Recurrence of substitutive Sturmian words

Pablo Rotondo
LITIS, Université de Rouen, France
pablo.rotondo@univ-rouen.fr

Brigitte Vallée
GREYC, CNRS and Université de Caen.
brigitte.vallee@unicaen.fr

Abstract

This article studies the probabilistic behaviour of Sturmian words. Sturmian words are the simplest infinite words that are not periodic, and the paper focuses on a key subfamily of Sturmian words, the substitutive Sturmian words, built by the application of letter substitutions. With each Sturmian word, one associates a real number of the unit interval, called its slope. Whereas generic Sturmian words are related to general irrational numbers, substitutive Sturmian words have their slope that is a quadratic irrational number. Thus, substitutive Sturmian words form a discrete subset of the whole set of Sturmian words.

The recurrence function can be viewed as a waiting time to discover all of the factors of a given length of an infinite word. The recurrence function has been already studied by the authors in [5] and [17] in the case of a random generic Sturmian word. We wish to compare the behaviour of the recurrence function in the two cases –generic Sturmian words and substitutive Sturmian words–. This article thus performs the probabilistic analysis of the recurrence function in the case of a random substitutive Sturmian word, and proves that the distribution of this function behaves in a strongly similar way in both cases –generic Sturmian words and substitutive Sturmian words–.

Even though the behaviour of the recurrence function is proven to be similar in the two cases, the methods are completely different in the two papers. Many of the characteristics of a Sturmian word is described by the continued fraction expansion of its slope. This is in particular the case for the recurrence function, as shown by Morse and Hedlund in [16]. Whereas the slope of a generic Sturmian word is any (irrational) number of the unit interval, the slope of a substitutive Sturmian is a quadratic irrational, whose continued fraction expansion is ultimately periodic. We are led to study (in a probabilistic setting) fine parameters of (ultimate) periodic continued fraction expansions.

In this discrete setting, we then use tools from Analytic Combinatorics, in particular, generating functions. The dynamical system which produces continued fraction expansion (related to the celebrated Gauss map) plays an important role in the analysis, via the transfer operators that are associated to this system. Such operators play the role of generating operators that themselves generate the generating functions of interest.

2012 ACM Subject Classification Theory of computation → Pattern matching; Theory of computation → Randomness, geometry and discrete structures

Keywords and phrases Combinatorics of words, Sturmian words, Substitutive Sturmian words, Recurrence, Probabilistic analysis, Continued fraction expansion, Transfer operator, Analytic combinatorics, Generating functions, Mellin transform

Digital Object Identifier 10.4230/LIPIcs...

Funding Projects DynA3S and CoDys (ANR), Project AleaEnAmSud (STICAmSud), Projet RIN Alenor.

Acknowledgements We thank Valérie Berthé and Eda Cesaratto for many interesting discussions.

1 Introduction

The recurrence function of an infinite word measures the cost of discovering all of its factors of a given length n. This makes the recurrence function a fundamental measure of complexity, that is widely studied in Combinatorics on Words, particularly for a fundamental family of words known as *Sturmian words*. Sturmian words are defined as the simplest infinite words that are not eventually periodic. Such words turn up naturally in relation to digital geometry and quasicrystals, and have been the subject of many studies over the years.
Morse and Hedlund [16] provide a powerful arithmetic description of Sturmian words. They prove that each Sturmian word is associated with an irrational number α called its slope. Many characteristics of the Sturmian words depend on the slope α only, notably via its continued fraction expansion $\texttt{Cfe}(\alpha)$. This is in particular the case for the recurrence.

In the present paper, we study an important *subfamily* of Sturmian words that are constructed through the use of letter subtitutions. An instance of such a word is the celebrated Fibonacci word obtained as a fixed point of the substitution $\sigma\colon \mathbf{0} \mapsto \mathbf{01}$, $\mathbf{1} \mapsto \mathbf{0}$. The Sturmian words thus produced are called *substitutive* Sturmian words and constitute a very attractive subfamily of Sturmian words, that is central in Combinatorics on Words. The slope of a *substitutive* Sturmian word is proven to be a *quadratic irrational* number (qi in short)[1]. Then, the Cfe of its slope is ultimately periodic. More precisely, the qi numbers that arise as slopes of substitutive Sturmian words (called here sqi numbers) are characterized in [10]: their Cfe have a (small) preperiod, with a slight constraint between period and preperiod.

Results. As the recurrence function highly depends on $\texttt{Cfe}(\alpha)$, the recurrence of substitutive Sturmian words may appear *a priori* very particular. We thus perform a probabilistic study of the recurrence function of substitutive Sturmian words, comparing it to that of general Sturmian words, that has been already studied by the authors in [5] and [17]. We obtain in Theorem 8 a new result that describes the probabilistic behaviour of the recurrence on this class of substitutive Sturmian words and exhibits a strong similarity of the behaviour recurrence on the two classes (generic and substitutive).

Methods. Even if the results are similar, the methods are completely different. Substitutive Sturmian words form a discrete structure that is completely invisible to our previous analysis of [17] which deals with general Sturmian words, associated with real slopes. We then work here inside the framework of Analytic Combinatorics, dealing with three main tools described in Sections 5.1 and 5.2: generating functions, Mellin transforms and transfer operators. First, we introduce (as usual) the generating functions (of Dirichlet type) associated with our cost of interest (related to the distribution of the recurrence function); we are interested in the behaviour of the recurrence when the length n of factors tends to ∞. Mellin transform is a great tool for dealing with such asymptotics, and we are led to nest Mellin transforms into generating functions.
As continued fraction expansions are built with the Gauss map described in (5), we deal with the underlying dynamical system, and its associated transfer operators defined in (20); such operators naturally follow the construction of the $\texttt{Cfe}'s$, and the main operations on these $\texttt{Cfe}'s$. This is why the generating functions (of Mellin transforms) are expressed with the associated transfer operator, in the lines of other works, such as [2, 20, 19]. Moreover, such transfer operators, via their trace, are also well adapted to the description of cyclic structures, as in [9]. In conclusion, our main object is a *generating* function of the *Mellin*

[1] This means : an irrational number which is a solution of a quadratic equation with integer coefficients

transform of the cost, that is expressed in terms of *transfer operators*.

Plan of the paper. Sections 2 and 3 describe the general framework and Section 4 states our main result. Then, Section 5 introduces the toolbox and describes (in an informal style) the two main steps of the proof: first the algebraic step, second, the analytic step. The Annex provides detailed and formal statements as well their precise proofs.

2 Sturmian words, substitutive Sturmian words and recurrence.

This section first presents Sturmian words, then substitutive Sturmian words. It focuses on the main function of interest, the recurrence function. Then, it describes our point of view on continuant fraction expansion in 3 and introduces the notion of a continuant function, together with the general framework of our study.
We consider a finite set $\mathcal{A}$ of *symbols*, called *alphabet*. Let $\boldsymbol{u} = (u_n)_{n\in\mathbb{N}}$ be an infinite word in $\mathcal{A}^{\mathbb{N}}$. A finite word w of length n is a factor of $\boldsymbol{u}$ if there exists an index m for which $w = u_m \dots u_{m+n-1}$. Let $\mathcal{L}_{\boldsymbol{u}}(n)$ stand for the set of factors of length n of $\boldsymbol{u}$.

The (*factor*) *complexity function* of the infinite word $\boldsymbol{u}$ is defined as the sequence $n \mapsto p_u(n) := |\mathcal{L}_{\boldsymbol{u}}(n)|$. The eventually periodic words are the simplest ones, in terms of the complexity function. They are characterized by the inequality $p_{\boldsymbol{u}}(n) \le n$ for some n.

2.1 Sturmian words.

The simplest words that are not eventually periodic satisfy the equality $p_{\boldsymbol{u}}(n) = n+1$ for each $n \ge 0$. Such words do exist, they are called *Sturmian words*. Moreover, Morse and Hedlund provided in [16] a powerful arithmetic description of Sturmian words (see also [15] for more on Sturmian words).

▶ **Proposition 1.** [Morse and Hedlund] *Associate with a pair* $(\alpha, \beta) \in [0,1]^2$ *the two infinite words* $\underline{\mathfrak{S}}(\alpha,\beta)$ *and* $\overline{\mathfrak{S}}(\alpha,\beta)$ *whose* n*-th symbols are respectively*

$$\underline{u}_n = \lfloor \alpha(n+1) + \beta \rfloor - \lfloor \alpha n + \beta \rfloor, \qquad \overline{u}_n = \lceil \alpha(n+1) + \beta \rceil - \lceil \alpha n + \beta \rceil.$$

Then a word $\boldsymbol{u} \in \{0,1\}^{\mathbb{N}}$ *is Sturmian if and only if it equals* $\underline{\mathfrak{S}}(\alpha,\beta)$ *or* $\overline{\mathfrak{S}}(\alpha,\beta)$ *for a pair* (α,β) *formed with an irrational* $\alpha \in]0,1[$ *and a real* $\beta \in [0,1[$.

2.2 Substitutive Sturmian words.

There is an important *subfamily* of Sturmian words that gathers the *substitutive* Sturmian words.

▶ **Definition 2** (Substitutive words)**.** *A word* $\boldsymbol{u}$ *is said to be* substitutive *if and only if it is the fixed point of a* primitive *morphism* σ*, i.e.,* $\sigma(\boldsymbol{u}) = \boldsymbol{u}$.

We recall some facts about morphisms; A morphism $\sigma\colon \{0,1\}^* \to \{0,1\}^*$ is defined by the values of $\sigma(0)$ and $\sigma(1)$, as well as the property $\sigma(a \cdot b) = \sigma(a)\cdot\sigma(b)$. Such a morphism may be extended to act on infinite words from $\{0,1\}^{\mathbb{N}}$. An infinite word $\boldsymbol{u}$ that satisfies $\sigma(\boldsymbol{u}) = \boldsymbol{u}$ is called a fixed point of σ.
For $w \in \{0,1\}^*$ and $a \in \{0,1\}$, the number of occurrences of symbol a in w is denoted by $|w|_a$. We associate to σ the matrix M_σ with coefficients $\sigma_{i,j} := |\sigma(i)|_j$. A morphism $\sigma\colon \{0,1\}^* \to \{0,1\}^*$ is said to be primitive if and only if the matrix M_σ is primitive. In this case, the underlying graph is strongly connected, and this eliminates the trivial morphisms.

We now recall the characterization of slopes of substitutive Sturmian words. It is first shown that such a slope is a quadratic irrational number (qi number), namely the solution of a quadratic equation with integer coefficients. A famous theorem due to Lagrange (see [13, pp.183–184]) states that quadratic irrational numbers α coincide with numbers α whose continued fraction expansion $\texttt{Cfe}(\alpha)$ is ultimately periodic. Then, the following result obtained in [10] (Thm 3) gives a complete characterization of $\texttt{Cfe}(\alpha)$ for the slope α of a substitutive Sturmian word: it has a short preperiod (of length at most 2), and a period, and there is only a condition between the preperiod and the last digit of the period. Then, the slope α is closely related to the qi number β defined by the period. Such a qi β (without period) is called a *reduced* quadratic irrational (rqi).

▶ **Proposition 3.** *The two assertions are equivalent for a Sturmian word* $\underline{\mathfrak{S}}(\alpha, \alpha)$

(i) *It is substitutive;*

(ii) *Its slope α is a quadratic irrational number, and:*

- (a) *if $\alpha > 1/2$, then $\texttt{Cfe}(\alpha)$ has a preperiod $[1, m_0]$ of length 2, with $m_0 \geq 1$ and a period which ends with a digit $m \geq m_0$;*
- (b) *if $\alpha < 1/2$, then $\texttt{Cfe}(\alpha)$ has a preperiod $[m_0]$ of length 1, with $m_0 \geq 2$ and a period which ends with a digit $m \geq m_0 - 1$;*

Partial proof. We consider a morphism σ, and the Sturmian word $\boldsymbol{u}$ that is its fixed point. We analyze the frequenciy of '1' after an application of σ: one has, for any prefix $\boldsymbol{v}$ of $\boldsymbol{u}$,

$$\frac{|\sigma(\boldsymbol{v})|_1}{|\sigma(\boldsymbol{v})|_0} = \frac{|\boldsymbol{v}|_0\, \sigma_{0,1} + |\boldsymbol{v}|_1\, \sigma_{1,1}}{|\boldsymbol{v}|_0\, (\sigma_{0,0} + \sigma_{0,1}) + |\boldsymbol{v}|_1\, (\sigma_{1,1} + \sigma_{1,0})}\,, \qquad \text{with } \sigma_{i,j} := |\sigma(i)|_j\,.$$

Now, Prop. 1 proves that the slope α is the limit frequency of '1'; then, the following holds

$$\lim_{|\boldsymbol{v}|\to\infty} \frac{|\boldsymbol{v}|_1}{|\boldsymbol{v}|_0} = \lim_{|\boldsymbol{v}|\to\infty} \frac{|\sigma(\boldsymbol{v})|_1}{|\sigma(\boldsymbol{v})|_0} = \frac{\alpha}{1-\alpha},$$

and entails that $\alpha/(1-\alpha)$ is solution of the equation $x = \dfrac{\sigma_{0,1} + x\sigma_{1,1}}{(\sigma_{0,1} + \sigma_{0,0}) + x\,(\sigma_{1,1} + \sigma_{1,0})}$. ◀

2.3 Recurrence

It is also important to study where finite factors occur inside the infinite word u. An infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$ is *uniformly recurrent* if every factor of $\boldsymbol{u}$ appears infinitely often and with bounded gaps. More precisely, denote by $w_{\boldsymbol{u}}(q, n)$ the minimal number of symbols u_k with $k \geq q$ which have to be inspected for discovering the whole set $\mathcal{L}_{\boldsymbol{u}}(n)$ from the index q. Then, the integer $w_{\boldsymbol{u}}(q, n)$ is a sort of "waiting time" and $\boldsymbol{u}$ is uniformly recurrent if each set $\{w_{\boldsymbol{u}}(q, n) \mid q \in \mathbb{N}\}$ is bounded, and the *recurrence function* $n \mapsto R_{\boldsymbol{u}}(n)$ is defined as

$$R_{\boldsymbol{u}}(n) := \max\{w_u(q, n) \mid q \in \mathbb{N}\}.$$

We then recover the usual definition: Any factor of length $R_{\boldsymbol{u}}(n)$ of u contains all the factors of length n of $\boldsymbol{u}$, and the length $R_{\boldsymbol{u}}(n)$ is the smallest integer which satisfies this property. The inequality $R_{\boldsymbol{u}}(n) \geq p_{\boldsymbol{u}}(n) + n - 1$ thus holds.

Any Sturmian word is uniformly recurrent. Its recurrence function only depends on the slope α and is thus denoted by $n \mapsto R(\alpha, n)$. Furthermore, Morse and Hedlund prove in [16] that it only depends on α via the sequence $\texttt{Cfe}(\alpha)$, and more precisely via the sequence of its continuants. The continuant $q_k(\alpha)$ is the denominator of the k-th convergent of α and the sequence $k \mapsto q_k(\alpha)$ is strictly increasing. We will return to continuants in Section 3.

▶ **Proposition 4** (Morse and Hedlund)**.** *For a Sturmian word with slope α, the recurrence function $R(\alpha, n)$ satisfies*

$$R(\alpha, n) = n - 1 + q_k(\alpha) + q_{k-1}(\alpha)\,, \qquad \textit{for every } n \in [q_{k-1}(\alpha), q_k(\alpha)[\,. \tag{1}$$

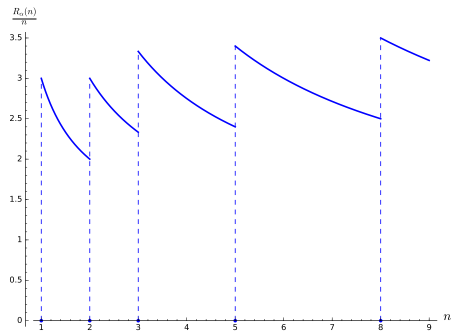
The recurrence function is our main subject of interest. It proves more useful to study the recurrence quotient defined, for $n \in [q_{k-1}(\alpha), q_k(\alpha)[$ as

$$T(\alpha, n) := \frac{R(\alpha, n) + 1}{n} = 1 + \frac{q_{k-1}(\alpha)}{n} + \frac{q_k(\alpha)}{n} . \tag{2}$$

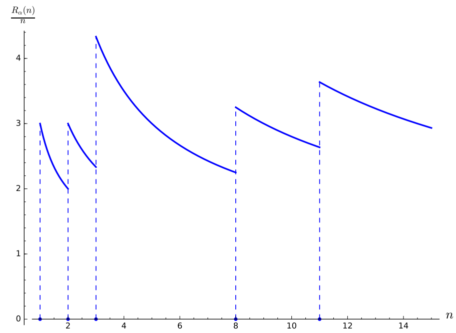
Most of the classical studies of the recurrence function deal with its extremal cases for a *fixed* irrational $\alpha \in [0, 1]$. Morse and Hedlund exhibit in [16] the worst case growth of the recurrence function for general α:

▶ **Proposition 5** (Morse and Hedlund). *For almost any irrational $\alpha \in [0, 1]$, and any $\epsilon > 0$ the recurrence quotient $T(\alpha, n)$ defined in (2) satisfies*

$$\limsup_{n\to\infty} \frac{T(\alpha, n)}{\log n} = \infty , \qquad \limsup_{n\to\infty} \frac{T(\alpha, n)}{(\log n)^{1+\epsilon}} = 0 .$$



(a) Recurrence quotient for $\alpha = \varphi^2$, with $\varphi = (\sqrt{5} - 1)/2$.



(b) Recurrence quotient for $\alpha = 1/e$.

Figure 1 Examples of recurrence quotients for two Sturmian words.

As shown in Fig. 1, $T(\alpha, n)$ is maximal when n is on the *left* of the interval $[q_{k-1}(\alpha), q_k(\alpha)[$ and highly depends on the behaviour of the digits m_i in the $\texttt{Cfe}(\alpha)$. On the other hand, when n is close to the *right* border of the intervals $[q_{k-1}(\alpha), q_k(\alpha)[$, the recurrence quotient satisfies $T(\alpha, q_k(\alpha) - 1) \leq 3$. This highlights the influence of the position of the integer n within the interval $[q_{k-1}(\alpha), q_k(\alpha)[$.

2.4 $\mathcal{Q}$-functions.

The recurrence quotient defined in (2) satisfies, for $n \in [q_{k-1}(\alpha), q_k(\alpha)[$,

$$T(\alpha, n) = 1 + \frac{q_{k-1}(\alpha)}{n} + \frac{q_k(\alpha)}{n} = 1 + \frac{q_{k-1}(\alpha)}{q_k(\alpha)} \frac{q_k(\alpha)}{n} + \frac{q_k(\alpha)}{n} .$$

This function only depends there on the pair (x, y) with $x = \frac{q_{k-1}(\alpha)}{q_k(\alpha)}$, $y = \frac{q_k(\alpha)}{n}$. Moreover, the pair (x, y) belongs to the domain $\mathcal{D} := \{(x, y) \mid 0 \leq xy \leq 1, y \geq 1\}$.

This leads to more general functions, of the same flavour as the recurrence quotient, called continuant functions, and denoted as $\mathcal{Q}$-functions. Such a function, denoted by Λ, is defined via its base function $f : \mathcal{D} \to \mathbb{R}_{\geq 0}$ assumed to be of class $\mathcal{C}^1$ as follows : given a general slope $\alpha \in [0, 1]$, its sequence of continuants $\ell \mapsto q_\ell(\alpha)$ and an integer n, we let

$$\Lambda(\alpha, n) := \Lambda_f(\alpha, n) := f\left(\frac{q_{k-1}(\alpha)}{q_k(\alpha)}, \frac{q_k(\alpha)}{n}\right), \qquad \text{for} \quad n \in [q_{k-1}(\alpha), q_k(\alpha)[. \tag{3}$$

The main function of interest, the recurrence function, is a continuant function associated with $f(x,y) = 1 + y(x+1)$. Remark that the function $n \mapsto \Lambda(\alpha, n)$ is easily extended to real numbers of the interval $[1,\infty]$ and is zero on $[0,1]$. We are interested in pairs (α, n) for which the function $\Lambda(\alpha, n)$ satisfies $\Lambda(\alpha, n) \le \lambda$. This leads to domains $\mathcal{D}_\lambda$ represented in Fig. 2,

$$\mathcal{D}_\lambda := \{(x,y) \in \mathcal{D} \mid f(x,y) \le \lambda\}, \qquad \mathcal{D} := \{(x,y) \mid 0 \le xy \le 1, y \ge 1\}. \tag{4}$$

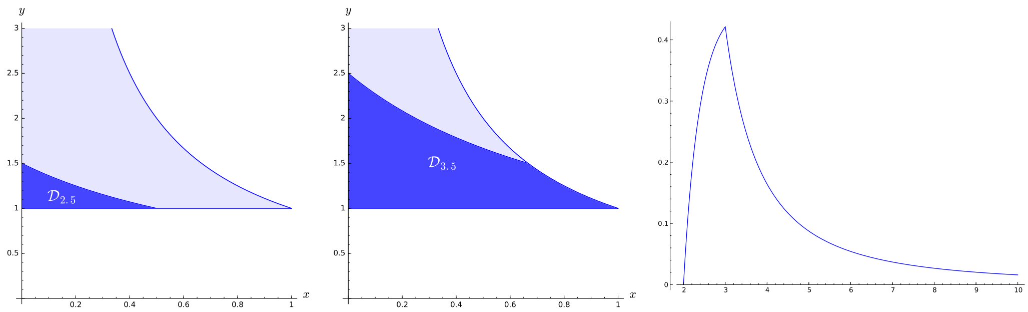


Figure 2 Various examples in the case of the base function $f(x,y) = 1 + y(1+x)$ associated with the recurrence quotient. On the left and middle, the domain $\mathcal{D}$ in light blue, and two examples of $\mathcal{D}_\lambda$ in darker blue – for $\lambda = 5/2$ on the left, and $\lambda = 7/2$ in the middle –. Notice the change of behavior at $\lambda = 3$ due to the *cusp* at the point $(1,1)$. On the right, we consider the graph of the density function $\lambda \mapsto R'(\lambda)$ [where $R(\lambda)$ is defined in (13)]. The function $R'(\lambda)$ satisfies

$$\frac{12}{\pi^2}\frac{1}{\lambda-1}\log(\lambda-1) \quad \text{if } 2 \le \lambda \le 3, \qquad \frac{12}{\pi^2}\frac{1}{\lambda-1}\log(1+\frac{1}{\lambda-2}) \quad \text{if } \lambda \ge 3\,.$$

3 Sturmian words and continued fraction expansions

3.1 Continued fraction expansions

They are built with the Gauss map $T : [0,1] \to [0,1]$ defined by

$$T(0) = 0, \qquad T(x) := \left(\frac{1}{x}\right) - \left\lfloor\frac{1}{x}\right\rfloor \qquad \text{with} \quad m(x) := \left\lfloor\frac{1}{x}\right\rfloor \qquad (\text{for } x \neq 0)\,. \tag{5}$$

(Here $\lfloor\cdot\rfloor$ is the integer part). With any (irrational) number $\alpha \in]0,1]$, this gives rise to its (infinite) continued fraction expansion

$$\texttt{Cfe}(\alpha) := \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}} = [m_1, m_2, \ldots, m_k \ldots] \qquad\qquad m_i = m(T^{i-1}(\alpha))\,.$$

We deal with linear fractional transformations (in shorthand LFT), namely mappings of the form $g : x \mapsto (ax+b)/(cx+d)$, defined with the truncatures of $\texttt{Cfe}(\alpha)$: With any integer $m \ge 1$, we associate the LFT $h_m : x \mapsto 1/(m+x)$, and, then, with each prefix $\boldsymbol{m} := [m_1, m_2, \ldots m_k]$

of $\mathtt{Cfe}(\alpha)$ of length $\langle \boldsymbol{m} \rangle = k$, we associate the *prefix* LFT $h_{\boldsymbol{m}} := h_{m_1} \circ h_{m_2} \circ \ldots \circ h_{m_k}$. We then deal with the following sets of LFT's

$$\mathcal{H} := \{h_m : m \geq 1\}, \quad \mathcal{H}^k := \{h_{\boldsymbol{m}} \mid \langle \boldsymbol{m} \rangle = k\}, \quad \mathcal{H}^+ = \bigcup_{k \geq 1} \mathcal{H}^k \,, \tag{6}$$

and we associate with α the sequence $\mathtt{Seq}(\alpha) := \{h_{\boldsymbol{m}} \mid \boldsymbol{m} \text{ is a prefix of } \mathtt{Cfe}(\alpha)\}$.

For $g \in \mathtt{Seq}(\alpha)$, the *continuant* $q(g)$ is defined as the denominator of the rational $g(0)$. The LFT $b(g)$ is the prefix of g of depth $\langle g \rangle - 1$. For $g \in \mathtt{Seq}(\alpha)$, the intervals $Q(g) := [q(b(g)), q(g)[$ form a partition of the interval $[1, \infty[$.

The mirror operation is very important in the context: it maps a prefix $\boldsymbol{m} := [m_1, m_2, \ldots, m_k]$ to its mirror $\hat{\boldsymbol{m}} := [m_k, \ldots, m_2, m_1]$. It is extended to LFT's, and, the mirror $\hat{g}$ is the LFT associated with prefix $\hat{\boldsymbol{m}}$, i.e., $\hat{g} = h_{\hat{\boldsymbol{m}}}$ for $g := h_{\boldsymbol{m}}$. As any g is a linear fractional transformation of the form $g : x \mapsto (ax + b)/(cx + d)$ with (a, b, c, d) coprime and $\det g := ad - bc = \pm 1$, the equalities hold

$$q(g) = |g'(0)|^{-1/2}, \qquad g'(0) = \hat{g}'(0), \qquad \hat{g}(0) = q[b(g)]/q[g]\,; \tag{7}$$

3.2 Return To $\mathcal{Q}$-functions.

We now describe continuant functions with this point of view. We consider the function L_λ indicator of the set $\mathcal{D}_\lambda$ defined in (4), and introduce the functions A_g defined as

$$A_g(t, n, \lambda) := L_\lambda \left(\hat{g}(t), \frac{1}{n\, |\hat{g}'(t)|^{1/2}} \right) \tag{8}$$

With a pair (α, n), with $\alpha \in [0, 1]$ and $n \geq 1$, we associate the unique LFT $g \in \mathtt{Seq}(\alpha)$ for which n belongs to the interval $[q(b(g)), q(g)[$. Then, Eq. (7) entails the following

$$n \in [q(b(g)), q(g)] \Rightarrow [\![\Lambda(\alpha, n) \leq \lambda]\!] = A_g(0, n, \lambda), \quad [\![\Lambda(\alpha, n) \leq \lambda]\!] = \sum_{g \in \mathtt{Seq}(\alpha)} A_g(0, n, \lambda)\,. \tag{9}$$

3.3 Return to Substitutive Sturmian words.

Proposition 3 leads to a characterization of substitutive Sturmian words in this context:

▶ **Lemma 6.** *The* $\mathtt{Cfe}$ *of a substitutive Sturmian word, called a sqi number, is written as* $p \cdot (g \circ r)^\star$ *where the triple* (p, r, g) *satisfies the following constraints:*

(i) $p \in \mathcal{H} \cup \{h_1\} \times \mathcal{H}$; $\;g \in \mathcal{H}^\star$; $\;r \in \mathcal{H}$;

(ii) *The inequality* $q(r) \leq q(p)$ *holds between the continuants* $q(p)$ *and* $q(r)$ *of LFT's* p *and* r.

The size $\epsilon(\alpha)$ of a rqi α is mathematically defined as the fundamental unit of the associated quadratic field, related to the *primitive even* period which appears in $\mathtt{Cfe}(\beta)$. Here, we adopt a simplified form and always define the size $\epsilon(\alpha)$ from the primitive period h of $\mathtt{Cfe}(\alpha)$, as

$$\epsilon(\alpha) := |h'(h^\star)|^{-1/2}, \qquad h \;\text{ is the primitive LFT associated with } \alpha\,. \tag{10}$$

There is a natural size of a $\mathcal{S}$-Sturmian word of slope α for which $\mathtt{Cfe}(\alpha)$ is of the form $p \cdot h^\star = p \cdot (g \circ r)^\star$

$$\epsilon(\alpha) = |p'(\hat{h}^\star)|^{-1/2} \cdot |h'(h^\star)|^{-1/2} = |p'(\hat{h}^\star)|^{-1/2} \cdot |\hat{h}'(\hat{h}^\star)|^{-1/2} = \ldots\ldots \tag{11}$$

These mathematical notions of size are closely related to the usual sizes in the computer science meaning: in both cases, $\log \epsilon(\alpha)$ is proportional to the space needed to represent α via its $\mathtt{Cfe}$.

4 Main results : Probabilistic behaviour of the recurrence function.

We here analyze the probabilistic behaviour of a general $\mathcal{Q}$–function $\alpha \mapsto \Lambda(\alpha, n)$ in two cases. We first recall in Theorem 7, (Section 4.1), the result that deals with a general slope $\alpha \in [0,1]$, and thus studies the probabilistic behaviour of recurrence for generic Sturmian words. Then, Section 4.2 focuses on Substitutive Sturmian words and states our new result Theorem 8. This section thus deals with slopes that are quadratic irrational numbers with a small pre-period, and may be also applied when there is no preperiod.

Our results exhibit the existence of a universal density ϖ on the domain $\mathcal{D}$,

$$\varpi(x,y) := \frac{12}{\pi^2}\frac{1}{y(1+x)}, \qquad \mathcal{D} := \{(x,y) \in \mathbb{R}^2 \mid 0 \le xy \le 1, y > 1\}\,. \tag{12}$$

When we focus on a continuant function Λ associated with a base function f defined on $\mathcal{D}$, we consider the domain $\mathcal{D}_\lambda$ and the integral $R(\lambda)$ of ϖ on $\mathcal{D}_\lambda$, namely

$$\mathcal{D}_\lambda := \{(x,y) \in \mathcal{D} \mid f(x,y) \le \lambda\}, \qquad R(\lambda) := \iint_{\mathcal{D}_\lambda} \varpi(x,y)\,dxdy\,. \tag{13}$$

Then, the probabilistic behaviour of the continuant function Λ, associated with the measure of the events $[\![\Lambda(\alpha,n) \le \lambda]\!]$ will be described by the integrals $R(\lambda)$, and the derivative $R'(\lambda)$ is thus viewed as a limit density function. Its graph is depicted in Figure 2 in the case of the function $f(x,y) = 1 + y(x+1)$.

We exhibit a strong similarity between generic Sturmian words and substitutive Sturmian words, in the sense that any continuant function $\mathcal{Q}$ behaves in a similar way in each class. Our results are stated for a general $\mathcal{Q}$ function, that appears to be the convenient framework. However, we mainly apply our results to the recurrence quotient $T(\alpha,n)$ where the base function f is defined by $f(x,y) = 1 + y(x+1)$.

4.1 Case of general Sturmian words

We first consider a real α in the unit interval $[0,1]$, and the set

$$\mathcal{M}_\lambda(n) := \{\alpha \in \mathcal{I} \mid \Lambda(\alpha,n) \le \lambda\}\,. \tag{14}$$

When α is drawn with the uniform density, the measure of the set has been already studied (for $n \to \infty$) by the authors in [17]:

▶ **Theorem 7.** [Generic Sturmian words] *Consider a continuant function Λ associated with a function f and the domains $\mathcal{D}, \mathcal{D}_\lambda$ defined in* (12) *and* (13)*. The measure $M_\lambda(n)$ of the set $\mathcal{M}_\lambda(n)$ defined in* (14) *satisfies*[2]

$$M_\lambda(n) = R(\lambda)\,\left[1 + O(n^{-1})\right], \qquad n \to \infty$$

It involves the integral $R(\lambda)$ of the density ϖ on the domain D_λ, defined in (13)*. The constants hidden in the O-term are uniform for any λ with $\lambda \le \lambda_0$, and only depends on λ_0.*

[2] The remainder of the form $O(n^{-1})$ is only obtained for a class of base functions.

4.2 Case of substitutive Sturmian words

We now focus on the case when the slope α is a sqi number, i.e., a quadratic irrational number that satisfies the conditions described in Lemma 6. We are interested in the distribution of the function $\alpha \mapsto \Lambda(\alpha, n)$ restricted to $\alpha \in \mathcal{S}$, where $\mathcal{S}$ is the set of sqi numbers.
One has first to define our probabilistic model. Now, the slope α lives in an infinite discrete set $\mathcal{S}$ which is endowed the size ϵ defined in (10), and we will deal with the finite subsets $\mathcal{S}_D$

$$\mathcal{S}_D := \{\alpha \in \mathcal{S} \mid \epsilon(\alpha) \leq D\} \qquad D \to \infty\,. \tag{15}$$

We are interested in the asymptotic behaviour of the events $[\![\Lambda(\alpha, n) \leq \lambda]\!]$ for $n \to \infty$. It is natural to relate the sizes of the integer n and the sqi α, and the ultimate periodicity of $\texttt{Cfe}(\alpha)$ gives rise to the notion of *tour*, that we now describe more formally. With a pair formed with a sqi number of slope α and an integer n, we associate, as usual, the LFT g for which n belongs to the interval $[q(b(g), q(g)[$, but we only focus on pairs (α, n) which lead to LFT g that satisfy $ah^\ell \preceq g \prec ah^{\ell+1}$ (where a is the LFT associated with the preperiod, and h the primitive LFT associated with the period). The interval $Q_\ell(\alpha) := [q(a\,h^\ell), q(ah^{\ell+1})[$ is called the ℓ-th tour of α. The integers n of $Q_\ell(\alpha)$ satisfy $n \geq q[ah^\ell]$, and we focus on those integers n (at the "end" of the ℓ-th tour) for which $u := n/q[ah^\ell] \to \infty$. We are led to the set

$$\mathcal{M}_\lambda(u, \ell, D) := \left\{(\alpha, n) \mid \alpha \in \mathcal{S}_D,\, n \in Q_\alpha(\ell),\ n = q[ah^\ell]u,\ \ \Lambda\,(\alpha, n) \leq \lambda\right\},$$

and wish to study the behaviour of its cardinality $||\mathcal{M}_\lambda(u, \ell, D)||$ when u, ℓ, D tend to ∞ for a given λ. As later explained in Section 5.2, we consider instead an averaged version of the set, with an extra parameter $\mu > 1$,

$$\mathcal{M}_{\lambda,\mu}(u, \ell, D) := \left\{(\alpha, n) \mid \alpha \in \mathcal{S}_D,\, n \in Q_\alpha(\ell), \frac{n}{q[ah^\ell]} \in [u, \mu u],\, \Lambda\,(\alpha, n) \leq \lambda\right\}.$$

Since the length of the interval $[u, \mu u]$ is $(\mu - 1)u$, we are led to the normalized cardinality

$$M_{\lambda,\mu}(u, \ell, D) := \left(\frac{1}{(\mu - 1)u}\right) ||\mathcal{M}_{\lambda,\mu}(u, \ell, D)||\,. \tag{16}$$

In the following result, the occurrence of the factor $\log D$ is completely natural, since this is the mean value on $\mathcal{S}_D$ of the period of a sqi number –equal to the length of a *tour*–. This shows that the distribution function of the continuant function restricted to sqi numbers behaves in the same way as the distribution function in the real model.

▶ **Theorem 8.** [Substitutive Sturmian words] *Consider a continuant function Λ associated with a function f, the domains $\mathcal{D}, \mathcal{D}_\lambda$ defined in* (12),(13)*, and the set $\mathcal{S}_D$ defined in* (15)*. The following estimate holds for the normalized cardinality $M_{\lambda,\mu}(u, \ell, D)$ defined in* (16)*,*

$$\frac{1}{|\mathcal{S}_D|} M_{\lambda,\mu}(u, \ell, D) = R(\lambda)\,(\log D)\,[1 + O(\log D)^{-1}]\,[1 + O(u^{-a})]\,[1 + O(\theta^\ell)]\,,$$

where $R(\lambda)$ is the integral defined in (13)*, a is some strictly positive real, θ is a real with $\theta \in]0, 1[$ and the constants that are present in the remainder terms are uniform for any pair (λ, μ) with $\lambda \leq \lambda_0$, and $\mu \geq \mu_0 > 1$.*

5 Structure of the proof.

After a change of variables $u \mapsto 1/u$, the cost of interest is written as

$$C_\ell(\alpha, u, \lambda) := \sum_{g \in Q_\ell(\alpha)} A_g\left(0, \frac{q[ah^\ell]}{u}, \lambda\right) \qquad (u \to 0)\,.$$

It involves the function A_g defined in (8) and the pair (a, h) that defines the sqi α, where a is the *preperiod*, and h is the *primitive* LFT h associated with the period. We will show[3] that we can *forget* both the *primitivity* of h and also (at least at the beginning) the preperiod[4] a ; we are then led to a cost defined for any $h \in \mathcal{H}^+$, as

$$C_\ell(h, u, \lambda) := \sum_{g \in Q_\ell(h)} A_g\left(0, \frac{q[h^\ell]}{u}, \lambda\right) \quad (u \to 0)\,. \tag{17}$$

5.1 Algebraic study

We use three main tools:

A first tool: generating functions. We follow principles from Analytic Combinatorics. The set $\mathcal{H}^+$ is endowed with a size ϵ defined with $\epsilon(h) := |h'(h^\star)|^{-1/2}$, and $\mathcal{H}^+_D$ is the finite subset $\{h \in \mathcal{H}^+ \mid \epsilon(h) \le D\}$. For a cost C defined on $\mathcal{H}^+$, the asymptotic mean value $\mathbb{E}_D[C]$ on $\mathcal{H}^+_D$ (for $D \to \infty$) is given by the ratio, whose denominator and numerator are

$$\sum_{\substack{h \in \mathcal{H}^+ \\ \epsilon(h) \le D}} 1, \qquad \sum_{\substack{h \in \mathcal{H}^+ \\ \epsilon(h) \le D}} C(h) \qquad (D \to \infty)\,. \tag{18}$$

We are then led to asymptotic estimates of the previous sums, that are obtained via their generating functions (here of Dirichlet type, DGF in shorthand),

$$T(s) := \sum_{h \in \mathcal{H}^+} \epsilon(h)^{-s}, \qquad T_C(s) := \sum_{h \in \mathcal{H}^+} \epsilon(h)^{-s} C(h)\,.$$

However, the cost of interest in (17) does not depend only on the LFT h; it depends also on the parameter u and the index ℓ of the tour. As we are interested in the mean values on the set $\mathcal{H}^+_D$ when D, ℓ tend to ∞, $u \to 0$ and $\lambda \le \lambda_0$, we are led to

$$G_{D,\ell,\lambda}(u) := \mathbb{E}_D[C_\ell(h, u, \lambda)], \qquad D, \ell \to \infty,\ u \to 0\,. \tag{19}$$

We forget (for the moment) the dependence of G on λ and study $G_{D,\ell}$ via its Mellin transform.

A second tool: Mellin transforms. The Mellin transform (see [12]) associates with a function $G : [0, +\infty[\to \mathbb{C}$ the function, here denoted $\langle G \rangle_\rho$, defined by the integral

$$\langle G \rangle_\rho := \int_0^\infty u^{\rho-1} G(u) du\,.$$

The present context introduces the Mellin transform of the function $G_{D,\ell}$ defined in (19),

$$\langle G_{D,\ell} \rangle_\rho = \langle \mathbb{E}_D[C_\ell] \rangle_\rho = \mathbb{E}_D[\langle C_\ell \rangle_\rho] = \sum_{h \in \mathcal{H}^+} \mathbb{E}_D[\langle C_\ell(h) \rangle_\rho]\,.$$

As the sequence of functions $h \mapsto \langle C_\ell(h) \rangle_\rho$ *will be proven to have a limit* for $\ell \to \infty$, denoted as $h \mapsto F_h(\rho)$, we are led to the generating functions

$$S_{(\ell)}(s, \rho) := \sum_{h \in \mathcal{H}^+} \epsilon(h)^{-s} \langle C_\ell(h) \rangle_\rho\,, \qquad S(s, \rho) := \sum_{h \in \mathcal{H}^+} \epsilon(h)^{-s} F_h(\rho)\,.$$

[3] This will be proven via a precise comparison of their generating functions, that will be shown be *analytically equivalent*, as introduced in Definition 18

[4] We first perform the analysis without the preperiod and then show that we can *plug* the constraint via the preperiod operator $\mathbf{P}$ defined in (21)

As moreover these two series are analytically equivalent[5] when $\ell \to \infty$, the series $S(s,\rho)$ will be our main tool to estimate the expectation $\mathbb{E}_D[\langle C_\ell \rangle_\rho]$. Finally, we do not study the (initial) DGF of the cost C_ℓ, but, instead, the DGF associated to the Mellin transform of the (averaged)[6] version of the cost C_ℓ for an index ℓ *equal to* $\ell = \infty$. Transfer operators, together with their traces, will provide an alternative form for the main series $S(s,\rho)$.

A third tool: transfer operators. As in [2, 18, 9], the role of these operators is crucial. With each LFT $g \in \mathcal{H}^+$ defined in (6), we associate an operator $\mathbf{H}_{[g],s}$ and we deal with

$$\mathbf{H}_{[g],s}[f](x) := |g'(x)|^s f \circ g(x)\,, \quad \mathbf{H}_s := \sum_{h \in \mathcal{H}} \mathbf{H}_{[g],s}, \quad \mathbf{H}_s \circ (I - \mathbf{H}_s)^{-1} = \sum_{h \in \mathcal{H}^+} \mathbf{H}_{[g],s}\,. \tag{20}$$

These operators satisfy the key property $\mathbf{H}_{[g],s} \circ \mathbf{H}_{[h],s} = \mathbf{H}_{[h \circ g],s}$. There are also operators specific to the present framework: the *tour* operator $\mathbf{C}$ describes a *tour* and the *preperiod* operator $\mathbf{R}$ describes the constraint $\mathcal{R}$ relative to the preperiod defined in Lemma 6,

$$\mathbf{C}_{[h](c,d)} := \sum_{\substack{(u,g),\\ g \in \mathcal{H}^+\\ u \circ g = h}} \mathbf{H}_{[g],c} \circ \mathbf{H}_{[u],d}, \qquad \mathbf{R}_{(c,d)} := \sum_{(u,b) \in \mathcal{R}} \mathbf{H}_{[b],c} \circ \mathbf{H}_{[u],d} \quad \text{with } (c,d) \in \mathbb{C}^2\,. \tag{21}$$

As shown in the Annex, taking the index $\ell = \infty$ leads to the fixed point $h^\star$ of the LFT h, and fixed points themselves lead to the *trace* of operators. The DGF $S(s,\rho)$ involves the function $L_{\lambda,\rho}$ that is the ρ-Mellin transform of the function $x \mapsto L_\lambda(x,y)$, itself associated with the geometry of the $\mathcal{Q}$-function, together with the traces of tour operators $\mathbf{C}$. Finally, the DGF $S(s,\rho)$ is proven to be analytically equivalent to another DGF $T(s,\rho)$ which is expressed with the pair $(a,b) := (s/2, (s+\rho)/2)$, as

$$T(s,\rho) = \sum_{h \in \mathcal{H}^+} \operatorname{Tr}\left[L_{\lambda,\rho} \cdot \mathbf{C}_{[h],(a,b)}\right] = \operatorname{Tr}\left[L_{\lambda,\rho} \cdot (I - \mathbf{H}_a)^{-1} \circ \mathbf{H}_a \circ (I - \mathbf{H}_b)^{-1}\right]. \tag{22}$$

5.2 Analytic study

Analytic properties of the transfer operators. Eq. (22) shows that we need fine properties of the transfer operator $\mathbf{H}_s$, and notably of its quasi-inverse $(I - \mathbf{H}_s)^{-1}$. There are two useful spaces[7] on which the operator $\mathbf{H}_s$ acts, each of them playing a specific role: A space of analytic functions is needed for dealing with the trace of operators, whereas the space $\mathcal{C}^1([0,1])$ is well adapted to study the operator $\mathbf{H}_s$ for parameters s with large $|\Im s|$. When the operator $\mathbf{H}_s$ acts on each of the two spaces, it admits nice spectral properties, that are transfered to the quasi-inverse. First (on both spaces), the mapping $s \mapsto (I - \mathbf{H}_{s/2})^{-1}$ is meromorphic on the half-plane $\Re s > 2$, with a unique pole (simple) at $s = 2$, and satisfies

$$(I - \mathbf{H}_{s/2})^{-1}[F](t) \sim_{s \to 2} \frac{2}{\mathcal{E}} \frac{1}{s-2} \psi(t)\ I[F]\,, \qquad \mathcal{E} = \pi^2/(6 \log 2)\,, \tag{23}$$

where $I[F]$ is the integral of F on $[0,1]$, ψ is the Gauss density defined by the equality $\psi(x) = (1/\log 2) 1/(1+x)$ and $\mathcal{E}$ is the entropy of the dynamical system. This behaviour will be transfered to the trace of the quasi-inverse (see [9]), defined when $\mathbf{H}_s$ acts on the analytic space. Moreover, when $\mathbf{H}_s$ acts on $\mathcal{C}^1([0,1])$, deep results due to Dologopyat-Baladi-Vallée

[5] as introduced in Definition 18.
[6] We return to the need of averaging in Section 5.2.
[7] The study is parallel to this performed in [9]

[11] [2] prove that the quasi-inverse $(I - \mathbf{H}_{s/2})^{-1}$ well behaves on the left of $s = 2$: it admits a vertical strip free of poles with a polynomial growth on vertical lines.

Perron Formula. We consider the Dirichlet series $T_\rho : s \mapsto T(s, \rho)$ defined in (22) where ρ belongs to a ρ-vertical strip $|\Re\rho| \le \nu_0$ with ν_0 small enough. On a s-vertical strip $|\Re s - 2| < \sigma_0$, Eq. (22) shows that each $T_\rho(s)$ has two (simple) poles at $s = 2$ and $s = 2 - \rho$ that coalesce to a double pole for $\rho = 0$. With Dolgopyat bounds, the series $T_\rho(s)$ is of polynomial growth when $|\Im s|$ grows to ∞.
Then, the Landau theorem applies (see [14], [4], [6]) and provides asymptotic estimates (for $D \to \infty$) of the sums defined in (18), that involve the residues (A for $T(s)$ at $s = 2$ and two residues $a(\rho)$ et $b(\rho)$ for $s \mapsto T(s, \rho)$ at $s = 2$ and $s = 2 - \rho$). This leads to the final asymptotics (when $D \to \infty$) for the mean value $M_D(\rho) := \mathbb{E}_D[F(\rho)]$ where $F(\rho)$ is defined as the ρ-Mellin transform of the cost C_ℓ at $\ell = \infty$.

Inverse Mellin Transform. Need for averaging the cost. We need now to recover our initial cost and return from the asymptotics (for $D \to \infty$) of

$$M_D(\rho) := \mathbb{E}_D[F(\rho)] = \mathbb{E}_D[\lim_{\ell\to\infty} \langle C_\ell\rangle_\rho] = \lim_{\ell\to\infty} \langle \mathbb{E}_D[C_\ell]\rangle_\rho \ ,$$

to our initial study, where we are interested in the behaviour of the function $G_{D,\ell} := \mathbb{E}_D[C_\ell]$ defined in (19) when $D, \ell \to \infty$ and $u \to 0$. We then wish to use the inverse Mellin transform (see [12]) of the function $\rho \mapsto M_D(\rho)$.
The function $\rho \mapsto M_D(\rho)$ is meromorphic on the ρ-vertical strip $|\Re\rho| \le \nu_0$. It admits a main term –given by the residues $a(\rho)$ and $b(\rho)$, that coalesce at $\rho = 0$– which gives rise to a unique pole (simple) at $\rho = 0$, and a residue[8] $\Theta(\log D)$ at $\rho = 0$. There is also a remainder term uniform, both in D (for $D \to \infty$) and in ρ when ρ belongs to the domain $\{\rho \mid |\Re\rho| \le \nu_0, |\Im\rho| \ge b\}$.
For using the inverse Mellin transform, we need, on the previous domain, the (crucial) estimate $M_D(\rho) = O(|\Im\rho|^{-c})$ with $c > 1$; however, we *only obtain an exponent* $c \in]0, 1[$ *for our study, associated with the cost* $v \mapsto C(h, v)$. This is why, as already said in Section 4.2, we deal with the *averaged version* $M_\mu(C_\ell)$ of the cost C_ℓ whose ρ-Mellin transform satisfies

$$\langle M_\mu(C_\ell)\rangle_\rho = \left(\frac{1}{\rho+1}\right)\left(\frac{\mu^{\rho+1}-1}{\mu-1}\right)\langle C_\ell\rangle_\rho \,.$$

Then, the averaging brought a supplementary factor $O(|\Im\rho|^{-1})$ in the Mellin transforms and the inverse Mellin transform now applies: the behaviour of $\langle M(G_{D,\ell})\rangle_\rho$ on the left border on the half-plane provides a good knowledge of the behaviour of $u \mapsto M(G_{D,\ell})(u)$ near 0, and finally, remembering the parameter λ, we obtain for the averaged version the final estimate

$$M(G_{D,\ell,\lambda})(u) \sim R(\lambda)\, \log D[1 + O(\log D)^{-1}][1 + O(u^a)][1 + O(\theta^\ell)] \qquad u \to 0\,.$$

Here, $R(\lambda)$ occurs as a residue which coincides with the expression in (13), and all the constants are uniform for $\lambda \le \lambda_0$. We have then proven the main result for the cost $M_\mu(C_\ell)$. Remark that we can perform all our study with (the plain) cost C_ℓ, and consider the cost $M_\mu(C_\ell)$ only at the end.

Open problems. Is it possible to obtain our main result without averaging the cost?
In the previous *generic* study for generic Sturmian words [17], we deal with a parameter λ which may vary; this leads in Section 4 of [17] to results on the conditional expectation of the recurrence, depending on the position of n inside the reference interval Is it possible to perform the same type of study in the present framework ?

[8] The factor $\log D$ is due to the presence of a double pole at $s = 2$ for the series $S_0(s)$

References

1 Cyril Allauzen. Une caracterisation simple des nombres de Sturm. Journal de Théorie des Nombres de Bordeaux, 10(2):237–241, 1998.
2 Viviane Baladi and Brigitte Vallée, Euclidean Algorithms are Gaussian, Journal of Number Theory, Volume 110, Issue 2 (2005) pp 331–386
3 Valérie Berthé, Fréquences des facteurs des suites sturmiennes, Theoret. Comput. Sci., 165 (1996), pp. 295–309.
4 Valérie Berthé, Loick Lhote and Brigitte Vallée, Probabilistic analyses of the plain multiple GCD algorithm, Journal of Symbolic Computation 74 (2016) 425–474,
5 Valérie Berthé, Eda Cesaratto, Pablo Rotondo, Brigitte Vallée and A. Viola, Recurrence function of Sturmian Words: a probabilistic study, in Proceedings of the 40th MFCS, Springer, 2015, Part I, pp. 116–128
6 Eda Cesaratto, Julien Clément, Benoît Daireaux, Loïck Lhote, Véronique Maume-Deschamps and Brigitte Vallée, Regularity of the Euclid Algorithm: application to the analysis of fast gcd Algorithms, Journal of Symbolic Computation, 44 (2009) pp 726–767
7 Julien Cassaigne. Limit values of the recurrence quotient of Sturmian sequences. Theoretical Computer Science, 218(1):3–12, 1999.
8 Julien Cassaigne. Recurrence in infinite words. In Afonso Ferreira and Horst Reichel, editors, STACS 2001, pages 1–11, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
9 Eda Cesaratto and Brigitte Vallée. Gaussian behavior of quadratic irrationals, https://arxiv.org/pdf/1708.00051.pdf
10 D. Crisp, W. Moran, A. Pollington and P. Shiue, Substitution invariant cutting sequences. Journal de théorie des nombres de Bordeaux, Volume 5 (1993) no. 1, pp. 123–137.
11 Dimitri Dolgopyat, On decay of correlations in Anosov flows. Ann. of Math., 147(2):357–390, 1998.
12 Philippe Flajolet, Xavier Gourdon and Philippe Dumas, Mellin transforms and asymptotics: Harmonic sums. Theoretical Computer Science 144, 1- 2 (June 1995), 3-58.
13 N. Pytheas Fogg. Substitutions in dynamics, arithmetics and combinatorics, volume 1794 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.
14 Edmund Landau, Über die Anzahl der Gitterpunkte in Gewissen Bereichen. IV. Nachr. Ges. Wiss. Gött., Math.-Phys. Kl. :137–150, 1924.
15 M. Lothaire. Algebraic combinatorics on words, volume 90 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2002.
16 M. Morse and G. Hedlund, Symbolics dynamics II. Sturmian trajectories, American Journal of Mathematics, 62 (1940), pp. 1–42.
17 Pablo Rotondo and Brigitte Vallée. The recurrence function of a random Sturmian word. In ANALCO'2017 Proceedings, pages 101–114. SIAM, Philadelphia, PA, 2017.
18 Brigitte Vallée, Dynamical sources in Information Theory: Fundamental Intervals and Word prefixes, Algorithmica (2001), vol 29 (1/2) pp 262–306
19 Brigitte Vallée, Dynamique des fractions continues à contraintes periodiques. Journal of Number Theory, vol 72, pp. 183–235, 1998.
20 Brigitte Vallée. Euclidean Dynamics. Discrete and Continuous Dynamical Systems, vol. 15, no. 1, pp. 281–352, 2006.