_{a_{i_0}} I_{a_{i_0+1}} \cdots I_{a_{i_1-1}} \right\| \\
&\le Cn^{q-1} \sum_{p+q-k \le i_0 < i_1 \le n} \frac{1}{2^{i_1-i_0-1}} \\
&\le 2Cn^q.
\end{aligned}
$$

Since this computation is valid for any value of k, this yields:

$$
\sum_{M \in \mathcal{F}^{(n)}_{(\alpha^{p+1},\beta^q)}} \|M\| \le 2qCn^q,
$$

which proves the lemma. □

3.6. Computing all the moments

Let us write the expansion of $\widehat{\mu}_a$:

$$
\widehat{\mu}_a(\theta) = \sum_{k=0}^{N} \frac{i^k m_k(a)}{k!}\theta^k + o(\theta^n),
$$

where

$$m_k(a) = \sum_{d\in\mathbb{Z}} \mu_a(d)d^k$$

is the moment of order k of the probability measure μ_a (indeed, recall that μ_a is centered).

Now let us renormalize $\mu_{a_X(n)}$. From Proposition 3.2, we know that we have to look at $\widetilde{\mu}_{a_X(n)} \in l^1\left(\sqrt{\frac{2}{n}}\mathbb{Z}\right)$ defined by:

$$\forall d \in \sqrt{\frac{2}{n}}\mathbb{Z},\ \widetilde{\mu}_{a_X(n)}(d) = \mu_{a_X(n)}\left(\sqrt{\frac{n}{2}}d\right).$$

Now notice that the characteristic function of $\widetilde{\mu}_{a_X(n)}$ in θ is actually:

$$\widehat{\mu}_{a_X(n)}\left(\sqrt{\frac{2}{n}}\theta\right) = \sum_{k=0}^{N} \frac{\left(i\sqrt{2}\right)^k m_k(a_X(n))}{n^{\frac{k}{2}}k!}\theta^k + o\left(\theta^N\right).$$

Hence, for any $n \in \mathbb{N}$, the moments of order k of the probality measure $\widetilde{\mu}_{a_X(n)}$, denoted by $\widetilde{m}_k(a_X(n))$, are:

$$\widetilde{m}_k(a_X(n)) = \frac{\sqrt{2^k}m_k(a_X(n))}{n^{\frac{k}{2}}}$$

and thus, we wish to understand, if it exists, for any integer k,

$$\lim_{n\to+\infty} \frac{\sqrt{2^k}m_k(a_X(n))}{n^{\frac{k}{2}}}.$$

Lemma 3.8. *For almost every sequence $X \in \{0, 1\}^{\mathbb{N}}$ and for any $k \in \mathbb{N}$ we have:*

$$\lim_{n\to+\infty} \widetilde{m}_{2k}(a_X(n)) = \frac{(2k)!}{2^k k!}$$

and

$$\lim_{n\to+\infty} \widetilde{m}_{2k+1}(a_X(n)) = 0,$$

which are the moments of the normal law $\mathcal{N}(0, 1)$.

Proof. Remark that in this proof, we only consider terms of the following type:

$$\begin{pmatrix}1 & 0\end{pmatrix} \sum_{M\in\mathcal{F}^{(n)}_{(\alpha^p,\beta^q)}} M\begin{pmatrix}1\\1\end{pmatrix},$$

since taking a term of degree greater than 0 in the Taylor expansion of $\frac{e^{i\theta}}{2-e^{-i\theta}}$ will not contribute because it involves terms of smaller types on β.

Let us remark right away that for a moment of order $2k+1$, the type of terms which could contribute the most is $\left(\alpha^1, \beta^k\right)$. Thanks to Lemma 3.7, this type has weight at most k. Moreover, there is always a finite number of types contributing to a moment. Hence $m_{2k+1}(a_X(n)) = O(n^k)$, and thus:

$$\lim_{n\to+\infty} \frac{\sqrt{2^{2k+1}} m_{2k+1}(a_X(n))}{n^{\frac{2k+1}{2}}} = 0,$$

or, equivalently,

$$\lim_{n\to+\infty} \widetilde{m}_{2k+1}(a_X(n)) = 0.$$

Next, we consider the even moments.

For a moment m_{2k}, from Lemma 3.7, the only type potentially contributing to the limit is $\left(\alpha^0, \beta^k\right)$. More precisely, in order to get that

$$\lim_{n\to+\infty} \widetilde{m}_{2k}(a_X(n)) = \frac{(2k)!}{2^k k!},$$

one must show the following identity on limits (and prove that they exist):

$$\lim_{n\to+\infty} \frac{\left(i\sqrt{2}\right)^{2k} m_{2k}(a_X(n))}{n^k (2k)!} = \lim_{n\to+\infty} \left(\frac{2}{n}\right)^k \left(\frac{-1}{2}\right)^k (1\ 0) \sum_{M\in\mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}} M \begin{pmatrix}1\\1\end{pmatrix},$$

since the Taylor expansion of $\hat{A}_j$ near 0 is

$$\hat{A}_j(\theta) = I_j + i\theta\alpha_j - \frac{1}{2}\theta^2\beta_j + O\left(\theta^3\right).$$

This equality is equivalent to:

$$\lim_{n\to+\infty} \frac{2^k m_{2k}(a_X(n))}{n^k} = \lim_{n\to+\infty} \frac{(2k)!}{n^k} \left(1\ 0\right) \sum_{M\in\mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}} M \begin{pmatrix}1\\1\end{pmatrix}.$$

In short, we must show that:

$$\left(1\ 0\right) \sum_{M\in\mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}} M \begin{pmatrix}1\\1\end{pmatrix} = \frac{n^k}{2^k k!} + o(n^k),$$

so that

$$\lim_{n\to+\infty} \frac{2^k m_{2k}(a_X(n))}{n^k} = \frac{(2k)!}{2^k k!},$$

which is the moment of order $2k$ of the normal law $\mathcal{N}(0,1)$.

Let $\mathcal{D}_k(n) = \{(d_1, \ldots, d_k) \mid 0 \le d_1 < \ldots < d_k \le n\}$. For $d \in \mathcal{D}_k(n)$, denote $\Pi_d = \widetilde{I}_{a_0} \cdots \widetilde{I}_{a_{d_1-1}} \widetilde{\beta}_{a_{d_1}} \cdots \widetilde{I}_{a_{d_k-1}} \widetilde{\beta}_{a_{d_k}}$ (this is just a matrix $M \in \mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}$ after the change of basis described in the proof of Theorem 3.1). Let us prove by induction on k that

$$\Pi_d = \begin{pmatrix} \frac{1}{2^k} + A_d & 0 \\ B_d & 0 \end{pmatrix},$$

with A_d and B_d satisfying

$$\lim_{n \to +\infty} \frac{1}{n^k} \sum_{d \in \mathcal{D}_k(n)} |A_d| = 0$$

and

$$\lim_{n \to +\infty} \frac{1}{n^k} \sum_{d \in \mathcal{D}_k(n)} |B_d| = 0.$$

The case of $k = 1$ is treated in Lemma 3.3. Let us assume this is true up to an integer k.

Let $d \in \mathcal{D}_{k+1}(n)$. For clarity in the formulas, let us write $d = (d_1, \ldots, d_{k-1}, j, l)$ and $d' = (d_1, \ldots, d_{k-1}, j) \in \mathcal{D}_k(n)$. Then we have

$$\Pi_d = \Pi_{d'} \widetilde{I}_{a_{j+1}} \cdots \widetilde{I}_{a_{l-1}} \widetilde{\beta}_{a_l}.$$

Now compute:

$$\widetilde{I}_{a_{j+1}} \cdots \widetilde{I}_{a_{l-1}} \widetilde{\beta}_{a_l} = \begin{pmatrix} \frac{1}{2} - b_l \sum_{i=j+1}^{l-1} \frac{b_i}{2^{l-i}} & 0 \\ -\frac{b_l}{2^{l-j}} & 0 \end{pmatrix},$$

and by the induction hypothesis,

$$\Pi_{d'} = \begin{pmatrix} \frac{1}{2^k} + A_{d'} & 0 \\ B_{d'} & 0 \end{pmatrix}.$$

Thus

$$\Pi_d = \begin{pmatrix} \frac{1}{2^{k+1}} - \frac{1}{2^k} b_l \sum_{i=j+1}^{l-1} \frac{b_i}{2^{l-i}} + \frac{A_{d'}}{2} - A_{d'} b_l \sum_{i=j+1}^{l-1} \frac{b_i}{2^{l-i}} & 0 \\ \frac{1}{2} B_{d'} - B_{d'} b_l \sum_{i=j+1}^{l-1} \frac{b_i}{2^{l-i}} & 0 \end{pmatrix}$$

and we have to prove the following.

Claim.

$$\frac{1}{n^{k+1}}\sum_{d\in\mathcal{D}_{k+1}(n)}\left|-\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}+\frac{A_{d'}}{2}-A_{d'}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|\underset{n\to\infty}{\longrightarrow}0.$$

Proof of the claim.

- First, we have:

$$\sum_{d\in\mathcal{D}_{k+1}(n)}\left|\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|=\sum_{j=k}^{n-k}\sum_{0\le d_1<\ldots<d_{k-1}<j}\sum_{l=j+1}^{n}\left|\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|$$

$$\le n^{k-1}\sum_{j=k}^{n-k}\sum_{l=j+1}^{n}\left|\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|,$$

and also that

$$\sum_{l=j+1}^{n}\left|b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|\le C_{2,n}$$

so, for n big enough, according to Lemma 3.3:

$$n^{k-1}\sum_{j=k}^{n-k}\sum_{l=j+1}^{n}\left|\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|$$

$$\le\frac{n^{k-1}}{2^k}\sum_{j=k}^{n-k}n^{\frac{1}{2}+\varepsilon}$$

$$\le\frac{n^{k+\frac{1}{2}+\varepsilon}}{2^k};$$

- We have also that

$$\sum_{d\in\mathcal{D}_{k+1}(n)}\left|\frac{A_{d'}}{2}\right|=\sum_{d'\in\mathcal{D}_k(n)}\sum_{l=k}^{n}\left|\frac{A_{d'}}{2}\right|\le n\sum_{d'\in\mathcal{D}_k(n)}\left|\frac{A_{d'}}{2}\right|$$

and, by induction,

$$\lim_{n\to+\infty}\frac{1}{n^k}\sum_{d'\in\mathcal{D}_k(n)}\left|\frac{A_{d'}}{2}\right|=0;$$

hence

$$\lim_{n\to+\infty}\frac{1}{n^{k+1}}\sum_{d\in\mathcal{D}_{k+1}(n)}\left|\frac{A_{d'}}{2}\right|=0;$$

- Finally we have

$$\begin{aligned}\sum_{d\in\mathcal{D}_{k+1}(n)}\left|A_{d'}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right| &= \sum_{j=k}^{n-k}\sum_{0\le d_1<\ldots<d_{k-1}<j}\sum_{l=j+1}^{n}\left|A_{d'}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right| \\ &\le \sum_{j=k}^{n-k}\sum_{0\le d_1<\ldots<d_{k-1}<j}|A_{d'}|\sum_{l=j+1}^{n}\left|b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|\end{aligned}$$

and, as in the study of the first term, using Lemma 3.3, for n big enough

$$\begin{aligned}&\sum_{j=k}^{n-k}\sum_{0\le d_1<\ldots<d_{k-1}<j}|A_{d'}|\sum_{l=j+1}^{n}\left|b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right| \\ &\le \sum_{j=k}^{n-k}\sum_{0\le d_1<\ldots<d_{k-1}<j}|A_{d'}|\,n^{\frac{1}{2}+\varepsilon} \\ &\le n^{\frac{1}{2}+\varepsilon}\sum_{d'\in\mathcal{D}_k(n)}|A_{d'}| \\ &\le n^{k+\frac{1}{2}+\varepsilon}\end{aligned}$$

by induction hypothesis.

In the end,

$$\lim_{n\to+\infty}\frac{1}{n^{k+1}}\sum_{d\in\mathcal{D}_{k+1}(n)}\left|\frac{1}{2^k}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}+\frac{A_{d'}}{2}-A_{d'}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|=0.$$

And the same goes for proving that:

$$\frac{1}{n^{k+1}}\sum_{d\in\mathcal{D}_{k+1}(n)}\left|\frac{1}{2}B_{d'}-B_{d'}b_l\sum_{i=j+1}^{l-1}\frac{b_i}{2^{l-i}}\right|\xrightarrow[n\to\infty]{}0.$$

Hence, we get

$$\begin{aligned}(1\ 0)\sum_{M\in\mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}} M\begin{pmatrix}1\\1\end{pmatrix} &= \begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\end{pmatrix}\sum_{d\in\mathcal{D}_k(n)}\Pi_d\begin{pmatrix}2\\0\end{pmatrix}\\ &= \sum_{d\in\mathcal{D}_k(n)}\frac{1}{2^k}+A_d-\sum_{d\in\mathcal{D}_k(n)}B_d\\ &= \sum_{d\in\mathcal{D}_k(n)}\frac{1}{2^k}+\sum_{d\in\mathcal{D}_k(n)}A_d-\sum_{d\in\mathcal{D}_k(n)}B_d\\ &= \binom{n}{k}\frac{1}{2^k}+\sum_{d\in\mathcal{D}_k(n)}A_d-\sum_{d\in\mathcal{D}_k(n)}B_d\\ &= \frac{n!}{k!(n-k)!2^k}+\sum_{d\in\mathcal{D}_k(n)}A_d-\sum_{d\in\mathcal{D}_k(n)}B_d.\end{aligned}$$

Since

$$\lim_{n\to+\infty}\frac{1}{n^k}\sum_{d\in\mathcal{D}_k(n)}|A_d|=0$$

and

$$\lim_{n\to+\infty}\frac{1}{n^k}\sum_{d\in\mathcal{D}_k(n)}|B_d|=0,$$

we have that

$$(1\ 0)\sum_{M\in\mathcal{F}^{(n)}_{(\alpha^0,\beta^k)}} M\begin{pmatrix}1\\1\end{pmatrix} \underset{n\to+\infty}{=} \frac{n^k}{2^k k!}+o\big(n^k\big),$$

which yields

$$\lim_{n\to+\infty}\frac{2^k m_{2k}(a_X(n))}{n^k}=\frac{(2k)!}{2^k k!}. \qquad \square$$

Hence the moments of the probability measure $\widetilde{\mu}_{a_X(n)}$ converge towards the moments of the normal law $\mathcal{N}(0,1)$, which prove the [14, Central Limit Theorem 1.10].

References

[1] J. BÉSINEAU, *Indépendance statistique d'ensembles liés à la fonction "somme des chiffres"*, In: "Séminaire Delange-Pisot-Poitou, 13e année (1971/72), Théorie des nombres", Fasc. 2, Exp. No. 23, page 8. Secrétariat Mathematique, Paris, 1973.

[2] J. CASSAIGNE, C. MAUDUIT and A. SÁRKÖZY, *On finite pseudorandom binary sequences. VII. The measures of pseudorandomness*, Acta Arith. **103** (2002), 97–118.

[3] T. W. CUSICK, Y. LI and P. STĂNICĂ, *On a combinatorial conjecture*, Integers **11** (2011), 185–203.

[4] H. DELANGE, *Sur les fonctions q-additives ou q-multiplicatives*, Acta Arith. **21** (errata insert) (1972), 285–298.

[5] H. DELANGE, *Sur la fonction sommatoire de la fonction"somme des chiffres"*, Enseignement Math. (2) **21** (1975), 31–47.

[6] M. DRMOTA, *The joint distribution of q-additive functions*, Acta Arith. **100** (2001), 17–39.

[7] M. DRMOTA, M. KAUERS and L. SPIEGELHOFER, *On a conjecture of Cusick concerning the sum of digits of n and $n+t$*, SIAM J. Discrete Math. **30** (2016), 621–649.

[8] M. DRMOTA, C. MAUDUIT and J. RIVAT, *Primes with an average sum of digits*, Compos. Math. **145** (2009), 271–292.

[9] J. EMME and A. PRIKHODKO, *On the asymptotic behaviour of the correlation measure of sum-of-digits function in base 2*, preprint, 2015, arXiv:1504.01701.

[10] M. KEANE, *Generalized Morse sequences*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **10** (1968), 335–353.

[11] D.-H. KIM, *On the joint distribution of q-additive functions in residue classes*, J. Number Theory **74** (1999), 307–336.

[12] J. Y. KIM, *Integer addition and hamming weight*, preprint, 2015, arXiv:1503.01170v2.

[13] S. KOPPARTY, *On the complexity of powering in finite fields*, In: "STOC'11-Proceedings of the 43rd ACM Symposium on Theory of Computing", ACM, New York, 2011, 489–498.

[14] A. A. MARKOV, "Démonstration du second théorème-limite du calcul des probabilités par la méthode des moments", Saint-Petersbourg, 1913.

[15] J. F. MORGENBESSER and L. SPIEGELHOFER, *A reverse order property of correlation measures of the sum-of-digits function*, Integers **12** (2012), 5.

LMO, CNRS - UMR 8628
Université Paris-Sud 11
91405 Orsay, France
jordan.emme@math.u-psud.fr

Aix Marseille Unversité
CNRS, Centrale Marseille
Institut de Mathematiques de Marseille
12M – UMR 7373
13453 Marseille, France
pascal.hubert@univ-amu.fr