Regularity of the Euclid Algorithm, Application to the analysis of fast GCD Algorithms

Eda Cesaratto, Julien Clément, Benoît Daireaux, Loïck Lhote, Véronique Maume-Deschamps, Brigitte Vallée

REGULARITY OF THE EUCLID ALGORITHM. APPLICATION TO THE ANALYSIS OF FAST GCD ALGORITHMS.

EDA CESARATTO, JULIEN CLÉMENT, BENOÎT DAIREAUX, LOÏCK LHOTE, VÉRONIQUE MAUME-DESCHAMPS, AND BRIGITTE VALLÉE

Abstract. There exist fast variants of the gcd algorithm which are all based on principles due to Knuth and Schönhage. On inputs of size n, these algorithms use a Divide and Conquer approach, perform FFT multiplications with complexity $\mu(n)$ and stop the recursion at a depth slightly smaller than $\lg n$. A rough estimate of the worst–case complexity of these fast versions provides the bound $O(\mu(n)\log n)$. Even the worst-case estimate is partly based on heuristics and is not actually proven. Here, we provide a precise probabilistic analysis of some of these fast variants, and we prove that their average bit–complexity on random inputs of size n is $\Theta(\mu(n)\log n$, with a precise remainder term, and estimates of the constant in the Θ–term. Our analysis applies to any cases when the cost $\mu(n)$ is of order $\Omega(n\log n)$, and is valid both for the FFT multiplication algorithm of Schönhage–Stassen, but also for the new algorithm introduced quite recently by Fürer [12]. We view such a fast algorithm as a sequence of what we call interrupted algorithms, and we obtain two main results about the (plain) Euclid Algorithm which are of independent interest. We precisely describe the evolution of the distribution of numbers during the execution of the (plain) Euclid Algorithm, and we exhibit an (unexpected) density ψ which plays a central rôle since it always appear at the beginning of each recursive call. This strong regularity phenomenon proves that the interrupted algorithms are locally "similar" to the total algorithm. This finally leads to the precise evaluation of the average bit–complexity of these fast algorithms. This work uses various tools, and is based on a precise study of generalised transfer operators related to the dynamical system underlying the Euclid Algorithm.

1. Introduction

Gcd computation is a widely used routine in computations on long integers. It is omnipresent in rational computations, public key cryptography or computer algebra. Many gcd algorithms have been designed since Euclid. Most of them compute a sequence of remainders by successive divisions, which leads to algorithms with a quadratic bit–complexity (in the worst-case as well as in the average-case). Using Lehmer's ideas [20] (which replace large divisions by large multiplications and small divisions), computations can be speeded-up by a constant factor, but the asymptotic complexity remains quadratic. Major improvements in this area are due to Knuth [19], who designed the first subquadratic algorithm in 1970, and to Schönhage [24] who subsequently improved it the same year. They use Divide and Conquer techniques combined with Lehmer's ideas to compute in a recursive way the quotient sequence (whose total size is $O(n)$). Moreover, if a fast multiplication with subquadratic complexity (FFT, Karatsuba...) is performed, then one obtains a subquadratic gcd algorithm (in the worst-case). Such a methodology has been recently used by Stehlé and Zimmermann [25] to design a Least-Significant-Bit version of the Knuth-Schönhage algorithm. According to experiments due to [5] and [22], these algorithms (with an FFT multiplication) become efficient only for integers of size larger than 10000 words, whereas, with Karatsuba multiplication, they become efficient for smaller integers (around 100 words). A precise description of the Knuth-Schönhage algorithm can be found in [29, 22] for instance.

1.1. Previous results.

The average-case behaviour of the quadratic gcd algorithms is now well understood. First results are due to Heilbronn and Dixon in the seventies, who studied for the first time the mean number of iterations of the Euclid Algorithm. Then Brent analysed the Binary algorithm [4], and Hensley [14] provided the first distributional analysis for the number of steps of the Euclid Algorithm. Since 1995, the Caen Group [26, 28, 27] and its collaborators have performed an average-case analysis of various parameters of a large class of Euclidean algorithms. More recently, distributional results have also been obtained for the Euclid algorithm and some of

Date: 24 november 2007.

its variants: first Baladi and Vallée prove that a whole class of so–called additive costs of moderate growth follows an asymptotic gaussian law [2] (for instance, the number of iterations, the number of occurrences of a given digit, and so on...). In 2006, Lhote and Vallée [21] showed that a more general class of parameters also follows an asymptotic gaussian law. This class contains the length of a remainder at a fraction of the execution, and the bit-complexity. To the best of our knowledge, there are yet few results on "efficient" gcd algorithms. In [7], the authors perform an average-case analysis of Lehmer's algorithm, and exhibit the average speed-up obtained using these techniques. However, as far as we know, there does not exist any probabilistic analysis of subquadratic gcd algorithms. It is the goal of this paper to perform such a study.

1.2. **Our results.** There are two algorithms to be analyzed: the $\mathcal{HG}$ algorithm and the $\mathcal{G}$ algorithm. The $\mathcal{G}$ algorithm computes the gcd, and the $\mathcal{HG}$ algorithm (for "half-gcd" Algorithm) only simulates the "first half" of the $\mathcal{G}$ algorithm. We first show that these algorithms can be viewed as a sequence of the so–called Interrupted Euclidean algorithms. An Interrupted Euclidean algorithm is a subsequence formed by successive iterations of the plain algorithm, as we now explain: On an input (A, B), the plain Euclid algorithm builds a sequence of remainders A_i, a sequence of quotients Q_i, and a sequence of matrices $\mathcal{M}_i$ [see Section 2.1]. On an input (A, B) of binary size n, the Interrupted Euclidean algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}$ starts at the index k of the execution of the Euclid Algorithm, as soon as the remainder A_k has already lost $\delta\, n$ bits (with respect to the initial A which has n bits) and stops at index $k+i$ as soon as the remainder A_{k+i} has lost $\gamma\, n$ additional bits (with respect to the remainder A_k). The $\mathcal{HG}$ algorithm just simulates the interrupted algorithm $\mathcal{E}_{[0,1/2]}$. A quite natural question is: How many iterations are necessary to lose these $\gamma\, n$ bits? Of course, it is natural to expect that this subsequence of the Euclidean algorithm is just locally similar to the "total" Euclidean Algorithm; in this case, the number of iterations would be close to $\gamma\, P$ (where P is the number of iterations of the "total" Euclid algorithm). We prove in Theorem 1 that this is indeed the case: This is why we say that the algorithm is "regular".

For a probabilistic study of fast variants, a precise description of the evolution of the distribution during the execution of the plain Euclid Algorithm is of crucial interest. For real inputs, we know that the continued fraction algorithm does not terminate (except for rationals ...). Moreover, as the continued fraction algorithm is executed, the distribution of reals tends to the distribution associated to the Gauss density φ, defined as

$$\varphi(x) = \frac{1}{\log 2}\frac{1}{1+x}. \tag{1}$$

For rational inputs, we begin with a given distribution on the set of the inputs $x := A_1/A_0$ of size n, and we consider the rationals $x_k := A_{k+1}/A_k$. We focus on the first index k where the binary size of x_k is less than $(1-\delta)n$ and we denote the corresponding rational x_k by $x_{\langle\delta\rangle}$. What is the distribution of the rational $x_{\langle\delta\rangle}$? The evolution of this distribution is clearly more intricate than in the real case, since at the end of the Algorithm (when $\delta = 1$), the distribution is the Dirac measure at $x = 0$. We obtain here a precise description of this distribution (see Theorem 2 and Figure 1) which surprisingly involves the density function

$$\psi(x) := \frac{12}{\pi^2}\sum_{m\geq 1}\frac{\log(m+x)}{(m+x)(m+x+1)}. \tag{2}$$

We also need precise results on the distribution of some truncations of remainders. This is done in Theorem 3. Then, the choice of parameters in the fast algorithms must take into account this evolution of distribution. This is why we are led to introduce some variants of the classical algorithms, denoted by $\underline{\mathcal{HG}}$ and $\underline{\mathcal{G}}$ for which the precise analysis can be performed.

The fast versions also involve other functions, which are called the Adjust functions. Such functions perform a few steps of the (plain) Euclid Algorithm. However, the bit-complexity of the Adjust functions depends on the size of the quotients which are computed during these steps. Even for estimating the worst–case complexity of the fast variants, the Adjust functions are not precisely analyzed. The usual argument is "The size of a quotient is $O(1)$". Of course, this assertion is false in the worst-case, and only true on average, provided that the distribution on input pairs be made precise. Moreover, the Adjust functions are related to some specific steps, which happen just when the pairs have lost a fraction of their bits. We are then led to study the mean value of the size of

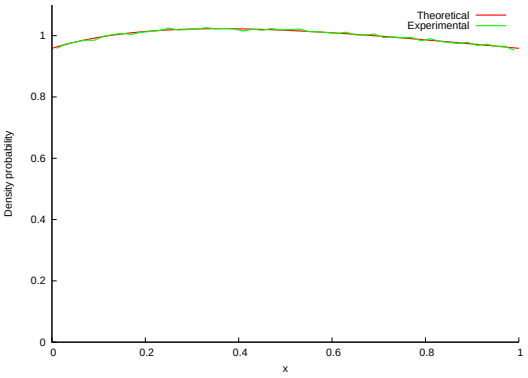

FIGURE 1. Density distribution of $x_{\langle\delta\rangle}$ in the case $\delta = 1/2$, corresponding to the density distribution of the rational $x_k := A_{k+1}/A_k$ obtained as soon as $\ell(A_k)$ is smaller than $(1/2)\ell(A_0)$. The diagram compares Monte-Carlo simulations to the exact value of $\psi(x)$. For simulations, we consider 3 537 944 rationals with 48 bits, drawn according to the Gauss density φ. For estimating the density, the interval $[0, 1]$ is subdivided into equal subintervals of length $1/50$.

the quotients computed at these specific steps, and we prove that it is asymptotic to a constant L which is defined in (19). And, we also need this type of result for our truncated data. This is covered by Theorem 4.

There are now two main fast multiplication algorithms, both based on FFT principles. We consider in fact a whole class of possible fast multiplication algorithms, for which the following is true:
There exist a function $a(n)$ satisfying[1] $a(n) = O(\log\log n)$, $a(n) = \Omega(1)$ *and two constants A_1, A_2 (probably large) such that, for any pair of integers u, v whose respective sizes satisfy $\ell(u) = n$ and $\ell(v) = Kn$ for some integer K, the bit–cost $M(u, v)$ of the product between two numbers u and v satisfies*

$$A_1\, K\, \mu(n) \le M(u,v) \le A_2\, K\, \mu(n) \qquad \textit{with} \qquad \mu(n) = n \log n\, a(n). \tag{3}$$

In particular, Fürer proved this year [12] that it is possible to choose $a(n) = 2^{O(\log^\star n)}$, and improves the previous function $a(n) = \log\log n$, due to Schönhage and Strassen.

Such a fast multiplication also leads to a fast division:
There exist two constants A_3, A_4 (larger than A_1, A_2) such that, for any pair of integers u, v whose respective sizes satisfy $\ell(u) = n$ and $\ell(v) = Kn$ for some integer $K > 1$, the bit–cost $D(u, v)$ of the division between two numbers v and u satisfies[2]

$$A_3\, (K-1)\, \mu(n) \le D(v,u) \le A_4\, (K-1)\, \mu(n) \qquad \textit{with} \qquad \mu(n) = n \log n\, a(n). \tag{4}$$

Finally, we obtain the exact average-case complexity of our versions of the two main algorithms of interest, the $\underline{\mathcal{HG}}$ algorithm, and the $\underline{\mathcal{G}}$ algorithm itself. When they use a fast multiplication which satisfies (3), we prove the following estimates [Theorems 6 and 7] for the average bit-complexity B, G of both algorithms, on the set of random inputs of size n:

$$\mathbb{E}_n[B] = \Theta(1)\, n \log^2 n\, a(n) \left[1 + O\left(\frac{1}{a(n)}\right)\right], \quad \mathbb{E}_n[G] = \Theta(1)\, n \log^2 n\, a(n) \left[1 + O\left(\frac{1}{a(\sqrt{n}\log n)}\right)\right].$$

Furthermore, we obtain precise information about the Θ–term, which involves two types of constants : first, the constants A_1, A_2, which intervene in the cost of the mutiplication [see (3)], second, together with the density ψ defined in (2), another mysterious "spectral" constant σ (defined in Section 1.3). Our proven average bit–complexity of the $\underline{\mathcal{HG}}, \underline{\mathcal{G}}$ algorithms then appears to be of the same order as the usual (heuristic) bound on the worst-case complexity of $\mathcal{HG}, \mathcal{G}$ algorithms.

[1] the notation $f = \Omega(g)$ means that there exists $B > 0$ such that, for n large enough, $f_n \ge B g_n$

[2] In this case $(K-1)n$ is the size of the quotient

1.3. **Methods.** All our main conclusions obtained here are "expected", and certainly will not surprise the reader. However, the irruption of the density ψ is unexpected, and an actual proof of this phenomenon is not straightforward. This is due to the fact that there are correlations between successive steps of the Euclid Algorithm. Accordingly, the tools which are usual in analysis of algorithms [11], like generating functions, are not well–suited in this case. All the analyses which will be described here are instances of the dynamical analysis paradigm, where one proceeds in three main steps: First, the (discrete) algorithm is extended into a continuous process, which can be defined in terms of the dynamical system related to the Gauss map. Then, the transfer operator $\mathbf{H}_s$ defined as

$$\mathbf{H}_s[f](x) := \sum_{m\geq 1} \frac{1}{(m+x)^{2s}} f\left(\frac{1}{m+x}\right)$$

serves to describe how the distribution evolves, in the continuous world. Finally, the executions of the gcd algorithm are now described by particular trajectories (i.e., trajectories of "rational" points), and a transfer "from the continuous to the discrete" must be performed, using Dirichlet series.

The present paper mainly uses two previous works, and can be viewed as an extension of them: first, the average-case analysis of the Lehmer-Euclid algorithm performed in [7]; second, the distributional methods described in [2, 21]. First, we again use the general framework that Daireaux and Vallée have developed for the analysis of the Lehmer-Euclid Algorithm, which explains how the Lehmer-Euclid algorithm can be viewed as a sequence of Interrupted Euclidean algorithms $\mathcal{E}_{[\delta,\delta+\gamma]}$. Whereas some "easy" properties of the transfer operator $\mathbf{H}_s$ were used in [7], we here need properties which were already crucial in previous distributional analysis [2, 1, 21] –namely, the *US* Property for the quasi-inverse $(I-\mathbf{H}_s)^{-1}$ of the transfer operator–. The *US*(α) Property can be summarized in an informal way as follows:

Property *US*(α). *When $\mathbf{H}_s$ acts on the functional space $\mathcal{C}^1(\mathcal{I})$ of functions with a continuous derivative on the unit interval $\mathcal{I} := [0,1]$, the following holds on the strip $\mathcal{S} := \{s, 1-\alpha \leq \Re s \leq 1\}$*

(*i*) *The quasi-inverse $(I-\mathbf{H}_s)^{-1}$ has a unique pôle located at $s=1$.*
(*ii*) *It is of polynomial growth with respect to $|\Im s|$ for s large enough.*

The main result of Dolgopyat, made more precise by Baladi and Vallée, proves that there exists an $\alpha > 0$ for which Property *US*(α) holds. The arguments which show the existence of such a strip are not all constructive, and we do not know any explicit strictly positive lower bound on α. In the paper, such a lower bound is denoted by σ, and the parameter $\underline{\sigma} := \min(\sigma, 1/2)$ plays a central rôle in our analyses: This is the mysterious constant which intervenes in the constants of our two main Theorems. It intervenes also in all the (exponential) remainder terms [see Theorems 1, 2, 3, 4, 5].

In order to establish our main results, we are led to studying parameters of various type, whose generating functions involve operators $\mathbb{G}_{s,t}$ which depend on two variables s,t. However, for small t's, all these operators can be viewed as a perturbation of the quasi-inverse $(I-\mathbf{H}_s)^{-1}$ and the *US* Property extends to these perturbed quasi-inverses. In particular, the existence of a strip $\mathcal{S}$ where the *US* property holds uniformly with respect to t is crucial in the analysis.

Plan and notations. Section 2 describes the main algorithms $\mathcal{HG}$ and $\mathcal{G}$. Section 3 presents the main steps towards a proven analysis. Then, we state our main results of general interest, without proofs. In Section 4, we describe the versions $\underline{\mathcal{HG}}$ and $\underline{\mathcal{G}}$ to be analyzed, and, with the results (yet unproved of Section 3), we show the two main results about their average bit-complexity. Section 5 describes the general framework of the Dynamic Analysis paradigm, and Section 6 is devoted to the proof of the main results stated in Section 3. Some technical results are gathered in an appendix (Section 7).
We denote the logarithm in base 2 by $\lg x$, and $\ell(x)$ denotes the binary size of integer x, namely $\ell(x) := \lfloor \lg x \rfloor + 1$.

2. Fast and Interrupted Euclidean algorithms

We present in this section the main algorithms studied in this paper. We first describe the general structure of the Knuth-Schönhage algorithm. We explain how the $\mathcal{HG}$ algorithm can be

seen as a sequence of interrupted Euclidean algorithms, where the sequence of divisions is stopped as soon as the integers have lost a fraction of their number of bits.

2.1. **Euclid's algorithm.** Let (A_1, A_0) be a pair of positive integers with $A_1 \leq A_0$. On input (A_1, A_0), the Euclid algorithm computes the remainder sequence (A_k) with a succession of divisions of the form

$$A_k = Q_{k+1}A_{k+1} + A_{k+2}, \quad \text{with} \quad Q_{k+1} = \left\lfloor \frac{A_k}{A_{k+1}} \right\rfloor, \tag{5}$$

and stops when $A_{p+1} = 0$. The integer Q_k is the k–th quotient and the successive divisions are written as

$$\mathcal{A}_k = \mathcal{Q}_{k+1}\mathcal{A}_{k+1}, \quad \text{with} \quad \mathcal{A}_k := \begin{pmatrix} A_{k+1} \\ A_k \end{pmatrix} \quad \text{and} \quad \mathcal{Q}_k := \begin{pmatrix} 0 & 1 \\ 1 & Q_k \end{pmatrix},$$

so that

$$\mathcal{A}_0 = \mathcal{M}_{(i)}\mathcal{A}_i \ \text{ with } \quad \mathcal{M}_{(i)} := \mathcal{Q}_1\mathcal{Q}_2\cdots\mathcal{Q}_i. \tag{6}$$

In the following, we consider a part of the plain Euclidean Algorithm $\mathcal{E}$, (which is sometimes called a "slice") between index i and index j, namely the interrupted algorithm $\mathcal{E}_{(i,j)}$ which begins with the pair $\mathcal{A}_i$ as its input and computes the sequence of divisions (5) with $i \leq k \leq j-1$. Its output is the pair $\mathcal{A}_j$ together with the matrix

$$\mathcal{M}_{(i,j)} = \prod_{k=i+1}^{j} \mathcal{Q}_k, \qquad \mathcal{M}_{(1,i)} = \mathcal{M}_{(i)}, \tag{7}$$

with matrix $\mathcal{M}_{(i)}$ defined in (6). We define the size of a matrix $\mathcal{M}$ as the maximum of the binary sizes of its coefficients. The size $\ell_{(i,j)}$ of the matrix $\mathcal{M}_{(i,j)}$ satisfies

$$\ell_{(i,j)} \leq 2(j-i) + \sum_{k=i+1}^{j} \ell(Q_k) \tag{8}$$

The (naive) bit-complexity $C_{(i,j)}$ of the algorithm $\mathcal{E}_{(i,j)}$ satisfies

$$C_{(i,j)} := \sum_{k=i+1}^{j} \ell(A_k)\cdot\ell(Q_k) \leq \ell(A_{i+1})\cdot\sum_{k=i+1}^{j}\ell(Q_k). \tag{9}$$

The Lehmer Algorithm [20, 18] replaces large divisions by large multiplications and small divisions. The fast algorithm applies recursively the principles of Lehmer, and using fast FFT multiplications of complexity $\Theta(\mu(n))$ (with $\mu(n) = n\log n\log\log n$) replaces the costly computation of the remainder sequence A_i (which requires $O(n^2)$ bit operations), by a sequence of matrix products: it divides the total Euclidean Algorithm into interrupted Euclidean algorithms, of the form $\mathcal{E}_{(i,j)}$ and computes matrices of the form $\mathcal{M}_{(i,j)}$, defined in (7). The recursion, based on Divide and Conquer techniques, is stopped when the integers are small enough, and, at this moment, the algorithm uses small divisions. One finally obtains a subquadratic gcd algorithm.

2.2. **How to replace large divisions by small divisions?** Lehmer remarked that, when two pairs (A, B) and (a, b) are sufficiently close (i.e., the rationals A/B and a/b are close enough), the Euclid algorithm on (A, B) or (a, b) produces (at least at the beginning) the same quotient sequence (Q_i). This is why the following definition is introduced:

Definition. *Consider a pair (A, B) with $A \leq B$ and an integer b of length $\ell(b) \leq \ell(B)$. We denote by $\pi_{[b]}(A, B)$ any pair (a, b) which satisfies*

$$\left|\frac{A}{B} - \frac{a}{b}\right| \leq \frac{1}{b}.$$

And the criterion (due to Lehmer and made precise by Jebelean) is:

Lemma 1. [Lehmer, Jebelean] *For a pair (A, B) with $A \leq B$ and $n := \ell(B)$, consider, for $m \leq n$, the small pair $(a, b) = \pi_{[b]}(A, B)$ of length $\ell(b) = m$, and the sequence of the remainders (a_i) of the Euclid Algorithm on the small input (a, b). Denote by k the first integer k for which a_k satisfies $\ell(a_k) \leq \lceil m/2 \rceil$. Then the sequence of the quotients q_i of the Euclid Algorithm on the small input*

(a, b) coincides with the sequence of the quotients Q_i of the Euclid Algorithm on the large input (A, B) for $i \leq k - 3$.

Usually, this criterion is used with a particular pair $\pi_{[b]}(A, B)$ where the integer b is obtained by the m-truncation of B, i.e., the suppression of its $(n - m)$ least significant bits. Then a is easy to compute since it may be chosen itself as the m-truncation of A. In this case, the $\pi_{[b]}$ function corresponds to truncation of both A and B and is denoted by $T_m(A, B)$. However, the Jebelean criterion holds for any choice of $(a, b) = \pi_{[b]}(A, B)$, even if the integer a is less easy to compute in the general case: the integer a can be chosen as the integer part of the rational $(Ab)/B$, and its computation needs a product and a division.

2.3. **Interrupted Algorithms.** In Jebelean's property (Lemma 1), the Euclid Algorithm on the small pair (a, b) of binary size m is stopped as soon the remainder a_k has lost $\lceil m/2 \rceil$ bits. This is a particular case of the so–called Interrupted Euclidean Algorithm of parameter δ (with $0 < \delta < 1$), which stops as soon as the current remainder has lost δm bits (with respect to the input which has m bits). This (general) interrupted Algorithm denoted by $\mathcal{E}_\delta$, and described in Figure 2, is defined as follows: On the input (A, B) of size n, this algorithm begins at the beginning of the Euclid Algorithm, and stops as soon as the remainder A_i has lost $\delta\, n$ bits (with respect to the input B). Then, with the notations defined in Section 2.1, one has $\mathcal{E}_\delta = \mathcal{E}_{(1,P_\delta)}$, with

$$P_\delta := \min\big\{k;\ \lg A_k \leq (1 - \delta)n\big\}. \tag{10}$$

Figure 2 also describes the $\widehat{\mathcal{E}}_\delta$ Algorithm, which is just a slight modification of the $\mathcal{E}_\delta$ Algorithm, where the last three steps are suppressed (in view of applications of Lemma 1), and $\widehat{P}_\delta$ denotes the variable $P_\delta - 3$. Then, P_δ, and $\widehat{P}_\delta$ are just the number of iterations of the $\mathcal{E}_\delta, \widehat{\mathcal{E}}_\delta$ algorithms and $P_1 = P$ is just the number of iterations of the Euclid Algorithm.

In the following, it will be convenient to consider more general interrupted algorithms, of the form $\mathcal{E}_{[\delta,\delta+\gamma]}$. The Algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}$ is defined as follows: On the input (A, B) of size n, this algorithm begins at the P_δ-th iteration of the Euclid Algorithm, as soon as the remainder A_k has lost $\delta\, n$ bits (with respect to the input B) and stops when the remainder A_i has lost $\gamma\, n$ additional bits (with respect to the input B). Then, $\mathcal{E}_{[0,\delta]} = \mathcal{E}_\delta = \mathcal{E}_{(0,P_\delta)}$ and $\mathcal{E}_{[\delta,\gamma+\delta]} = \mathcal{E}_{(P_\delta,P_{\delta+\gamma})}$, where P_δ is defined in (10). Of course, we can also design the variants with a hat, where the last three steps are suppressed.

```
Algorithm E_δ(A, B)
  n := ℓ(B)
  i := 1
  A_1 := A, A_0 := B
  M_0 := I
  While lg A_i > (1 − δ) · n
    Q_i := ⌊A_{i−1}/A_i⌋
    A_{i+1} := A_{i−1} − Q_i A_i
    M_i := M_{i−1} · Q_i
    i := i + 1
  Return (A_{i−1}, A_i, M_{i−1})
```

```
Algorithm Ê_δ(A, B)
  n := ℓ(B)
  i := 1
  A_1 := A, A_0 := B
  M_0 := I
  While lg A_i > (1 − δ) · n
    Q_i := ⌊A_{i−1}/A_i⌋
    A_{i+1} := A_{i−1} − Q_i A_i
    M_i := M_{i−1} · Q_i
    i := i + 1
  Return (A_{i−3}, A_{i−2}, M_{i−3})
```

FIGURE 2. The $\mathcal{E}_\delta$ Algorithm, and the $\widehat{\mathcal{E}}_\delta$ algorithm, which is a slight modification of the $\mathcal{E}_\delta$ Algorithm.

2.4. **Implementing the interrupted algorithms with the help of the $\mathcal{HG}$ Algorithm.** This is the $\widehat{\mathcal{E}}_{1/2}$ algorithm which is used in Jebelean's Lemma. This lemma is a main tool to compute (in a recursive way) a function $\mathcal{HG}$ [for Half-gcd]. On an input (A, B) of binary size n, this function returns exactly the same result as $\widehat{\mathcal{E}}_{1/2}$, but runs faster. With the algorithm $\mathcal{HG}$, it is possible to design a fast algorithm denoted $\mathcal{G}$ which computes the gcd itself. Let us explain the main principles how the $\mathcal{HG}$ algorithm can be used inside the $\mathcal{E}_{[\delta,\delta+\gamma]}$ algorithm. This is described in Figure 3 and we comment now this figure.

Algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}(A,B)$

Input. The pair (A',B') obtained after P_δ iterations of the $\mathcal{E}$ Algorithm on the input pair (A,B).
Output. The pair (C',D') which would be obtained after $P_{\delta+\gamma}$ iterations of the $\mathcal{E}$ Algorithm on the input pair (A,B).

(*i*) Truncate the input (A',B'), with a truncation length $m = 2\gamma n$, and obtain a small pair (a,b) of length m.
(*ii*) Perform the $\mathcal{HG}$ algorithm on the pair (a,b), which produces a pair (c,d) and a matrix $\mathcal{M}$.
(*iii*) Perform the product $\left(\begin{smallmatrix} C \\ D \end{smallmatrix}\right) := \mathcal{M}^{-1} \left(\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right)$.
(*iv*) Use the Adjust function, which performs some steps of the Euclid Algorithm from the pair (C,D) and stops as soon the current remainder pair (C',D') has a size equal to $\lfloor (1-\delta-\gamma)n \rfloor$.

FIGURE 3. An implementation of the $\mathcal{E}_{[\delta,\delta+\gamma]}$ Algorithm using the $\mathcal{HG}$ algorithm.

Suppose that the Euclid Algorithm, on an input (A,B) of length n, has already performed P_δ iterations. Now, the current pair, denoted by (A',B') has a binary size close to $(1-\delta)n$. We may use the Jebelean Property to continue. Then, we choose a length m for truncating of the form $m = 2\gamma n$, an integer b of length m, and consider the small pair $(a,b) = \pi_{[b]}(A',B')$ with $\pi_{[b]}$ defined in Section 2.1. The $\mathcal{HG}$ algorithm on this pair (a,b) will produce a matrix $\mathcal{M}$ which would have been produced by the Euclid algorithm on the pair (A',B'). Then, the pair (C,D) computed as $\left(\begin{smallmatrix} C \\ D \end{smallmatrix}\right) = \mathcal{M}^{-1} \left(\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right)$ is a remainder pair of the Euclid algorithm on the input (A,B). The size of the matrix $\mathcal{M}$ is approximately $m/2$, but smaller than $m/2$ (due to the three backward steps of Lemma 1), and thus of the form $(m/2) - r(A,B)$, where $r(A,B)$ is the number of bits which are "lost" for the matrix $\mathcal{M}$ during the three backward steps. Then, with (8), $r(A,B)$ satisfies,

$$3 \leq r(A,B) \leq Q(A,B) \qquad \text{with} \quad Q(A,B) := \sum_{i=P_{1/2}(a,b)-2}^{P_{1/2}(a,b)} \ell(q_i) + 1. \tag{11}$$

Here, q_i are the quotients that occur in $\mathcal{E}(a,b)$, and $P_\delta(a,b)$ is defined in (10). If the truncature length m is chosen as a linear function of the input size n, of the form $m = 2\gamma n$, then the size of the pair (C,D) is approximately equal to $[1-\delta-\gamma]n$, but slightly larger. If we wish to obtain a remainder pair (C',D') of length $[1-\delta-\gamma]n$, we have to perform, from the pair (C,D) a certain number of steps of the Euclid Algorithm, in order to cancel the loss due to the backward steps. This is the goal of the Adjust function, whose cost $R(A,B)$ will be estimated with (9) as

$$3(1-\delta)n \leq R(A,B) \leq (1-\delta)n \cdot Q(A,B). \tag{12}$$

We recall that, in the papers where the worst-case of fast GCD's is studied, the authors suppose that Q is $O(1)$ (in the worst case). We will *prove* that the *mean value* of Q on Ω_n will be indeed asymptotic to a precise constant η, which will be defined later. Then, the asymptotic cost of Step (*iv*) will be of order $O(n)$.

Step (*iii*) performs a matrix product and uses a fast multiplication of type (3). The integer pair (A',B') has size $\approx (1-\delta)n$, while the coefficients of the matrix $\mathcal{M}^{-1}$ have size $\approx \gamma n$. Then, if there exists an integer K for which $(1-\delta) = K\gamma$, the total cost $S(A,B)$ of Step (*iii*) is "expected" to satisfy

$$4A_1 \frac{1-\delta}{\gamma} \mu(\gamma n) \leq S(A,B) \leq 4A_2 \frac{1-\delta}{\gamma} \mu(\gamma n). \tag{13}$$

Finally, we have designed an algorithm which produces the same result as the interrupted algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}$, and is described in Figure 3.

In Section 3.4, we shall state a class of results which prove that these last estimates (13) hold in the average case, as soon as a convenient choice of parameters δ, γ is done. In the same vein, these

results will prove that the mean value of parameter R on Ω_n is of order $O(n)$, which will entail, with (11, 12), that the cost R of the Adjust functions will be negligible with respect to the cost of matrix products.

2.5. **The usual designs for the $\mathcal{HG}$ and $\mathcal{G}$ algorithms.** How to use this idea for computing (in a recursive way) the $\mathcal{HG}$ Algorithm? The usual choice for γ is $\gamma = 1/4$, more precisely $m = \lceil n/2 \rceil$. Then, the previous description provides a method to obtain $\mathcal{E}_{[0,1/4]}$ (with a first choice $\delta = 0$), then $\widehat{\mathcal{E}}_{[1/4,1/2]}$ (with a second choice $\delta = 1/4$). Since $\widehat{\mathcal{E}}_{[0,1/2]} = \mathcal{E}_{[0,1/4]} \cdot \widehat{\mathcal{E}}_{[1/4,1/2]}$, we are done. Remark that using the "hat" algorithm in the second step leads to modifying the Adjust function for this step, which may also perform some backward steps in the Euclid Algorithm on the large inputs.

The general structure of the algorithm $\mathcal{HG}$ is described in Figure 3. The recursion is stopped when the naive algorithm $\widehat{\mathcal{E}}_{1/2}$ becomes competitive. This defines a threshold for the binary size denoted by S (remark that $S = S(n)$ is a function of the input size n).

With this $\mathcal{HG}$ algorithm, we can obtain an algorithm named $\mathcal{G}$ which computes the gcd. The idea for designing such an algorithm is to decompose the total Euclid Algorithm into interrupted algorithms, as

$$\mathcal{E}_{[0,1]} = \mathcal{E}_{[0,1/2]} \cdot \mathcal{E}_{[1/2,3/4]} \cdot \ldots \cdot \mathcal{E}_{[1-(1/2)^k,1-(1/2)^{k+1}]} \cdot \ldots$$

Then, the $\mathcal{HG}$ algorithm, when running on inputs of size $n/(2^k)$ produced by the $\mathcal{E}_{[0,1-(1/2)^k]}$ algorithm can easily simulate the $\mathcal{E}_{[1-(1/2)^k,1-(1/2)^{k+1}]}$ algorithm.

This decomposition also stops when the naive algorithm `gcd` becomes competitive. This defines a threshold for the length denoted by T (remark that $T = T(n)$ is also a function of the input size n).

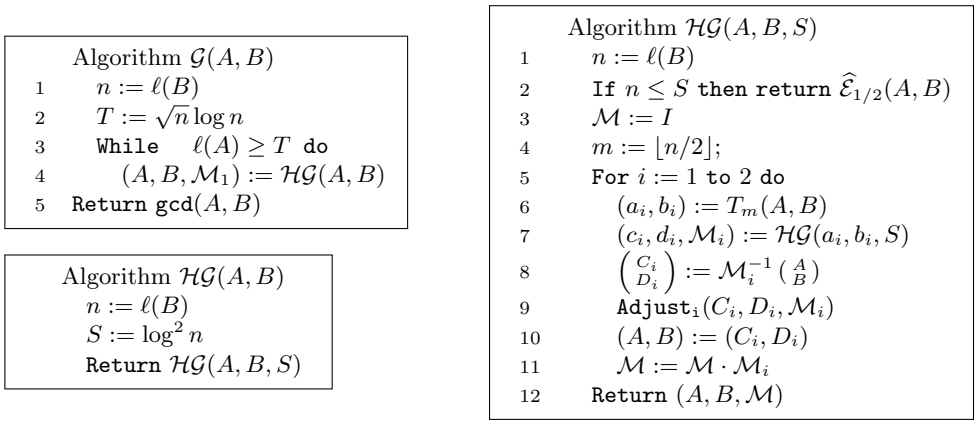

FIGURE 4. General structure of the classical algorithms $\mathcal{HG}$ and $\mathcal{G}$.

We now consider the $\mathcal{HG}$ Algorithm, where all the products use a FFT multiplication which satisfies (3). In this case, we choose the recursion depth H so that the main cost will be the "internal" cost, of order $\Theta(\mu(n)) \log n$, since the cost due to the leaves (where the naive $\widehat{\mathcal{E}}_{1/2}$ is performed) will be of asymptotic smaller order. Then, H satisfies the relation[3]

$$2^H \cdot \left(\frac{n}{2^H}\right)^2 \approx_{\leq} \mu(n) \log n,$$

$$\text{so that} \quad \frac{n}{2^H} \approx_{\leq} S(n) = \log^2 n, \quad H \approx_{\geq} \lg n - 2 \lg \lg n.$$

This is the "classical" version of the Knuth–Schönhage algorithm. Clearly, the cost of this algorithm comes from three types of operations:

(i) the two recursive calls of line 7;

[3]The notation $a(n) \approx_{\leq} b(n)$ means: There exist two constants A, B with $0 < A < B < 1$ for which
$$Ab(n) \leq a(n) \leq Bb(n)$$

(*ii*) the products done at lines 8 and 11: with a clever implementation, it is possible to use in line 8 the pair (c, d) just computed in line 7. If all the matrices and integer pairs have –on average– the expected size, the total expected cost due to the products is $[12 + 8 + 8]\ \mu(n/4) = 28 :, \Theta(1)\mu(n/4)$, where the constants hidden in the Θ–term are A_1, A_2 defined in (3);

(*iii*) the two functions `Adjust` performed at line 9, whose total average cost is $R(n)$.

We consider as the set of all possible inputs of the $\mathcal{HG}$ algorithm the set $\Omega := \{(u, v); \quad 0 \leq u \leq v\}$, and the set of all possible inputs of size n,

$$\Omega_n := \{(u, v); \quad 0 \leq u \leq v, \quad \ell(v) = n\} \tag{14}$$

is endowed with some probability $\mathbb{P}_n$. We denote by $B(n)$ the average number of bit operations performed by the algorithm $\mathcal{HG}$ on Ω_n. Since each of the two recursive calls is made on data with size $n/2$, it can be "expected" that $B(n)$ asymptotically satisfies

$$B(n) \approx 2B(\frac{n}{2}) + 28\,\Theta(1)\mu(\frac{n}{4}) + R(n) \qquad \text{for } n > S. \tag{15}$$

Moreover, the average cost $R(n)$ can be "expected" to be negligible with respect to the multiplication cost $\mu(n)$. If the FFT multiplication is used of type (3), the total average bit–cost is "expected" to be

$$B(n) \approx \Theta(\mu(n) \log n) = \Theta(n(\log n)^2\, a(n)),$$

where the constants hidden in the Θ-terms are $7A_1, 7A_2$, with A_1, A_2 defined in (3).

With this (heuristic) analysis of the $\mathcal{HG}$ algorithm, it is easy to obtain the (heuristic) average bit–complexity of the $\mathcal{G}$ algorithm which makes a recursive use of the $\mathcal{HG}$ algorithm and stops as soon as the naive algorithm becomes competitive. It then stops at a recursion depth M, when

$$\Big(\frac{n}{2^M}\Big)^2 \approx_{\leq} \mu(n) \log n,$$

so that

$$\frac{n}{2^M} \approx_{\leq} T(n) = \sqrt{n} \log n, \quad M \approx_{\geq} \frac{1}{2} \lg n - \lg\lg n.$$

The average bit–cost $G(n)$ of the $\mathcal{G}$ algorithm on data of size n satisfies

$$G(n) \approx \sum_{i=0}^{M-1} B(\frac{n}{2^i}) \quad \text{ so that } \quad G(n) \approx \Theta(B(n)).$$

3. The main steps towards a proven analysis.

The analysis is based on the Divide and Conquer equation (15), which is not a "true" equality. It is not clear why a "true" equality should hold, since each of the two recursive calls is done on data which do not possess a priori the same distribution as the input data. And, of course, the same problem will be asked at each depth of the recursion. If we wish a "Divide and Conquer" probabilistic approach to be possible, we have to make precise *the evolution of the distribution during the Euclid Algorithm, but also the distribution of the truncated data.*

We first state in Section 3.1 our main two results, Theorems 1 and 2, which are of general interest. In particular, Theorem 2 involves the density ψ already defined in (2) which plays a central rôle in our analysis. These theorems are stated here, but not proved. This will be done in Section 6. Then, in Section 3.3, we explain how Theorem 2 can be applied to truncated data, as soon as the truncation is a probabilistic one, defined in Section 3.2. Section 3.4 explains the analysis of the Adjust Functions, and provides estimates for the mean bit–complexity of the interrupted algorithms described in Section 2.4, in particular the mean–complexity of Steps (*iii*) and (*iv*).

3.1. **Evolution of densities.** Consider a density f on the unit interval $= [0,1]$, which is "extended" to the set $\Omega := \{0 \leq u < v\}$ via the equality $f(u,v) := f(u/v)$. The set Ω_n formed with the inputs of size n, already defined in (14), namely $\Omega_n := \{0 \leq u < v,\ \ell(v) = n\}$ is endowed with the restriction of f to Ω_n: for any pair $(u,v) \in \Omega_n$,

$$\mathbb{P}_{n,f}(u,v) := \frac{1}{|\Omega_n|_f} f(\frac{u}{v}), \qquad \text{where} \quad |\Omega_n|_f := \sum_{(u,v)\in\Omega_n} f(\frac{u}{v}) \tag{16}$$

is the total f–weight of the set Ω_n. Remark that, for $f \equiv 1$, we recover the uniform density on Ω_n. For reasons which will appear later, the subsets $\widetilde{\Omega}, \widetilde{\Omega}_n$ formed with coprime inputs

$$\widetilde{\Omega} := \{(u,v) \in \Omega, \quad \gcd(u,v) = 1\}, \tag{17}$$

$$\widetilde{\Omega}_n := \{(u,v) \in \Omega, \quad \gcd(u,v) = 1, \ell(v) = n\}, \tag{18}$$

play an important (intermediate) rôle. We endow $\widetilde{\Omega}_n$ with the probability $\widetilde{\mathbb{P}}_{n,f}$ defined in the same vein as in (16).

The evolution of the density during the execution of the Euclid Algorithm is of crucial interest. For $(u,v) \in \Omega$, the Euclid Algorithm creates a sequence of successive remainders u_k, with $u_0 := v, u_1 := u, \dots u_p := \gcd(u,v)$. The corresponding integer pairs are denoted by $U_k := (u_{k+1}, u_k)$, and the corresponding rationals are denoted by $x_k := u_{k+1}/u_k$. We recall that $P_\delta(u,v)$ is the smallest integer k for which $\lg u_k < (1-\delta)\ell(u_0)$. We are interested in describing the density of the pair $U_{\langle\delta\rangle}$ defined as

$$U_{\langle\delta\rangle} := U_k \qquad \text{when} \quad P_\delta(u,v) = k.$$

This integer pair is the input for all interrupted algorithms with a beginning parameter δ. Since the density on Ω_n is defined via the associated rationals, the position of rational

$$x_{\langle\delta\rangle} := x_k \qquad \text{when} \quad P_\delta(u,v) = k$$

inside the interval $[0,1]$ will be essential.

We are interested in the study of the random variable P_δ: Since the rational x loses $\ell(x)$ bits during $P(x)$ iterations, it can be expected that it loses $\delta\ell(x)$ bits during $\delta P(x)$ iterations, which would imply that $P_\delta(x)$ is sufficiently close to $\delta P(x)$. This is what we call the regularity of the algorithm.

We do not succeed to directly study these two variables $P_\delta, x_{\langle\delta\rangle}$, and we replace them by some of their probabilistic variants, as we now explain. Consider, for some $\rho > 0$ with $\rho \leq (1-\delta)$, the interval $[2^{(1-\delta)n}(1-2^{-\rho n}), 2^{(1-\delta)n}]$, and draw an integer W uniformly in this interval. Denote by $\underline{P}_\delta$ the first integer k for which u_k is less than W, and by $\underline{x}_{\langle\delta\rangle}$ the rational x_k. The two underlined variables define probabilistic variants of the plain variables. Since they depend on parameter ρ, we call them the ρ–probabilistic variants. Moreover, as soon as n is sufficiently large ($n > 1/\rho$), the interval is contained in an interval $]A/2, A]$ and contains at most two possible rationals x_k (this is due to the fact that $u_{k+2} \leq (1/2)u_k$). This proves, that in the case when $n > 1/\rho$, the probabilistic variable $\underline{x}_{\langle\delta\rangle}$ equals $x_{\langle\delta\rangle}$, $x_{\langle\delta\rangle+1}$, or $x_{\langle\delta\rangle+2}$, while the variables P_δ and $\underline{P}_\delta$ satisfy $|P_\delta - \underline{P}_\delta| \leq 2$.

With techniques close to the renewal methods, we prove a quasi-powers expression for the moment generating function of $\underline{P}_\delta$, from which we deduce an asymptotic gaussian law for $\underline{P}_\delta$ on Ω, then an asymptotic gaussian law for the deterministic variable P_δ on Ω. We then obtain an extension of the result of Baladi-Vallée [2] (which exhibits an asymptotic gaussian law for $P := P_1$), even if our proof cannot directly apply to $\delta = 1$.

Theorem 1. *Consider the set Ω_n endowed with a probability $\mathbb{P}_{n,f}$ relative to a strictly positive function f of class $\mathcal{C}^1$. Then, for any $\delta \in]0,1]$, the random variable P_δ is asymptotically gaussian on Ω_n [with a speed of convergence of order $O(n^{-1/3})$]. Moreover, if $\rho(\delta) := (1/2)\ \min(\sigma, 1/2)\ \min(\delta, 1-\delta)$ where σ is a strictly positive lower bound for the width of the* US *Strip, the $\rho(\delta)$–probabilistic variant $\underline{P}_\delta$ of P_δ satisfies*

$$\mathbb{E}_{n,f}[\underline{P}_\delta] = 2\log 2 \frac{1}{|\Lambda'(1)|}\,\delta n + D_1 + O(2^{-n\rho(\delta)}),$$

$$\mathbb{V}_{n,f}[\underline{P}_\delta] = 2\log 2 \left|\frac{\Lambda''(1)}{\Lambda'(1)^3}\right|\,\delta n + D_2 + O(2^{-n\rho(\delta)}).$$

Here, D_1, D_2 are some constants. The constants D_1, D_2 and the constant in the O-term only depend on the function f.

Our second result is related to the distribution of the probabilistic variant $\underline{x}_{\langle\delta\rangle}$, and, here, it does not seem possible to derive some information for the deterministic variable $x_{\langle\delta\rangle}$.

Theorem 2. *Denote by σ a strictly positive lower bound on the width of the* US *strip and let $\underline{\sigma} := \min(\sigma, 1/2)$. Denote by ψ the density defined in (2). Consider a strictly positive density f of class $\mathcal{C}^1$, a real δ with $0 < \delta < 1$, an interval $J \subset I$ whose length $|J|$ satisfies $|\lg(|J|)| < (1/2)\underline{\sigma}(1-\delta)$, and denote by $\rho(\delta)$ the real defined as $\rho(\delta) := (1/2)\underline{\sigma}\min(1-\delta, \delta)$. Then, the probability that the $\rho(\delta)$–probabilistic rational $\underline{x}_{\langle\delta\rangle}$ computed by the Euclid Algorithm belongs to the interval J satisfies*

$$\mathbb{P}_{n,f}[\underline{x}_{\langle\delta\rangle} \in J] = \left(\int_J \psi(t)dt\right) \cdot \left[1 + O\left(2^{-n\rho(\delta)}\right)\right].$$

The constant in the O-term only depends on the function f via its norm $||f||_1 := \sup|f| + \sup|f'|$.

3.2. **Probabilistic truncations.** Finally, we are also interested by the distribution of the truncated pairs. We recall that the truncated pairs classically used are obtained with truncations of "numerator" A and "denominator" B of pair (A, B). It is not clear how to reach the distribution of such truncated pairs. This is why we define a probabilistic truncation, which leads to more regular distributions, and also allows us to apply Jebelean's Property (Lemma 1).

For $x = (A, B) \in \Omega_n$, and $m \leq n$, we define $\pi_m(A, B)$ as follows:

(1) Choose a denominator b in the set $\{v, \ell(v) = m\}$ of integers of binary size m, with a probability proportional to b. More precisely, we choose a denominator b according to the law

$$\Pr[b = b_0] = \frac{1}{\theta_m} \cdot b_0 \qquad \text{with} \quad \theta_m = \sum_{b;\ell(b)=m} b.$$

(2) Compute the integer a which is the integer part of $x \cdot b$. This computation involves the product $A \cdot b$ then the division of the integer $A \cdot b$ by B. This can be done in $O(\mu(n))$ with a O–constant larger than the constant of the multiplication(see Equation (4). Of course, this does not give rise to a very efficient algorithm. However, we will see that using this probabilistic truncation does not change the order of the average complexity of the $\mathcal{HG}$ algorithm. We return to this remark in Theorem 5.

(3) Define $\pi_m(A, B)$ as the pair (a, b), and remark that the set $\pi_m^{-1}(a, b)$ is the pairs (C, D) of Ω_n for which the associated rational C/D belongs to the interval

$$J\big(\frac{a}{b}\big) := \big[\frac{a}{b}, \quad \frac{a}{b} + \frac{1}{b}\big[, \quad \text{with} \quad \left|J\big(\frac{a}{b}\big)\right| = \frac{1}{b} = \Theta(2^{-m}).$$

This is sufficient for applying Jebelean's criterion (Lemma 1).

We start with a strictly positive density f of class $\mathcal{C}^1$ on $[0, 1]$, and for any integer m, the function $g_m = g_m[f]$ defined on Ω_m as

$$g_m[f](u, v) = \frac{1}{|J(y)|}\int_{J(y)} f(t)dt, \qquad \text{with} \quad y := \frac{u}{v}$$

only depends on the rational u/v and satisfies $\mathbb{P}_{n,f}[(A, B); \pi_m(A, B) = (a, b)] = \mathbb{P}_{m,g_m[f]}(a, b)$. Furthermore, for any $(u, v) \in \Omega_m$, the relation

$$g_m[f](u, v) = f(\frac{u}{v}) + O(|J(\frac{u}{v})| \cdot \|f\|_1)$$

proves that the function $g_m[f]$ (viewed as a function defined on $\mathbb{Q}$) is a smoothed version of the initial function f. Furthermore,

$$\frac{\mathbb{P}_{m,g_m[f]}}{\mathbb{P}_{m,f}} = 1 + O(2^{-m}).$$

Since f is a density on $[0, 1]$, the cumulative sum of $g_m[f](x)$ on Ω_m satisfies

$$\sum_{(u,v)\in\Omega_m} g_m[f](u, v) = \sum_{\ell(v)=m} v\left[\sum_{u<v}\left(\int_{J(\frac{u}{v})} f(t)\right)\right] = \theta_m\left(\int_I f(t)dt\right) = \theta_m.$$

This allows a comparison between two probabilities:

Lemma 2. *Consider a strictly positive density f of class $\mathcal{C}^1$ on I. For any n, for any $m \leq n$, for any $(a,b) \in \Omega_m$, one has*

$$\mathbb{P}_{n,f}[(A,B); \pi_m(A,B) = (a,b)] = \mathbb{P}_{m,f}(a,b) \cdot [1 + O(2^{-m})],$$

where the constant in the O-term only depends on f via its norm $||f||_1 := \sup|f| + \sup|f'|$.

3.3. **Truncations and evolution of densities.** In our framework, the truncation length m is linear with respect to the input size n, of the form $m = 2\gamma n$, and, in this case, we denote π_m by $\pi_{\langle 2\gamma\rangle}$. With Theorem 2 and the previous comparison of densities done in Lemma 2, we obtain the following result which will be a central tool in our analysis. When

Theorem 3. *Denote by σ a strictly positive lower bound on the width of the* US *strip and let $\underline{\sigma} := \min(\sigma, 1/2)$. Denote by ψ the density defined in (2). Consider a real $\delta \in [0,1[$, and a parameter γ strictly less than $(1/2)(1-\delta)\underline{\sigma}$, and denote by $\rho(\delta,\gamma)$ the real defined by*

$$\rho(0,\gamma) = 2\gamma, \qquad \rho(\delta,\gamma) := \min\{\underline{\sigma}(1-\delta) - 2\gamma,\ (1/2)\underline{\sigma}\delta,\ 2\gamma\} \qquad \text{for } \delta > 0,\ 2\gamma < \underline{\sigma}(1-\delta)\ .$$

Then, the distribution of the $\langle 2\gamma\rangle$–truncation of the $\rho(\delta,\gamma)$–probabilistic rational $\underline{x}_{\langle\delta\rangle}$ computed by the Euclid Algorithm satisfies

$$\mathbb{P}_{n,\psi}[x;\ \pi_{\langle 2\gamma\rangle}(\underline{x}_{\langle\delta\rangle}) = y_0] = \mathbb{P}_{m,\psi}[y_0] \cdot \left[1 + O(2^{-n\rho(\delta,\gamma)})\right].$$

3.4. **Mean bit–complexity of the interrupted algorithm $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$.** We return now to the algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}$ defined in Figure 3 and we use the notations of Section 2.4. We will study a probabilistic version of the algorithm $\mathcal{E}_{[\delta,\delta+\gamma]}$ which will be denoted by $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$. We now describe the main differences between $\mathcal{E}_{[\delta,\delta+\gamma]}$ and its probabilistic version. In the probabilistic version $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$:

- (a) the input pair of the algorithm is the pair $\underline{U}_{\langle\delta\rangle}$ relative to the parameter $\rho(\delta,\gamma)$
- (b) the output pair of the algorithm is the pair $\underline{U}_{\langle\delta+\gamma\rangle}$ relative to the parameter $\rho(\delta+\gamma,\gamma)$
- (c) Step (i) uses the probabilistic truncature $\pi_{\langle 2\gamma\rangle}$ defined In Section 3.2 and 3.3.

As in the initial $\mathcal{E}_{[\delta,\delta+\gamma]}$, Step ($iii$) uses any fast multiplication of type (3).

We first analyse the mean cost R of the Adjust function performed in Step (iv), which deals with the probabilistic version $\underline{Q}$ of parameter Q defined in (11). In fact, we study a more general parameter $\underline{Q}_\delta$ which involves the size of quotients, when the pair (u,v) has already lost a fraction δ of its bits,

$$\underline{Q}_\delta(u,v) := \sum_{i=\underline{P}_\delta(u,v)-2}^{\underline{P}_\delta(u,v)} \ell(q_i),$$

and the (initial) parameter $\underline{Q}$ is obtained for $\delta = 1/2$. A central result is :

Theorem 4. *Consider the set Ω_n endowed with a probability $\mathbb{P}_{n,f}$ relative to a strictly positive function f of class $\mathcal{C}^1$. Then, for any $\delta \in]0,1]$, the mean value of the cost $\underline{Q}_\delta$ is asymptotic to a constant η, which does not depend on δ and density f, and involves the Gauss density φ defined in (1), together with the operators $\mathbf{H}_{s,[\ell]}$ and $\mathbf{H}'_s$ defined in (30) and (31), under the form*

$$\mathbb{E}_{n,f}[\underline{Q}_\delta] = \eta\ \left[1 + O(2^{-n\rho(\delta)})\right] \qquad \text{with} \quad \eta := \frac{-6\log 2}{\pi^2} \int_I \mathbf{H}'_1 \circ \mathbf{H}^3_{1,[\ell]}[\varphi](t)dt, \tag{19}$$

where $\rho(\delta) := (1/2)\underline{\sigma}\min(1-\delta,\delta)$ is the constant of Theorem 2.

This following result studies the bit-complexity of the Interrupted Algorithm $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$ and proves two facts: First, the cost of the multiplications performed in Step (iii) is exactly of the same order as this expected. Second, the cost of the Adjust function performed in Step (iv) is negligible with respect to costs of Step (iii).

Theorem 5. *Consider two parameters γ, δ satisfying $\gamma < (1/2)\underline{\sigma}(1-\delta)$, with the constant $\rho(\delta,\gamma)$ from Theorem 3. Then, the probabilistic version $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$ of the $\mathcal{E}_{[\delta,\delta+\gamma]}$ algorithm described in the beginning of this Section 3.4. satisfies the following:*

(*i*) *In the case when the ratio* $(1-\delta)/\gamma$ *is integer, the mean bit–complexity cost* $\mathbb{E}_{n,\psi}[S]$ *of Step* (*iii*) *satisfies :*

$$\mathbb{E}_{n,\psi}[S] = \Theta(1)\frac{1-\delta}{\gamma}\mu(\gamma n)[1+O(2^{-n\rho(\delta,\gamma)})],$$

where the hidden constants in the Θ*–term are independent on the pair* (γ,δ) *and can be chosen as* $4A_1, 4A_2$ *for constants* A_1, A_2 *relative to the fast multiplication defined in (3).*

(*ii*) *The mean bit–complexity cost* $\mathbb{E}_{n,\psi}[R]$ *of Step* (*iv*) *satisfies :*

$$\mathbb{E}_{n,\psi}[R] = (1-\delta)n\eta\ \left[1+O(2^{-n\rho(\delta,\gamma)})\right]$$

and involves the constant L *defined in (19).*

(*iii*) *The mean bit–complexity cost* $\mathbb{E}_{n,\psi}[T]$ *of Step* (*i*) *satisfies*

$$\mathbb{E}_{n,\psi}[T] = \Theta(1)\frac{1-\delta}{\gamma}\mu(\gamma n)[1+O(2^{-n\rho(\delta,\gamma)})],$$

where the hidden constants in the Θ*–term are independent on the pair* (γ,δ) *and can be chosen as* $2\max(A_1,A_3), 2\max(A_2,A_4)$ *for constants* A_1, A_2 *relative to the fast multiplication defined in (3), and constants* A_3, A_4 *relative to the fast division defined in (4).*

(*iv*) *The total bit-complexity of Steps* (*i*), (*iii*) *and* (*iv*) *is*

$$\mathbb{E}_{n,\psi}[S+R+T] = \Theta(1)\frac{1-\delta}{\gamma}\mu(\gamma n)\left[1+O\left(\frac{1}{\log(\gamma n)a(\gamma n)}\right)\right]$$

and involves the functions $\mu(n)$ *and* $a(n)$ *associated to the fast multiplication. As previously, the hidden constants in the* Θ*–term are independent on the pair* (γ,δ) *and can be chosen as* $4A_1', 4A_2'$,

$$A_1' := \max\left(A_1, \frac{A_3}{2}\right), \qquad A_2' := \max\left(A_2, \frac{A_4}{2}\right) \tag{20}$$

and involve constants A_i, A_i' *defined in (3,4). The hidden constants in the O–term is independent on the pair* (γ,δ) *too.*

4. The algorithms to be analyzed.

There are three main differences between the usual $\mathcal{HG}$ and $\mathcal{G}$ Algorithm and our versions to be analyzed which are denoted as $\underline{\mathcal{HG}}$ and $\underline{\mathcal{G}}$. See Figure 5.

(*i*) Our algorithms are randomized, since we will use the probabilistic variants $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$ of the interrupted algorithms $\mathcal{E}_{[\delta,\delta+\gamma]}$.

(*ii*) For the $\underline{\mathcal{HG}}$ algorithm, the number L of recursive calls and the degree 2γ of truncatures (i.e., the ratio m/n) are not the same as in the $\mathcal{HG}$ Algorithm. The algorithm $\underline{\mathcal{HG}}$ is also built as a Divide and Conquer Algorithm; however, the relation which relates the two parameters γ,δ with $\underline{\sigma}$, crucial for applying Theorems 3 and 5, leads to a recursive algorithm $\underline{\mathcal{HG}}$ with L recursive calls, where L depends on parameter $\underline{\sigma}$ of the *US* strip and satisfies $(L+1) > 2/\underline{\sigma}$.

(*iii*) The study is done when the initial density equals ψ, since it is quasi-invariant under the recursive calls. This choice makes easier the study of various recursions. The constants which appear in Theorems 6 and 7 are relative to this particular case. Since any other strictly positive density f satisfies

$$\frac{\min f}{\max \psi} \le \frac{\mathbb{E}_{n,f}[C]}{\mathbb{E}_{n,\psi}[C]} \le \frac{\max f}{\min \psi},$$

Theorems 6 and 7 hold with any strictly positive density, with other constants, which depend on f.

As before, the recursive calls in the $\underline{\mathcal{HG}}$ Algorithm are stopped when the naive $\widehat{\mathcal{E}}_{1/2}$ Algorithm becomes competitive. The calls of the $\underline{\mathcal{G}}$ Algorithm to the $\underline{\mathcal{HG}}$ algorithm are stopped when the naive gcd algorithm becomes competitive.

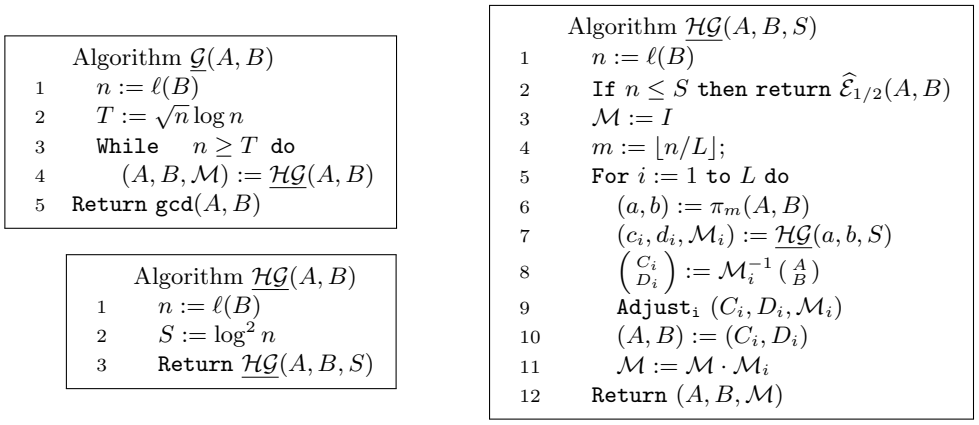

FIGURE 5. General structure of the algorithms $\mathcal{HG}$ and $\mathcal{G}$ to be analyzed. The number of recursive calls L satisfies $L > (2/\underline{\sigma}) - 1$.

4.1. **The first recursive call.** Inside the first recursive call of $\underline{\mathcal{G}}$ to $\underline{\mathcal{HG}}$, the parameter δ belongs to $[0, 1/2]$. We suppose that there are $L \geq 2$ recursive calls of $\underline{\mathcal{HG}}$ to himself. We denote by B_L the bit–complexity of the $\underline{\mathcal{HG}}$ Algorithm when it performs L recursive calls, and we analyse the asymptotic behaviour of the mean value $\mathbb{E}_{n,\psi}[B_L]$ (for $n \to \infty$).

Suppose indeed $L \geq 2$. Then, the possible values for pairs (δ, γ) of the first recursive call satisfy

$$\delta \in \Delta_1 := \left\{ \frac{i}{2L}, \quad \text{with} \quad 0 \leq i \leq L-1 \right\}, \qquad \gamma_1 := \frac{1}{2L}, \tag{21}$$

and the pairs relative to the h-th recursive call are

$$\delta \in \Delta_h := \left\{ \frac{i}{2L^h}, \quad \text{with} \quad 0 \leq i \leq L^h - 1 \right\} \qquad \gamma_h := \frac{1}{2L^h}.$$

We stop the recursion at a level H for which the total bit–cost $P(n)$ of the naive gcd computations is negligible with respect to the total cost of the algorithm. More precisely, if $a(n)$ is the function which intervenes in the multiplication cost, we ask

$$P(n) = \Theta\left(L^H \cdot \left(\frac{n}{L^H}\right)^2\right) = n\log^2 n = \frac{\mu(n)\log n}{a(n)}, \qquad H \sim \left(\frac{\log n}{\log L}\right), \quad \frac{n}{L^H} = \Theta(\log^2 n). \tag{22}$$

The parameter $\rho(\delta, \gamma)$ must be strictly positive, first for $\delta \in \Delta_1$. This is only possible if

$$L > \frac{2}{\underline{\sigma}} - 1,$$

and, in this case, the minimum value of $\rho(\delta, \gamma)$ at the h-th recursion level satisfies

$$\exists K > 0, \qquad \forall h \geq 1, \qquad \min\{\rho(\delta, \gamma_h), \delta \in \Delta_h\} \geq \frac{K}{L^h}. \tag{23}$$

With (23), Theorem 3 entails the following Divide and Conquer probabilistic equation,

$$\mathbb{E}_{n,\psi}[B_L] = \left(\sum_{\delta\in\Delta_1} \mathbb{E}_{\delta n,\psi}[B_L]\right) \cdot \left[1 + O\big(2^{-nK/L}\big)\right] + C_{n,1},$$

where $C_{n,1}$ is the total bit–complexity of steps Steps $(i), (iii)$ and (iv) performed during the executions of the $\underline{\mathcal{E}}_{[\delta,\delta+\gamma]}$ Algorithm, together with the matrix product performed in Line 11, for $\delta \in \Delta_1$ easily estimated with Theorem 5. Expanding the recursion (always with Theorem 3) leads to the estimate

$$\mathbb{E}_{n,\psi}[B_L] = \left(P(n) + \sum_{h=1}^{H} C_{n,h}\right) \left[\prod_{h=1}^{H} 1 + O\left(2^{-nK/L^h}\right)\right]$$

where $C_{n,h}$ is the total mean cost of all the Steps $(i), (iii)$ and (iv) of the interrupted algorithms at the h–th level, corresponding to $\delta \in \Delta_h, \gamma := \gamma_h$. The error term comes from the comparison of the distributions made with Theorem 3, and is of the form, with (22) and (23)

$$1 + O(\varepsilon(n)), \quad \text{with} \quad \varepsilon(n) = \sum_{h=1}^{H} 2^{-nK/L^h} \leq H 2^{-nK/L^H} = \Theta(\log n)\, 2^{-K \log^2 n} = O(n^{-K_1 \log n}).$$

The cost $C_{n,h}$ at the h–th recursion level is easily evaluated with Theorem 5. We let $b(n) := a(n) \log n$. For $h = 1$, Theorem 5 entails the estimate

$$C_{n,1} = \Theta(1) \left[\sum_{i=1}^{L} 2L(1 - \frac{i}{2L})\right] \mu(\frac{n}{2L}) \left[1 + O\left(\frac{1}{b(n/L)}\right)\right] + \Theta(1) \left[\sum_{I=1}^{L} i\right] \mu(\frac{n}{2L}) \left[1 + O\left(\frac{1}{b(n/L)}\right)\right]$$

where the first term is due to the cost of the interrupted algorithms and the second term to matrix products of Line 11. One has

$$C_{n,1} = \Theta(L^2)\, \mu(\frac{n}{2L}) \left[1 + O\left(\frac{1}{b(n/L)}\right)\right]$$

where the hidden constants are now respectively $6A_1' + 8A_1, 6A_2' + 8A_2$, with (A_1', A_2') defined in (20) and A_1, A_2 defined in (3). In the same vein,

$$C_{n,h} = \Theta(L^{h+1}) \mu(\frac{n}{2L^h}) \left[1 + O\left(\frac{1}{b(n/L^h)}\right)\right],$$

and finally

$$\sum_{h=1}^{H} C_{n,h} = \Theta(1) \frac{L}{2 \log L} \mu(n) \log n \cdot \left[1 + O\left(\frac{1}{b(\log^2 n)}\right)\right] =$$

where the constants in the Θ–term are always respectively $6A_1' + 8A_1, 6A_2' + 8A_2$. Now, with (22), the error term due to the leaves is of the form $1/a(n)$, and the function $b(\log^2 n)$ is larger than $a(n)$. Finally,

$$\mathbb{E}_{n,\psi}[B_L] = \Theta(1) \frac{L}{2 \log L} \mu(n) \log n \cdot \left[1 + O\left(\frac{1}{a(n)}\right)\right]$$

where the constants in the Θ–term are always respectively $6A_1' + 8A_1, 6A_2' + 8A_2$.

Theorem 6. *Consider the $\underline{\mathcal{HG}}$ algorithm defined in Figure 3, relative to a parameter L which satisfies $L > (2/\underline{\sigma}) - 1$, and involves $\underline{\sigma} := \max(\sigma, 1/2)$, where σ is a strictly positive lower bound for the* US *strip. Suppose that the algorithm uses a fast multiplication of type (3). Then, the mean bit–complexity B_L of this $\underline{\mathcal{HG}}$ algorithm on the set Ω_n endowed with the density ψ defined in (2) satisfies*

$$\mathbb{E}_{n,\psi}[B_L] = \Theta\left(\frac{L}{\log L}\right) n\, (\log n)^2\, a(n) \cdot \left[1 + O\left(\frac{1}{a(n)}\right)\right].$$

Here, the constants in the Θ–term can be chosen as $3A_1' + 4A_1, 3A_2' + 4A_2$, where A_1', A_2' defined in (20) are the constants related to the fast multiplication and the fast division .
The mean bit–complexity B_L of this $\underline{\mathcal{HG}}$ algorithm on the set Ω_n endowed with any density f of class $\mathcal{C}^1$ satisfies

$$\mathbb{E}_{n,f}[B_L] = \Theta(1)\, n\, (\log n)^2\, a(n) \cdot \left[1 + O\left(\frac{1}{a(n)}\right)\right].$$

Here, the constants in the Θ–term can be chosen as

$$\frac{\min f}{\max \psi} \max(7A_1, 4A_1 + \frac{3}{2}A_3), \qquad \text{and} \qquad \frac{\max f}{\min \psi} \max(7A_2, 4A_2 + \frac{3}{2}A_4),$$

where A_1, A_2 are the constants related to the fast multiplication and A_3, A_4 are the constants related to the fast division.

4.2. **The k-th recursive call.** The k-th recursive call of $\underline{\mathcal{G}}$ to $\underline{\mathcal{HG}}$ is made on integers with size $n_k = n(1/2)^{k-1}$. It deals with values $\delta^{(k)}$ which belong to the interval $[1-(1/2)^{k-1}, 1-(1/2)^k]$, so that the values $(1-\delta^{(k)})n$ belong to the interval $[n_k, n_k/2]$. If we wish to perform at the k-th level an algorithm $\underline{\mathcal{HG}}$ homothethic to the algorithm of the first level [with a ratio $(1/2)^{k-1}$], we deal with a truncation m_k of the form $m_k = 2\gamma^{(1)}n_k = 2\gamma^{(k)}n$ with $\gamma^{(k)} = 1/(2^{k-1}L)$. Now the parameter $\rho(\delta^{(k)}, \gamma^{(k)})$ relative to values $\delta^{(k)}, \gamma^{(k)}$ used in the kth recursive call of $\underline{\mathcal{G}}$ to $\underline{\mathcal{HG}}$ is related to the parameter $\rho(\delta^{(1)}, \gamma^{(1)})$ relative to values $\delta^{(1)}, \gamma^{(1)}$ used in the first recursive call of $\underline{\mathcal{G}}$ to $\underline{\mathcal{HG}}$, via the inequality

$$n\,\rho(\delta^{(k)}, \gamma^{(k)}) \geq n_k\,\rho(\delta^{(1)}, \gamma^{(1)}).$$

Then, all the previous study performed for the first recursive call can be applied to the k-th recursive call, as soon as n is replaced by n_k.

We then choose L equal to 2, as previously, and the bit-complexity $B_{k,L}$ of the k-th recursive call is, with Theorem 6,

$$\mathbb{E}_{n,\psi}[B_{k,L}] = \Theta\left(\frac{L}{\log L}\right) n_k(\log n_k)^2\, a(n_k) \cdot \left[1 + O\left(\frac{1}{a(n_k)}\right)\right]. \tag{24}$$

with the same constants involved as in Theorem 6.

4.3. **End of the recursion.** We stop calling the algorithms $\underline{\mathcal{HG}}$ inside the $\underline{\mathcal{G}}$ algorithm when the naive gcd algorithm becomes competitive, with a complexity $P_1(n) = \Theta(n\log^2 n)$. Then, the level of recursion M is defined by

$$n_M^2 = n\log^2 n \quad \text{so that} \quad n_M = \sqrt{n}\log n, \quad M = (1/2)(\log n).$$

Then the total cost G of the $\underline{\mathcal{G}}$ Algorithm satisfies

$$\mathbb{E}_{n,\psi}[G] = \sum_{k=1}^{M} \mathbb{E}_{n,\psi}[B_{k,L}] = \Theta\left(\frac{L}{\log L}\right) n\ (\log n)^2\, a(n)\ \cdot \left[1 + O\left(\frac{1}{a(\sqrt{n}\log n)}\right)\right]$$

where the constants in the Θ–term are equal to two times the constants of Theorem 6. Finally, we have proven the following:

Theorem 7. *Consider the $\underline{\mathcal{HG}}$ algorithm defined in Figure 3, relative to a parameter L which satisfies $L > (2/\underline{\sigma}) - 1$, and involves $\underline{\sigma} := \max(\sigma, 1/2)$, where σ is a strictly positive lower bound for the* US *strip. Suppose that the algorithm uses a fast multiplication of type (3). Then, the mean bit–complexity G_L of this $\underline{\mathcal{G}}$ algorithm on the set Ω_n endowed with the density ψ defined in (2) satisfies*

$$\mathbb{E}_{n,\psi}[G_L] = \Theta\left(\frac{L}{\log L}\right) n\ (\log n)^2\, a(n)\ \cdot \left[1 + O\left(\frac{1}{a(\sqrt{n}\log n)}\right)\right].$$

Here, the constants in the Θ–term can be chosen as $\max(14A_1, 8A_1+3A_3), \max(14A_2, 8A_2+3A_4)$, where A_1, A_2 are the constants related to the fast multiplication and A_3, A_4 are the constants related to the fast division.
The mean bit–complexity G_L of this $\underline{\mathcal{G}}$ algorithm on the set Ω_n endowed with any density f of class $\mathcal{C}^1$ satisfies

$$\mathbb{E}_{n,f}[G_L] = \Theta\left(\frac{L}{\log L}\right) n\,(\log n)^2\, a(n) \cdot \left[1 + O\left(\frac{1}{a(\sqrt{n}\log n)}\right)\right].$$

Here, the constants in the Θ–term can be chosen as

$$\frac{\min f}{\max \psi}\max(14A_1, 8A_1+3A_3), \qquad \text{and} \qquad \frac{\max f}{\min \psi}\max(14A_2, 8A_2+3A_4),$$

where A_1, A_2 are the constants related to the fast multiplication and A_3, A_4 are the constants related to the fast division.

5. Description of the Dynamical Analysis method.

Here, we present the main tools which will be used in the proof of Theorems 1 and 2. These tools come from analysis of algorithms (generating functions, here of Dirichlet types, described in 5.1) or dynamical systems theory (mainly transfer operators $\mathbf{H}_s$, described in 5.3 and 5.4). We introduce the main costs C of interest (in 5.2), and their related Dirichlet series, for which we provide an alternative expression with the transfer operator (in 5.5). For obtaining the asymptotic estimates of Theorems 1, 2, we extract coefficients from these Dirichlet series, in a "uniform way". Then, Property *US* (already described in 1.3) is crucial here for applying with success the Perron Formula, as in previous results of Baladi and Vallée [2].

5.1. **Dirichlet series.** For analysing a cost C, we deal with the generating Dirichlet series of this cost C. We recall that we deal with the sets $\Omega, \widetilde{\Omega}$ of all possible inputs, and their subsets $\widetilde{\Omega}_n, \Omega_n$ which gather the inputs (u,v) with $\ell(v)=n$ defined in (17). We will explain later why it is easier and also sufficient to deal with inputs of $\widetilde{\Omega}$ (which is, from the algorithmic point of view, the set of trivial inputs...). We consider these sets endowed with probability $\mathbb{P}_{n,f}$ or $\widetilde{\mathbb{P}}_{n,f}$ defined from a positive function f of the interval $\mathcal{I}$ as

$$\mathbb{P}_{n,f}(u,v) := \frac{1}{|\Omega_n|_f} f(\frac{u}{v}), \qquad \widetilde{\mathbb{P}}_{n,f}(u,v) := \frac{1}{|\widetilde{\Omega}_n|_f} f(\frac{u}{v}), \qquad \text{for any } (u,v)\in\Omega_n,$$

where

$$|\Omega_n|_f := \sum_{(u,v)\in\Omega_n} f(\frac{u}{v}), \qquad |\widetilde{\Omega}_n|_f := \sum_{(u,v)\in\widetilde{\Omega}_n} f(\frac{u}{v})$$

are the total f-weights of the sets $\Omega_n, \widetilde{\Omega}_n$.
To any cost C, defined on Ω (or $\widetilde{\Omega}$), we associate Dirichlet series

$$F_C(s) = \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}}\, C(u,v)\, f(\frac{u}{v}), \qquad \widetilde{F}_C(s) = \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{v^{2s}}\, C(u,v)\, f(\frac{u}{v}),$$

whose alternative expressions are

$$F_C(s) = \sum_{v\geq 1} \frac{c_v}{v^{2s}}, \qquad \widetilde{F}_C(s) = \sum_{v\geq 1} \frac{\widetilde{c}_v}{v^{2s}},$$

where c_v, $\widetilde{c}_v$ denote the cumulative costs of C on $\omega_v := \{(u,v)\in\Omega\}$, $\widetilde{\omega}_v := \{(u,v)\in\widetilde{\Omega}\}$, namely,

$$c_v = \sum_{(u,v)\in\omega_v} C(u,v)\, f(\frac{u}{v}), \qquad \widetilde{c}_v = \sum_{(u,v)\in\widetilde{\omega}_v} C(u,v)\, f(\frac{u}{v}).$$

For the trivial cost ($C\equiv 1$), the corresponding cumulative costs a_v or $\widetilde{a}_v$ are just the f-weights of subsets $\omega_v, \widetilde{\omega}_v$, namely

$$a_v = \sum_{(u,v)\in\omega_v} f(\frac{u}{v}), \qquad \widetilde{a}_v = \sum_{(u,v)\in\widetilde{\omega}_v} f(\frac{u}{v}).$$

The mean values of the cost C on $\Omega_n, \widetilde{\Omega}_n$ are then given by the ratio of partial sums,

$$\mathbb{E}_{n,f}[C] = \frac{\sum_{\ell(v)=n} c_v}{\sum_{\ell(v)=n} a_v}, \qquad \widetilde{\mathbb{E}}_{n,f}[C] = \frac{\sum_{\ell(v)=n} \widetilde{c}_v}{\sum_{\ell(v)=n} \widetilde{a}_v}. \tag{25}$$

We are mainly interested by some particular costs C.

5.2. **Costs of interest.** We now describe the main costs that intervene in this paper, defined on the set Ω of all the possible inputs. For each Theorem, we consider two costs, the deterministic cost that we wish to study and the probabilstic cost (underlined) that we succeed to study. For Theorem 1, we consider the costs $C_1 := P_\delta, \underline{C}_1 = \underline{P}_\delta$ for $\delta\in[0,1]$, defined by the relation (10). This means that

$$P_\delta(u,v) = k \qquad \text{iff} \qquad \lg u_k \leq (1-\delta)\ell(u_0) < \lg u_{k-1}.$$

For Theorem 2, we consider the cost C_2 (which depends on the interval J),

$$C_2 = [\![x_{\langle\delta\rangle}\in J]\!] \qquad \text{with} \quad x_{\langle\delta\rangle} := \frac{u_{k+1}}{u_k} \qquad \text{for} \quad k = P_\delta, \qquad \text{and} \quad \underline{C}_2 := [\![\underline{x}_{\langle\delta\rangle}\in J]\!]$$

Finally, for Theorem 4, we consider the cost C_4 (which depends on the interval J),

$$C_4(u,v) = Q_\delta(u,v) = \sum_{i=P_\delta(u,v)-2}^{P_\delta(u,v)} \ell(q_i), \qquad \text{and} \quad \underline{C}_4(u,v) := \sum_{i=\underline{P}_\delta(u,v)-2}^{\underline{P}_\delta(u,v)} \ell(q_i)$$

and, for Theorem 5, the costs $C_5 = \ell(u_{\langle\delta\rangle}$, $\underline{C}_5 := \ell(\underline{u}_{\langle\delta\rangle})$.

We first provide alternative expressions for Dirichlet series $\widetilde{F}_C(s)$, as a function of the transfer operator $\mathbf{H}_s$ relative to the Euclidean dynamical system. We first recall some basic facts about dynamical systems and transfer operators.

5.3. **The Euclidean Dynamical system.** When computing the gcd of the integer-pair (u,v), Euclid's algorithm performs a sequence of divisions. A division $v = uq + r$ replaces the pair (u,v) with the new pair (r,u). If we consider now rationals instead of integer pairs, there exists a map T which replaces the (old) rational u/v by the (new) rational r/u, defined as

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, \qquad T(0) = 0.$$

When extended to the real interval $I = [0,1]$, the pair (I,T) defines the dynamical system relative to Euclid algorithm. We denote by $\mathcal{H}$ the set of the inverse branches of T,

$$\mathcal{H} = \{h_{[q]} : x \to \frac{1}{q+x}; \quad q \geq 1\},$$

and by $\mathcal{H}^p$ the set of inverse branches of depth p (i.e., the set of inverse branches of T^p), namely $\mathcal{H}^p = \{h = h_1 \circ \cdots \circ h_p; h_i \in \mathcal{H}, \forall i\}$. The set $\mathcal{H}^\star := \cup_p \mathcal{H}^p$ is the set of all the possible inverse branches of any depth. Then, the sequence (5) builds a continued fraction

$$\frac{u}{v} = h(0) \quad \text{with} \quad h = h_1 \circ h_2 \circ \ldots \circ h_p \in \mathcal{H}^p. \tag{26}$$

One then associates to each execution of the algorithm a unique LFT $h \in \mathcal{H}^\star$ whose depth is exactly the number p of divisions performed. Remark that the i-th LFT h_i used by the algorithm is exactly the LFT relative to matrix $\mathcal{Q}_i$ of Section 2.1, so that the LFT $h_1 \circ h_2 \circ \ldots \circ h_i$ is relative to matrix $\mathcal{M}_{(i)}$ of Section 2.1. Then, the *CF*–expansion (26) of u/v, when splitted at depth i, creates two LFT's $b_i := h_1 \circ h_2 \circ \ldots \circ h_{i-1}$ and $e_i := h_i \circ \ldots \circ h_p$, defining each a rational number: the "beginning" rational $b_i(0)$, and the "ending" rational $e_i(0)$. The "ending" rational $e_i(0)$ can be expressed with the remainder sequence (u_i)

$$e_i(0) := h_{i+1} \circ h_{i+2} \circ \cdots \circ h_p(0) = \frac{u_{i+1}}{u_i},$$

while the "beginning" rational $b_i(0)$ can be expressed with the twosequences $(p_i), (r_i)$ related to coefficients of matrix $\mathcal{M}_{(i)}$ defined in (6),

$$b_i(0) := h_1 \circ h_2 \circ \cdots \circ h_{i-1}(0) = \frac{|p_i|}{|r_i|}.$$

The main parameters of interest of the Euclid Algorithm involve the denominators sequences u_i, r_i, which are called the continuants. The continuants are closely related to derivatives of LFT's, as we now explain. For any LFT h, the derivative $h'(x)$ can be expressed with the denominator function D: If the function D is defined by

$$D[g](x) = cx + d, \qquad \text{for } g(x) = \frac{ax+b}{cx+d} \quad \text{with} \quad \gcd(a,b,c,d) = 1,$$

then

$$h'(x) = \frac{\det h}{D[h](x)^2}. \tag{27}$$

Finally, since any LFT $h \in \mathcal{H}^\star$ has a determinant of absolute value equal to 1, one has:

$$u_i = |b_i'(0)|^{-1/2}, \qquad r_i = |e_i'(0)|^{-1/2}. \tag{28}$$

5.4. **Transfer operators.** One of the main tool in dynamical systems theory is the transfer operator [23], denoted by $\mathbf{H}_s$. It generalizes the density transformer $\mathbf{H}$ that describes the evolution of the density: if $f = f_0$ denotes the initial density on I, and f_1 the density on $\mathcal{I}$ after one iteration of T, then f_1 can be written as $f_1 = \mathbf{H}[f_0]$, where $\mathbf{H}$ is defined by

$$\mathbf{H}[f](x) = \sum_{h\in\mathcal{H}} |h'(x)|\, f\circ h(x). \tag{29}$$

It is useful to introduce a more general operator that depends on a complex parameter s,

$$\mathbf{H}_s[f](x) = \sum_{h\in\mathcal{H}} |h'(x)|^s\, f\circ h(x) = \sum_{m\geq 1} \frac{1}{(m+x)^{2s}} f\left(\frac{1}{m+x}\right),$$

and multiplicative properties of derivatives entail that

$$\mathbf{H}_s^p[f](x) = \sum_{h\in\mathcal{H}^p} |h'(x)|^s\, f\circ h(x), \qquad (I-\mathbf{H}_s)^{-1}[f](x) = \sum_{h\in\mathcal{H}^\star} |h'(x)|^s\, f\circ h(x).$$

Now, relation (27) between the denominator and the derivative of a LFT, and the fact that any element of $\mathcal{H}^\star$ has a determinant equal to ± 1, entail an alternative expression for the transfer operator,

$$\mathbf{H}_s^p[f](x) = \sum_{h\in\mathcal{H}^p} \frac{1}{D[h](x)^{2s}}\, f\circ h(x), \quad (I-\mathbf{H}_s)^{-1}[f](x) = \sum_{h\in\mathcal{H}^n} \frac{1}{D[h](x)^{2s}}\, f\circ h(x),$$

which will show, with (28) that the transfer operator can be viewed as a generating operator for denominator sequences u_i, r_i. This is the main idea on which is based the dynamical analyses. We now explain the relation between Dirichlet series and transfer operators.

5.5. **The Dirichlet series $F_C(s)$.** We describe alternative expression of the Dirichlet series $\widetilde{F}_C(s), \widetilde{F}_C(s)$, as a function of operator $\mathbf{H}_s$. Let us begin with the trivial cost:

Cost $C_0 \equiv 1$. The Euclid algorithm writes each rational $u/v \in \widetilde{\Omega}$ in a unique way as $u/v = h(0)$ with $h \in \mathcal{H}^\star$. Then,

$$\widetilde{F}_0(2s) := \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{v^{2s}}\, f(\frac{u}{v}) = \sum_{k\geq 0} \sum_{h\in\mathcal{H}^k} |h'(0)|^s \cdot f\circ h(0) = (I-\mathbf{H}_s)^{-1}[f](0),$$

from which we deduce an alternative expression of $F_0(2s)$, with the help of the Riemann ζ function:

$$F_0(2s) = \sum_{d\geq 1} \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{(dv)^{2s}} f\left(\frac{du}{dv}\right) = \zeta(2s)\, \widetilde{F}_0(2s) = \zeta(2s)\ (I-\mathbf{H}_s)^{-1}[f](0).$$

All the studies of the paper are based on refinements of the (simple) equality.

Cost $C_1, \underline{C}_1$ for Theorem 1. We will show in Section 6.4 that a main tool for studying the second cost P_δ on Ω, via its moment generating function $\mathbb{E}_{n,f}[\exp(wP_\delta)]$, is the Dirichlet series $G(2s, 2t, w)$ which depends on three parameters s, t, w and is equal to

$$G(2s,2t,w) = e^w \zeta(2s+2t)\, (I-\mathbf{H}_{s+t})^{-1} \circ (\mathbf{H}_s - \mathbf{H}_{s+t}) \circ (I - e^w \mathbf{H}_s)^{-1}[f](0).$$

Cost $C_2, \underline{C}_2$ for Theorem 2. We will show in Section 6.1 that a main tool for studying the distribution of $x_{\langle\delta\rangle}$ on Ω (via the estimate of $\mathbb{P}_{n,f}[x_{\langle\delta\rangle} \in J]$) is the Dirichlet series which depends on two parameters s, t, together with the interval J,

$$F(2s,2t,J) = \zeta(2s+2t)\, (I-\mathbf{H}_{s+t})^{-1} \left[1_J \cdot (\mathbf{H}_s - \mathbf{H}_{s+t}) \circ (I-\mathbf{H}_s)^{-1}[f]\right](0).$$

Cost $C_4, \underline{C}_4$ for Theorem 4. We will show in Section 6.7 that a main tool for studying the mean value of Q_δ is the Dirichlet series which depends on two parameters s, t,

$$\zeta(2s+2t)(I-\mathbf{H}_{s+t})^{-1}(\mathbf{H}_s - \mathbf{H}_{s+t}) \circ \mathbf{H}_{s,[\ell]}^3 \circ (I-\mathbf{H}_s)^{-1}[f](0),$$

and involves the weighted transfer operator $\mathbf{H}_{s,[\ell]}$ relative to the binary size ℓ and defined as

$$\mathbf{H}_{s,[\ell]}[f](x) := \sum_{m\geq 1} \frac{\ell(m)}{(m+x)^{2s}} f\left(\frac{1}{m+x}\right). \tag{30}$$

Cost $C_5, \underline{C}_5$ for Theorem 5. We will show in Section 6.8 that a main tool for studying the mean value of $\ell(u_{\langle\delta\rangle})$ is the Dirichlet series which depends on three parameters s, t,

$$\zeta(2s+2t)(I-\mathbf{H}_{s+t})^{-1}\circ\mathbf{H}'_{s+t}\circ(I-\mathbf{H}_{s+t})^{-1}\circ\mathbf{H}'_{s+t}\circ(I-\mathbf{H}_s)^{-1}[f](0),$$

and involves the operator $\mathbf{H}'_s := d/(ds)\mathbf{H}_s$ defined as

$$\mathbf{H}'_s[f](x) := -2\sum_{m\geq 1}\frac{\log(m+x)}{(m+x)^{2s}}f\left(\frac{1}{m+x}\right). \tag{31}$$

With alternative expressions of these Dirichlet series at hand, we now perform the second step: we find the dominant singularities of these Dirichlet series and their nature, and then transfer this information for obtaining asymptotic expressions of their coefficients. The expressions obtained in prove that the singularities of the Dirichlet series will be related to the dominant spectral objects of the transfer operator $\mathbf{H}_s$. A precise study of these spectral properties will lead to the asymptotic study of the coefficients of these Dirichlet series.

5.6. **Spectral properties of the transfer operator $\mathbf{H}_s$.** We now recall the main properties of the transfer operator $\mathbf{H}_s$ and its quasi-inverse $(I-\mathbf{H}_s)^{-1}$. These properties depend on the Banach space where the operator acts. Here, the Banach space is $\mathcal{C}^1(\mathcal{I})$, and we recall now the main properties of the operator $\mathbf{H}_s$, when acting on this functional space.

For $\Re(s) > 1/2$, the operator $\mathbf{H}_s$ acts on $\mathcal{C}^1(\mathcal{I})$ and the map $s \to \mathbf{H}_s$ is analytic. For $s = 1$, the operator is quasi–compact: there exists a spectral gap between the unique dominant eigenvalue (that equals 1, since the operator is a density transformer) and the remainder of the spectrum. By perturbation theory, these facts —existence of a dominant eigenvalue $\lambda(s)$ and of a spectral gap— remain true in a complex neighborhood $\mathcal{V}$ of $s = 1$. There, the operator splits into two parts: the part relative to the dominant eigensubspace, denoted $\mathbf{P}_s$, and the part relative to the remainder of the spectrum, denoted $\mathbf{N}_s$, whose spectral radius is strictly less than $\eta|\lambda(s)|$(with $\eta < 1$). This leads to the following spectral decomposition

$$\mathbf{H}_s[f](x) = \lambda(s)\mathbf{P}_s[f](x) + \mathbf{N}_s[f](x),$$

which extends to the powers $\mathbf{H}_s^n$ of the operator

$$\mathbf{H}_s^n[f](x) = \lambda^n(s)\mathbf{P}_s[f](x) + \mathbf{N}_s^n[f](x), \tag{32}$$

and finally to the quasi-inverse $(\mathbf{I}-\mathbf{H}_s)^{-1}$

$$(I-\mathbf{H}_s)^{-1}[f](x) = \frac{\lambda(s)}{1-\lambda(s)}\mathbf{P}_s[f](x) + (\mathbf{I}-\mathbf{N}_s)^{-1}[f](x). \tag{33}$$

The first term on the right admits a pole (of order 1) at $s = 1$, while the second term is analytic on the half–plane $\{\Re(s) > 1\}$. The dominant eigenvalue $\lambda(s)$ is analytic in a neighborhood of $s = 1$, and the pressure function $\Lambda(s) := \log\lambda(s)$ plays an important rôle. In particular, near $s = 1$, one has

$$(I-\mathbf{H}_s)^{-1}[f](x) \sim \frac{-1}{\lambda'(1)}\varphi(x)\int_I f(t)dt, \tag{34}$$

where $-\lambda'(1)$ is the entropy of the system, equal to $\pi^2/(6\log 2)$ and φ is the Gauss density, already mentioned in (1).

For Theorem 2, the Dirichlet series $(1/t)F(2s, 2t, J)$ defined in Section 5.5 can be viewed as a perturbation of

$$F_1(2s, J) := -(I-\mathbf{H}_s)^{-1}[\mathbf{1}_J\cdot\mathbf{H}'_s\circ(I-\mathbf{H}_s)^{-1}[f]](0),$$

for small t. This Dirichlet series $F_1(2s, J)$ involves the operator $\mathbf{H}'_s := (d/ds)\mathbf{H}_s$, has a pôle of order 2 at $s = 1$, and satisfies for s close to 1, with (34)

$$F_1(2s, J) \sim \frac{-1}{(s-1)^2}\left(\frac{1}{\lambda'(1)}\right)^2\varphi(0)\left(\int_J \mathbf{H}'[\varphi](t)dt\right),$$

where $\mathbf{H}' := \mathbf{H}'_1$ and φ is the Gauss density defined in (1). This explains why $\psi = \mathbf{H}'[\varphi]$ introduced in (2) plays a central rôle in our analyses.

5.7. ***US* Property for the Dirichlet series $F_C(s)$.** We have obtained a first information about the singularities of the quasi-inverse $(I-\mathbf{H}_s)^{-1}$ and an alternative expression of $\widetilde{F}_C(s)$ as a function of this quasi-inverse. We now wish to perform the second step and transfer this information for obtaining asymptotic expressions of the coefficients of the Dirichlet series. As a main tool, we rely on convenient "extractors" which express coefficients of series as a function of the series itself. There exist an easy "extractor" for Dirichlet series: the (plain) Tauberian Theorems. However, they do not provide remainder terms, and they are not adapted for our study, since we wish to obtain uniform estimates with respect to auxiliary parameters δ, w, t, J. We then adopt the Perron Formula, which may provide remainder terms, as soon as we have a precise knowledge of $\widetilde{F}_C(s)$ on vertical strips.

The Perron Formula of order two (see [10]) is valid for a Dirichlet series $F(s) = \sum_{n\geq 1} a_n n^{-2s}$ and a vertical line $\Re s = D > 0$ inside the convergence domain of F,

$$\Psi(T) := \sum_{n\leq T} a_n(T-n) = \frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} F(s)\frac{T^{2s+1}}{s(2s+1)}\,ds\,. \tag{35}$$

It is next natural to modify the integration contour $\Re s = D$ into a contour which contains a unique pole of $F(s)$, and it is thus useful to know that the Property *US* [Uniform Estimates on Strips] holds. We have already described this Property in an informal way in Section 1.3. It is now necessary to describe it more precisely.

Theorem A. [US Property for the Euclidean Dynamical System] [Dolgopyat, Baladi, Vallée] [9, 2] *When the transfer operator $\mathbf{H}_s$ relative to the Euclidean dynamical system acts on the functional space $\mathcal{C}^1(\mathcal{I})$ of functions with a continuous derivative on the unit interval $\mathcal{I} := [0,1]$, there exists $\alpha > 0$ for which the following holds on the strip $\mathcal{S} := \{s, 1-\alpha \leq \Re s \leq 1\}$.*

- *(i) The quasi-inverse $(I-\mathbf{H}_s)^{-1}$ has a unique pôle in the vertical strip $\mathcal{S} := \{s, |\Re s - 1| \leq \alpha\}$, located at $s = 1$.*
- *(ii) There exist $t_0 > 0, \xi < 1/5, C > 0$, such that, on the truncated strip $\{s, |\Re s - 1| \leq \alpha, |\Im s| \geq t_0\}$, letting $t := \Im s$,*

$$||(I-\mathbf{H}_s)^{-1}||_{1,t} = O\left(|\Im s|^{\xi}\right) \qquad \textit{with} \quad ||f||_{1,t} := \sup|f| + (1/t)\sup|f'|.$$

From works of Dolgopyat [9] and Baladi-Vallée [2], we know that $(I-\mathbf{H}_s)^{-1}$ satisfies the *US* Property, with a strip of width $\alpha > 0$. With this *US*–Property, we can shift the integration contour in (35). If, for instance

$$F(s) = (1-\mathbf{H}_{s+t})^{-1}[g](0) := \sum_{n\geq 1}\frac{a_n(t)}{n^{2s}},$$

we obtain

$$\Psi(T) := \sum_{n\leq T} a_n(T-n) = \mathrm{Res}_{s=1-t}\left(\frac{T^{2s+1}}{s(2s+1)}F(s)\right) + \frac{1}{2i\pi}\int_{\Re s = 1-t-\alpha} F(s)\frac{T^{2s+1}}{s(2s+1)}\,ds\,.$$

Finally, if the pole is simple, the residue is not zero, and the following estimate shows the importance of the parameter σ, defined as a lower bound for this width α, since it intervenes in the remainder term, as

$$\Psi(T) = \frac{T^{3-2t}}{(1-t)(3-2t)}\mathrm{Res}_{s=1-t}F(s)\left[1+O(T^{-2\sigma})\right]. \tag{36}$$

The real σ mentioned in all our Theorems 1–7 is a lower bound for this width α.

6. Proofs of Theorems 1 and 2

Here, we provide the complete proofs of Theorems 1 and 2. We first recall some notations. On an input (u, v), the Euclid algorithm builds a sequence of remainders (u_k) and a sequence of rationals $x_k = u_{k+1}/u_k$.

We recall that $P_\delta(u,v)$ is the smallest integer k for which $\lg u_k$ is less than $(1-\delta)\ell(u_0)$. We are interested in describing the position of the rational

$$x_{\langle\delta\rangle} := x_k \qquad \text{when} \quad P_\delta(u,v) = k.$$

6.1. **Proof of Theorem 2 – Step 1. The Dirichlet series of interest.** We here provide an estimate of the distribution of the rational $\underline{x}_{\langle\delta\rangle}$, which is a probabilistic version of the rational $x_{\langle\delta\rangle}$, as we shall now explain.

We first deal with intermediate sets $\mathcal{V}^{(k)}_{N,M}(J)$, $\mathcal{U}^{(k)}_{N,M}(J)$, defined as

$$\mathcal{V}^{(k)}_{N,M}(J) := \{(u,v) \in \Omega;\quad v = N,\, u_{k+1} = M,\, x_{k+1} \in J\},$$

$$\mathcal{U}^{(k)}_{N,M}(J) := \{(u,v) \in \Omega,\quad v = N,\, u_k = M,\, x_{k+1} \in J\},$$

and the set

$$\mathcal{A}_N(W,J) := \sum_{k\geq 0} \left[\left(\sum_{M\leq W} \mathcal{V}^{(k)}_{N,M}(J) \right) \setminus \left(\sum_{M\leq W} \mathcal{U}^{(k)}_{N,M}(J) \right) \right]. \tag{37}$$

gathers the pairs (u,v) of Ω with $v = N$ for which the following is true: "if k denotes the smallest index for which the remainder u_k has a denominator less than W, the rational x_k belongs to J ". This shows that these intermediate sets will be closely related to our problem.

We now observe two facts: The f-weights $\widetilde{u}^{(k)}_{N,M}(J), \widetilde{v}^{(k)}_{N,M}(J)$ of the tilded version of the intermediate sets

$$\widetilde{\mathcal{V}}^{(k)}_{N,M}(J) := \mathcal{V}^{(k)}_{N,M}(J) \cap \widetilde{\Omega}, \qquad \widetilde{\mathcal{U}}^{(k)}_{N,M}(J) := \mathcal{U}^{(k)}_{N,M}(J) \cap \widetilde{\Omega}$$

are easily generated by the transfer operator, since the two following equalities hold

$$\widetilde{U}(2s,2t,J,k) := \sum_{N\geq 1}\sum_{M\geq 1} \frac{\widetilde{u}^{(k)}_{N,M}(J)}{N^{2s}M^{2t}} = (I - \mathbf{H}_{s+t})^{-1}\left[1_J \cdot \mathbf{H}_{s+t} \circ \mathbf{H}_s^k[f]\right](0), \tag{38}$$

$$\widetilde{V}(2s,2t,J,k) := \sum_{N\geq 1}\sum_{M\geq 1} \frac{\widetilde{v}^{(k)}_{N,M}(J)}{N^{2s}M^{2t}} = (I - \mathbf{H}_{s+t})^{-1}\left[1_J \cdot \mathbf{H}_s^{k+1}[f]\right](0). \tag{39}$$

On the other hand, there are nice relations between $\mathcal{V}^{(k)}_{N,M}(J)$, $\mathcal{U}^{(k)}_{N,M}(J)$ and their tilded versions, as we now explain. Each of these two sets $\mathcal{V}^{(k)}_{N,M}(J)$, $\mathcal{U}^{(k)}_{N,M}(J)$ decomposes as a disjoint union

$$\mathcal{V}^{(k)}_{N,M}(J) = \bigcup_{d\geq 1} \left(\mathcal{V}^{(k)}_{N,M}(J) \cap \Omega_{[d]}\right), \qquad \mathcal{U}^{(k)}_{N,M}(J) = \bigcup_{d\geq 1} \left(\mathcal{U}^{(k)}_{N,M}(J) \cap \Omega_{[d]}\right),$$

which involves the set $\Omega_{[d]}$ of pairs (u,v) of Ω for which $\gcd(u,v) = d$; the map $(u,v) \mapsto (du,dv)$ defines two bijections which preserve the f-weights,

first from $\widetilde{\mathcal{V}}^{(k)}_{N,M}(J)$ onto $\left(\mathcal{V}^{(k)}_{dN,dM}(J) \cap \Omega_{[d]}\right)$, second from $\widetilde{\mathcal{U}}^{(k)}_{N,M}(J)$ onto $\left(\mathcal{U}^{(k)}_{dN,dM}(J) \cap \Omega_{[d]}\right)$.

Then, the Dirichlet series U, V and their tilded versions $\widetilde{U}, \widetilde{V}$ are related via the Riemann ζ function, as follows:

$$U(s,t,J,k) := \sum_{N\geq 1}\sum_{M\geq 1} \frac{u^{(k)}_{N,M}(J)}{N^s M^t} = \zeta(s+t)\, \widetilde{U}(s,t,J,k), \tag{40}$$

$$V(s,t,J,k) := \sum_{N\geq 1}\sum_{M\geq 1} \frac{v^{(k)}_{N,M}(J)}{N^s M^t} = \zeta(s+t)\, \widetilde{V}(s,t,J,k). \tag{41}$$

Finally, the series $F(s,t,J)$ defined as

$$F(s,t,J) := \sum_{k\geq 0} [V(s,t,J,k) - U(s,t,J,k)] \tag{42}$$

admits with (40, 41, 38, 39) the alternative expression which involves the ζ function and the transfer operator $\mathbf{H}_s$

$$F(s,t,J) := \zeta(2s+2t)(I - \mathbf{H}_{s+t})^{-1}\left[1_J \cdot (\mathbf{H}_s - \mathbf{H}_{s+t}) \circ (I - \mathbf{H}_s)^{-1}[f]\right](0). \tag{43}$$

On the other hand, $F(s,t,J)$ is a Dirichlet series of the form

$$F(s,t,J) = \sum_{N\geq 1}\sum_{M\geq 1}\frac{a_{N,M}(J)}{N^s M^t}$$

whose coefficient $a_{N,M}(J)$ satisfies the following, with the definition of F given in (42),

$$\sum_{M\leq W} a_{N,M}(J) = \sum_{M\leq W}\sum_{k\geq 0}\left(v_{N,M}^{(k)}(J) - u_{N,M}^{(k)}(J)\right) = \sum_{k\geq 0}\left[\left(\sum_{M\leq W} v_{N,M}^{(k)}(J)\right) - \left(\sum_{M\leq W} u_{N,M}^{(k)}(J)\right)\right],$$

and the last expression is exactly the f-weight of the set $\mathcal{A}_N(W,J)$ defined in (37). Finally, the equality

$$\sum_{N=2^{n-1}}^{2^n} \mathcal{A}_N(2^{(1-\delta)n}, J) = \{(u,v)\in\Omega_n;\quad x_{\langle\delta\rangle}\in J\}$$

holds and entails the equality

$$\mathbb{P}_{n,f}[x_{\langle\delta\rangle}\in J] = \frac{1}{|\Omega_n|_f}\sum_{N=2^{n-1}}^{2^n-1}\sum_{M\leq 2^{(1-\delta)n}} a_{N,M}(J)$$

where $|\Omega_n|_f$ is just the f-weight of Ω_n. Comparing the Riemann sum to the integral entails

$$|\Omega_n|_f := \sum_{(u,v)\in\Omega_n} f(\frac{u}{v}) = \sum_{v=2^{n-1}}^{2^n-1}\sum_{u<v} f(u/v) = |\Omega_n|\left[1 + 2^{-n}O(||f||_1)\right].$$

We have finally to evaluate

$$\frac{1}{|\Omega_n|}\sum_{N=2^{n-1}}^{2^n-1}\sum_{M\leq 2^{(1-\delta)n}} a_{N,M}(J)$$

It is then sufficient to extract coefficients from the Dirichlet series $F(s,t,J)$ given in (43). However, it is not possible to directly deal with the characteristic function of the interval J, since it does not belong to the "convenient" functional space $\mathcal{C}^1(I)$ where the Property US holds. Then, for a function ε positive which satisfies $\varepsilon(x)\leq x$, we replace the function $\mathbf{1}_J$ by two functions $\psi^+_{(J,\varepsilon)}$ and $\psi^-_{(J,\varepsilon)}$ of $\mathcal{C}^1(I)$ which are good approximations of $\mathbf{1}_J$, and satisfy

$$\psi^-_{(J,\varepsilon)} \leq \mathbf{1}_J \leq \psi^+_{(J,\varepsilon)}, \qquad ||\psi^+_{(J,\varepsilon)} - \psi^-_{(J,\varepsilon)}||_{1,1} \leq \frac{1}{\varepsilon(|J|)}, \qquad \int_I |\psi^+_{(J,\varepsilon)} - \psi^-_{(J,\varepsilon)}|(u)du \leq \varepsilon(|J|). \tag{44}$$

We replace the Dirichlet series $F(s,t,J)$ by the series $F_+(s,t,J,\varepsilon), F_-(s,t,J,\varepsilon)$ defined as

$$F_\pm(2s,2t,J,\varepsilon) = \zeta(2s+2t)(I-\mathbf{H}_{s+t})^{-1}\left[\psi^\pm_{(J,\varepsilon)}\cdot(\mathbf{H}_s - \mathbf{H}_{s+t})\circ(I-\mathbf{H}_s)^{-1}[f]\right](0). \tag{45}$$

The coefficients of these series, denoted by $a^\pm_{N,M}(J,\varepsilon)$, have the following combinatorial sense : The sum of these coefficients

$$\sum_{M\leq W} a^\pm_{N,M}(J,\varepsilon) \tag{46}$$

equals the sum, taken over all pairs (u,v) with $v=N$, of the quantities $f(x_k)\cdot\psi^\pm_{(J,\varepsilon)}(x_k)$, where x_k is the rational relative to the smallest index k for which u_k is less than W. Then, the inequalities

$$\sum_{M\leq W} a^-_{N,M}(J,\varepsilon) \leq \sum_{M\leq W} a_{N,M}(J) \leq \sum_{M\leq W} a^+_{N,M}(J,\varepsilon) \tag{47}$$

hold, and show that it is sufficient to deal with the series $F_\pm(s,t,J,\varepsilon)$, denoted in the following by $F(s,t,J,\varepsilon)$.

6.2. **Proof of Theorem 2 – Step 2. Extraction via the Perron Formula.** The series F defined in (45) depends of two complex variables s and t (with J and ε as parameters). We will use the Perron Formula, two times.

First suppose that the complex s is fixed, satisfies $\Re s > 1$ and consider the Dirichlet series F as a function of t, which has an only pôle at $t = 1 - s$ in the strip $1 - \alpha < \Re(s+t) < 1 + \alpha$. Then, with the Perron formula,

$$\sum_{W_1 \le W} \sum_{M \le W_1} \sum_{N \le 1} \frac{a^{\pm}_{N,M}(J,\varepsilon)}{N^{2s}} =$$

$$= \zeta(2) \frac{W^{2(1-s)+1}}{(3-2s)} \frac{\varphi(0)}{\lambda'(1)} \int_I \psi_{(J,\varepsilon)}(u) \cdot \left[\left(\frac{\mathbf{H}_s - \mathbf{H}_1}{s-1} \right) \circ (1 - \mathbf{H}_s)^{-1}[f] \right](u) du$$

$$+ \frac{1}{2i\pi} \int_{\Re(s+t)=1-\alpha} \zeta(2s+2t) \frac{W^{2t+1}}{t(2t+1)} F(2s, 2t, J, \varepsilon) dt.$$

This is now a Dirichlet series with respect to s, which has an only pôle at $s = 1$ in the strip $1 - \beta < \Re s < 1 + \beta$, and using again the Perron Formula for extracting coefficients, we obtain finally four terms for the sum of coefficients

$$E_1^{\pm}(T, W, \varepsilon) := \sum_{T_1 \le T} \sum_{N \le T_1} \sum_{W_1 \le W} \sum_{M \le W_1} a^{\pm}_{N,M}(J,\varepsilon),$$

namely

$$-\zeta(2) \frac{\varphi(0)}{\lambda'(1)^2} \frac{T^3}{3} W \int_I \psi^{\pm}_{(J,\varepsilon)}(u) \mathbf{H}'[\varphi](u) du$$

$$+ \frac{1}{2i\pi} \frac{T^3}{3} \int_{\Re t = -\alpha} \zeta(2+2t) \frac{W^{2t+1}}{t(2t+1)} (I - \mathbf{H}_{1+t})^{-1} \left[\psi^{\pm}_{(J,\varepsilon)} \cdot (\mathbf{H}_1 - \mathbf{H}_{1+t}) \left[\frac{\varphi}{-\lambda'(1)} \right] \right](0)$$

$$- \frac{\zeta(2)\varphi(0)}{2i\pi\lambda'(1)} \int_{\Re s = 1-\beta} \frac{T^{2s+1}}{s(2s+1)} \frac{W^{2(1-s)+1}}{(3-2s)} \left(\int_I \psi^{\pm}_{(J,\varepsilon)}(u) \left(\frac{\mathbf{H}_s - \mathbf{H}_1}{1-s} \right) \circ (I - \mathbf{H}_s)^{-1}[f](u) du \right) ds$$

$$- \int_{\substack{\Re t = \beta - \alpha \\ \Re s = 1 - \beta}} \frac{\zeta(2s+2t)}{4\pi^2} \frac{T^{2s+1}}{s(2s+1)} \frac{W^{2t+1}}{t(2t+1)} (I - \mathbf{H}_{s+t})^{-1} \left[\psi^{\pm}_{(J,\varepsilon)} \cdot (\mathbf{H}_s - \mathbf{H}_{s+t}) \circ (I - \mathbf{H}_s)^{-1}[f] \right](0) ds dt.$$

If we choose $\alpha = \beta$, it seems that the fourth term has a pôle at $t = 0$, but this is not a "true" pôle, since there is an occurrence of a secant operator, of the form $(1/t)(\mathbf{H}_{s+t} - \mathbf{H}_s)$ which tends to the operator $\mathbf{H}'_s$ when $t \to 0$. We then choose $\alpha = \beta$, and, for reasons which will appear later, due in particular to possible applications of Proposition A, we choose $\alpha = \beta = \underline{\sigma} := \min(\sigma, 1/2)$

The first term will provide the main term, which is $\Theta(T^3 W)$, more precisely

$$E_1^{\pm}(T, W) = a(J) \frac{T^3}{3} W + F_1^{\pm}(T, W) \qquad \text{with} \qquad a(J) = \frac{1}{\lambda'(1)} \int_J \mathbf{H}'[\varphi](t) dt = \int_J \psi(t) dt. \tag{48}$$

(For the computation of the constant $a(J)$, we used the equality $\zeta(2) = -\lambda'(1) \log 2$ which comes from spectral properties at $s = 1$ described in Section 5.6). Then Theorem A entails estimates for the four terms of $F_1^+(T, W, \varepsilon) - F_1^-(T, W, \varepsilon)$, respectively

$$O(1) \qquad O(T^3 W^{1-2\underline{\sigma}}) \qquad O(T^{3-2\underline{\sigma}} W^{2\underline{\sigma}+1}) \qquad O(T^{3-2\underline{\sigma}} W).$$

Here, the constants involved in the O-terms depend only on J and ε, but not in the same way for all the terms: In the first and the third term, the interval J intervenes via the integral of the function $\psi^+_{(J,\varepsilon)} - \psi^-_{(J,\varepsilon)}$, and, with (44), the constants are $O(\varepsilon(|J|))$. In the second and fourth terms, the interval J intervenes via the norm $||\cdot||_{1,1}$ of the function $\psi^+_{(J,\varepsilon)} - \psi^-_{(J,\varepsilon)}$, and, with (44), the constants in the second and the fourth term are $O(1/\varepsilon(|J|))$. Finally

$$F_1^+(T, W) - F_1^-(T, W) = \left[a(J) \frac{T^3}{3} W \right] \; C(J, W, \varepsilon) \left[1 + O\left(\frac{T}{W} \right)^{-2\underline{\sigma}} \right] \tag{49}$$

$$\text{with} \qquad C(J, W, \varepsilon) := O\left(\frac{\varepsilon(|J|)}{|J|} + \frac{1}{|J|\varepsilon(|J|)} W^{-2\underline{\sigma}} \right). \tag{50}$$

We now return to our object of interest, the quadruple sum

$$E_1(T,W) := \sum_{T_1\le T}\sum_{N\le T_1}\sum_{W_1\le W}\sum_{M\le W_1} a_{N,M}(J), \tag{51}$$

for which equations (47,48) and (49) entail the estimate

$$E_1(T,W) = a(J)\frac{T^3}{3}W\left[1+C(J,W,\varepsilon)\right]\left[1+O\left(\frac{T}{W}\right)^{-2\underline{\sigma}}\right] \tag{52}$$

where the function $C(J,W,\varepsilon)$ is defined in (50).

6.3. **Proof of Theorem 2 – Step 3. Final estimates for variable $\underline{x}\langle\delta\rangle$ on Ω.** The sum of coefficients $\sum_{M\le W} a_{N,M}(J)$ is positive, so that

$$T\mapsto B(T,W_1) := \sum_{N\le T}\sum_{M\le W_1} a_{N,M}(J)$$

is increasing. We first consider the corresponding estimates (52) with respect to variable T, each value of the triple (W,J,ε) being fixed. Then, it is possible to transform in $E_1(T,W)$ the double sum over indices N into a simple sum with Proposition B of the Appendix (Section 7) and deduce from the estimate of $E_1(T,W)$ an estimate for the sum

$$E_2(T,W) := \sum_{W_1\le W} B(T,W_1) = a(J)\,T^2W\ \left[1+C(J,W,\varepsilon)\right]\left[1+O\left(\frac{T}{W}\right)^{-\underline{\sigma}}\right].$$

We will be interested in the following by $E(T,W) := E_2(T,W) - E_2(T/2,W)$ for which we get the estimate

$$E(T,W) = \frac{3}{4}\,a(J)\,T^2W\left[1+C(J,W,\varepsilon)\right]\left[1+O\left(\frac{T}{W}\right)^{-\underline{\sigma}}\right]. \tag{53}$$

Applying now Proposition A of Section 7, with the choice $(T-T_-)/T = \Theta((T/W)^{\underline{\sigma}/2})$ (always, for each value of the triple (W,J,ε) fixed) provides the estimate

$$\frac{E(T,W)-E(T_-,W)}{T-T_-} = \frac{3}{2}\,a(J)\,TW\left[1+C(J,W,\varepsilon)\right]\left[1+O\left(\frac{T}{W}\right)^{-\underline{\sigma}/2}\right]. \tag{54}$$

We now consider that T and W are polynomially related, (but J and ε fixed) and we let $T=W^\nu$, with $\nu\ge 1$, and we wish to obtain an estimate of

$$\frac{E(W^\nu,W)-E(W^\nu,W_-)}{W-W_-}$$

First, observe the following decomposition

$$E(W^\nu,W)-E_1(W^\nu,W_-) = \left[E(W^\nu,W)-E(W_-^\nu,W_-)\right]-\left[E(W^\nu,W_-)-E(W_-^\nu,W_-)\right]. \tag{55}$$

Applying Proposition A to the first term, remarking that

$$E(W^\nu,W) = \frac{3}{4}\,a(J)\,W^{2\nu+1}\left[1+C(J,W,\varepsilon)\right]\left[1+O\left(W^{-\delta\nu\underline{\sigma}}\right)\right]$$

and choosing $(W-W_-)/W := \Theta(W^{-\tau/2})$ with $\tau := \min(\delta\nu\underline{\sigma},1)$, gives

$$\frac{E(W^\nu,W)-E(W_-^\nu,W_-)}{W-W_-} = \frac{3}{4}\,(2\nu+1)\,a(J)\,W^{2\nu}\left[1+O\left(C(J,\varepsilon,W)+W^{-\tau/2}\right)\right]. \tag{56}$$

For the second term, we take the same choice for $(W-W_-)/W$, and we remark that, in this case

$$\frac{T-T_-}{T_-} = \frac{W^\nu-W_-^\nu}{W_-^\nu} = \nu\frac{W-W_-}{W}[1+O(W^{-\tau/2})] = \Theta\left(\frac{W-W_-}{W}\right).$$

Using now (54), we obtain

$$\frac{E(W^\nu,W_-)-E(W_-^\nu,W_-)}{W-W_-} = \frac{3}{2}a(J)\,W^{\nu+1}\left[\nu W^{\nu-1}\right]\left[1+O\left(C(J,W,\varepsilon)+W^{-\tau/2}\right)\right]. \tag{57}$$

Finally, using (55, 56, 57) leads to

$$\frac{E(W^\nu, W) - E(W^\nu, W_-)}{W - W_-} = \frac{3}{4} a(J) W^{2\nu} \left[1 + O\left(R(W)\right)\right] \tag{58}$$

with

$$R(W) := \max\left(\frac{\varepsilon(|J|)}{|J|}, \frac{1}{|J|\varepsilon(|J|)} W^{-2\underline{\sigma}}, W^{-\tau/2} \right). \tag{59}$$

We now consider the case when the function ε (which quantifies the approximation of the characteristic function $\mathbf{1}_J$) is a power function, of the form $x \mapsto x^{1+\theta}$, We suppose that all our parameters $X \in \{|J|, \varepsilon(|J|), T, W\}$ have an exponential dependence on n (now J and ε vary), and we fix their exponents $e(X) := n^{-1} \lg X$ as

$$e(T) = 1, \qquad e(W) = (1-\delta) = \frac{1}{\nu}, \qquad e(|J|) = -2\gamma, \qquad e(\varepsilon(|J|)) = -2\gamma(1+\theta).$$

Then, the exponents of the terms in $R(W)$ are all at least equal to

$$\rho := \min\{2\gamma\theta,\ 2\underline{\sigma}(1-\delta) - 2\gamma(\theta+2),\ \delta\underline{\sigma}/2, (1-\delta)/2\}.$$

We first choose the best exponent of the function $\varepsilon : x \mapsto x^{1+\theta}$ in order to equalize the first two terms in the expression of ρ. Since the exponent θ must be strictly positive, this leads to choose $\gamma < (1/2)\underline{\sigma}(1-\delta)$, and finally

$$\rho \geq \rho_0 := \min\{\underline{\sigma}(1-\delta) - 2\gamma, \underline{\sigma}\delta/2\}.$$

[remark that the fourth term in ρ has now "disappeared" due to the inequality $\underline{\sigma} < 1/2$]. Suppose now that the interval J is large enough (with respect to $\underline{\sigma}$, and the fraction $(1-\delta)$), with $e(|J|) = 2\gamma < (1/2)\underline{\sigma}(1-\delta)$. Then, there is a lower bound $\rho(\delta)$ for ρ, which depends only on $\underline{\sigma}$ and δ, with

$$\rho(\delta) := (1/2)\,\underline{\sigma}\, \min(1-\delta, \delta).$$

Observe that $\rho(\delta)$ is always less than $(1/4)(1-\delta)$.

Finally, we return to our initial problem, and with (58, 59), together with the definition of $\rho(\delta)$ and the expression of $a(J)$ in (48), we obtain an estimate for

$$\sum_{N=2^{n-1}}^{2^n-1} \frac{1}{2^{(1-\delta-\rho(\delta))n}} \sum_{W=2^{(1-\delta)n}[1-2^{-\rho(\delta)n}]}^{2^{(1-\delta)n}} \sum_{M \leq W} a_{N,M}(J)$$
$$= \left(\frac{3}{4} 2^{2n}\right) \left(\int_J \psi(t) dt\right) \left[1 + O(2^{-n\rho(\delta)}\right].$$

The first term equals the cardinal of Ω_n, and "disappears" when we return to probabilities.

This not exactly the expression (46) for $M \leq 2^{(1-\delta)n}$, but a smoothed version of it. Then, we do not exactly study the variable $x_{\langle\delta\rangle}$ but a probabilistic variant of this variable that we now recall. Consider the interval $[2^{(1-\delta)n}(1-2^{-\rho(\delta)n}), 2^{(1-\delta)n}]$. Choose an integer W uniformly in this interval. Denote by $\underline{x}_{\langle\delta\rangle}$ the rational x_k associated to the first index k for which $\lg u_k$ is less than W. We have studied this (probabilistic) variable $\underline{x}_{\langle\delta\rangle}$ and evaluate

$$\mathbb{P}_{n,f}[\underline{x}_{\langle\delta\rangle} \in J].$$

Remark that in any interval $]A/2, A]$, there are at most two elements of the sequence x_k, so that, for n large enough, $\underline{x}_{\langle\delta\rangle}$ equals $x_{\langle\delta\rangle+i}$ with $0 \leq i \leq 2$.

6.4. **Proof of Theorem 1– Step 1. The Dirichlet series of interest.** We study, in the same vein as before, a probabilistic version $\underline{P}_\delta$ of P_δ. We prove that it follows an asymptotic gaussian law on Ω, from which it will be easy to deduce an asymptotic gaussian law for the deterministic version P_δ on Ω.

We wish to use the Quasi-Powers Theorem which provides sufficient conditions, which entail an asymptotic gaussian behaviour.

Theorem B. [Quasi-Powers Theorem.] (Hwang) [15] *Assume that the moment generating functions $\mathbb{E}_{n,f}[\exp(wR)]$ for a cost R are analytic in a complex neighborhood $\mathcal{W}$ of $w = 0$, and satisfy*

$$\mathbb{E}_{n,f}[\exp(wR] = \exp[\beta_n C(w) + D(w)] \left(1 + O(\kappa_n^{-1})\right), \tag{60}$$

with β_n, $\kappa_n \to \infty$ as $n \to \infty$, $C(w)$, $D(w)$ analytic on $\mathcal{W}$ and the O–term uniform in $\mathcal{W}$. Then, the mean and the variance satisfy

$$\mathbb{E}_{n,f}[R] = C''(0)\cdot\beta_n + D'(0) + O(\kappa_n^{-1})\,, \qquad \mathbb{V}[R_n] = C''(0)\cdot\beta_n + D''(0) + O(\kappa_n^{-1})\,.$$

Furthermore, if $C''(0) \neq 0$, the distribution of R is asymptotically Gaussian on Ω_n with speed of convergence $O(\kappa_n^{-1} + \beta_n^{-1/2})$,

$$\mathbb{P}_{n,f}\left[x \;\middle|\; \frac{R(x) - C'(0)n}{\sqrt{C''(0)n}} \leq Y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Y} e^{-y^2/2}\,dy + O(\kappa_n^{-1} + \beta_n^{-1/2})\,.$$

We shall show that Theorem B can be applied to our framework, with

$$\beta_n = n, \qquad \kappa_n = 2^{-n\rho(\delta)}, \qquad C(w) = 2\delta\lg(\tau(w) - 1),$$

where $\tau(w)$ is the solution of the equation $\Lambda(s) = -w$ which involves the pressure function $\Lambda(s) := \log\lambda(s)$. This will entail Theorem 1.

We first wish to estimate the generating function $\mathbb{E}_{n,f}(\exp[w\underline{P}_\delta])$, as a quasi-powers. We deal with the function $G(s,t,w) := \zeta(s+t)\,\widetilde{G}(s,t,w)$ with

$$\widetilde{G}(2s,2t,w) = e^w\,(I - \mathbf{H}_{s+t})^{-1}(\mathbf{H}_s - \mathbf{H}_{s+t}) \circ (I - e^w\mathbf{H}_s)^{-1}[f](0).$$

The series G can be written as a Dirichlet series which depends on two variables s, t, together with a parameter w

$$G(s,t,w) = \sum_{k\geq 0} e^{w(k+1)} \sum_{N\geq 1}\sum_{M\geq 1} \frac{b_{N,M}^{(k)}}{N^s M^t} = \sum_{k\geq 0} e^{wk}[V(s,t,I,k) - U(s,t,I,k)], \tag{61}$$

where the functions U and V are defined in (40, 41). Here, the coefficient $\sum_{M\leq W} b_{N,M}^{(k)}$ equals the f-weight of pairs (u,v) with $v = N$ for which u_{k+1} is at most W, while u_k is greater than W. Then, the quantity

$$\sum_{N=2^{n-1}}^{2^n} \sum_{M\leq 2^{(1-\delta)n}} b_{N,M}^{(k)}$$

equals the f-weight of the subset of pairs (u,v) of size n for which P_δ equals $k+1$, and the expression

$$\sum_{N=2^{n-1}}^{2^n} \sum_{M\leq 2^{(1-\delta)n}} \sum_{k\geq 0} e^{w(k+1)} b_{N,M}^{(k)}$$

is the cumulative generating function of parameter P_δ on Ω_n. As previously, it is then sufficient to extract coefficients from the Dirichlet series $G(s,t,w)$.

6.5. **Proof of Theorem 1 – Step 2. Extraction with the Perron Formula.** This series G defined in (61) depends of two complex variables s and t (with w as a parameter). We will use the Perron Formula, two times.
We proceed in two steps, as previously. We first consider the Dirichlet series as a function of t, which has an only pôle at $t = 1-s$ in the vertical strip in the strip $1-\alpha < \Re(s+t) < 1+\alpha$. Then

$$\sum_{W_1\leq W}\sum_{M\leq W_1}\sum_{N\leq 1}\sum_{k\geq 0} e^{tk}\frac{b_{N,M}^{(k)}}{N^{2s}} =$$

$$= \zeta(2)\frac{W^{2(1-s)+1}}{(3-2s)}\frac{\varphi(0)}{\lambda'(1)}\int_I \left(\frac{\mathbf{H}_1 - \mathbf{H}_s}{1-s}\right)\circ(1 - e^w\mathbf{H}_s)^{-1}[f](u)du$$

$$+\frac{1}{2i\pi}\int_{\Re(s+t)=1-\alpha} \zeta(2s+2t)\frac{W^{2t+1}}{t(2t+1)}G(2s,2t,w)dt.$$

This is now a Dirichlet series with respect to s, which has an only pôle at $s=\tau(w)$ in the vertical strip $1-\beta<\Re s<1+\beta$, where w is defined by the relation $w=-\Lambda(s)$. Using again the Perron Formula for extracting coefficients, we obtain finally four terms for this sum of coefficients

$$e^{-w}D(T,W,w):=\sum_{T_1\le T}\sum_{N\le T_1}\sum_{W_1\le W}\sum_{M\le W_1}\sum_{k\ge 0}e^{wk}b_{N,M}^{(k)}, \tag{62}$$

namely

$$\zeta(2)\frac{\nu_{\tau(w)}[f]}{\lambda'(1)\lambda'(\tau(w))}\frac{W^{3-2\tau(w)}}{(3-2\tau(w))}\frac{T^{2\tau(w)+1}}{\tau(w)(2\tau(w)+1)}\int_I\left(\frac{\mathbf{H}_{\tau(w)}-\mathbf{H}_s}{1-\tau(w)}\right)[\varphi_{\tau(w)}](u)du$$

$$+C_w[f]\,T^{2\tau(w)+1}\int_{\Re[\tau(w)+t]=1-\alpha}\frac{W^{2t+1}}{t(2t+1)}(I-\mathbf{H}_{\tau(w)+t})^{-1}\circ\left(\mathbf{H}_{\tau(w)}-\mathbf{H}_{\tau(w)+t}\right)\left[\frac{\varphi_w}{-\lambda'(\tau(w))}\right](0)$$

$$+\frac{\zeta(2)\varphi(0)}{2i\pi\,\lambda'(1)}\int_{\Re s=\tau(w)-\beta}\frac{T^{2s+1}}{s(2s+1)}\frac{W^{2(1-s)+1}}{(3-2s)}\left(\int_I\left(\frac{\mathbf{H}_s-\mathbf{H}_1}{1-s}\right)\circ(I-e^w\mathbf{H}_s)^{-1}[f](u)du\right)ds$$

$$-\int_{\substack{\Re t=1-\tau(w)+\beta-\alpha\\ \Re s=\tau(w)-\beta}}\frac{\zeta(2s+2t)}{4\pi^2}\frac{T^{2s+1}}{s(2s+1)}\frac{W^{2t+1}}{t(2t+1)}(I-\mathbf{H}_{s+t})^{-1}\circ(\mathbf{H}_s-\mathbf{H}_{s+t})\circ(I-e^w\mathbf{H}_s)^{-1}[f](0)dsdt.$$

Here, the first term involves the dominant eigenfunction φ_s of $\mathbf{H}_s$ and the dominant eigenmeasure ν_s of the dual $\mathbf{H}_s^\star$ at $s=\tau(w)$ and the second term involves

$$C_w[f]:=\frac{\nu_{\tau(w)}[f]\,\zeta(2\tau(w)+2)}{2i\pi\tau(w)(2\tau(w)+1)}.$$

We first choose, as in Theorem 2, $\alpha=\beta=\underline{\sigma}$. The first term will provide the main term, which is of the form $A(w)W^{3-2\tau(w)}T^{2\tau(w)+1}$ while Theorem A entails the following estimates for the other three terms : [here, $t(w)$ denotes the real part of $\tau(w)$]

$$O(W^{3-2\rho(w)-2\underline{\sigma}}T^{1+2t(w)}) \qquad O(T^{1+2t(w)-2\underline{\sigma}}W^{3-2t(w)-2\underline{\sigma}}) \qquad O(T^{1+2t(w)-2\underline{\sigma}}W^{3-2t(w)})$$

Here, the constants involved in the O-terms are uniform when w is near 0. Finally

$$D(T,W,w)=R(w)W^{3-2\tau(w)}T^{2\tau(w)+1}\left[1+O(W^{-2\underline{\sigma}})\right]\left[1+O\left(\frac{T}{W}\right)^{-2\underline{\sigma}}\right],$$

where $R(w)$ is analytic and not zero when $w\in\mathcal{W}$.

6.6. **Proof of Theorem 1– Step 3. Final estimates for variable P_δ on Ω.** We now follow the same lines as in the proof of Theorem 2. We first consider W as fixed. For transforming the double sum over indices N into a simple sum, it is more involved than above, since the positivity argument cannot be used here, because w is a complex parameter. However, it is possible to apply the Proposition C' of Appendix, and transform the double sum over indices N into a simple sum. We then deduce from the estimate of $D(T,W,w)$ in (62) an estimate for the sum

$$D_1(T,W,w):=\sum_{N=T/2}^{T}\sum_{W_1\le W}\sum_{M\le W_1}\sum_{k\ge 0}e^{wk}b_{N,M}^{(k)}$$

$$=R_1(w)T^{2(\tau(w)-1)}W^{2(1-\tau(w))}\left[1+O(W^{-2\underline{\sigma}})\right]\left[1+O\left(\frac{T}{W}\right)^{-\underline{\sigma}}\right].$$

Then, as in Section 6.3, we consider that T and W are polynomially related, and use two times Proposition A as in Section 6.3. We obtain an estimate for

$$\frac{1}{|\Omega_n|}\sum_{N=2^{n-1}}^{2^n-1}\frac{1}{2^{(1-\delta-\rho(\delta))n}}\sum_{W=2^{(1-\delta)n}}^{2^{(1-\delta)n}[1-2^{-\rho(\delta)n}]}\sum_{M\le W}\sum_{k\ge 0}e^{wk}b_{N,M}^{(k)}$$

$$=R_2(w)2^{2n\delta(\tau(w)-1)}\left[1+O(2^{-n\rho(\delta)}\right] \qquad \text{with} \quad d(0)=1, \qquad \rho(\delta)=(1/2)\underline{\sigma}\min(1-\delta,\delta).$$

Finally, we obtain an estimate for the moment generating function of the $\rho(\delta)$–probabilistic variant $\underline{P}_\delta$ on Ω_n, namely

$$\mathbb{E}_{n,f}[\exp(w\underline{P}_\delta)]=R_3(w)2^{2n\delta(\tau(w)-1)}\left[1+O(2^{-\rho(\delta)n})\right].$$

Then, the Quasi-Powers theorem, applied with

$$C(w) := 2\delta \lg(\tau(w) - 1), \qquad D(w) := \lg R_3(w)$$

entails an asymptotic gaussian law for the probabilistic variant $\underline{P}_\delta$ on Ω_n. Furthermore, we have already remarked that, in any interval $]A/2, A]$, there are at most two elements of the sequence x_k. Then, for n large enough, the two variables – the probabilistic variable $\underline{P}_\delta$ and its deterministic version P_δ– are closely related since they satisfy $|\underline{P}_\delta - P_\delta| \leq 2$.

Finally, Proposition 1 of the paper of Lhote–Vallée [21], together with the inequality $|\underline{P}_\delta - P_\delta| \leq 2$ proves that the asymptotic gaussian law also holds for P_δ on Ω, with a speed of convergence of order $O(n^{-1/3})$.

6.7. **Proof of Theorem 4. Sketch.** We study here the parameter Q_δ. We introduce here the Dirichlet series which depends on two parameters s, t,

$$H(2s, 2t) := \zeta(2s + 2t)(I - \mathbf{H}_{s+t})^{-1}(\mathbf{H}_{s+t} - \mathbf{H}_s) \circ \mathbf{H}^3_{s,[\ell]} \circ (I - \mathbf{H}_s)^{-1}[f](0),$$

It involves the weighted transfer operator $\mathbf{H}_{s,[\ell]}$ relative to the binary size ℓ and already defined in (30)

$$\mathbf{H}_{s,[\ell]}[f](x) := \sum_{m\geq 1} \frac{\ell(m)}{(m + x)^{2s}} f\left(\frac{1}{m + x}\right).$$

Applying the same principles as in Section 6.1 and 6.4 proves that it is well adapted to the study of cost $\underline{Q}_\delta$.

6.8. **Proof of Theorem 5. Sketch.** We study here the parameter $\ell(u_{\langle\delta\rangle})$. We introduce here the Dirichlet series which depends on two parameters s, t,

$$L(2s, 2t) := \zeta(2s + 2t)(I - \mathbf{H}_{s+t})^{-1} \circ \mathbf{H}'_{s+t} \circ (I - \mathbf{H}_{s+t})^{-1} \circ \mathbf{H}'_{s+t} \circ (I - \mathbf{H}_s)^{-1}[f](0),$$

It involves the derivative of the operator $\mathbf{H}_s$. Applying the same principles as in Section 6.1 and 6.4 proves that it is well adapted to the study of cost $\ell(\underline{u}_{\langle\delta\rangle})$.

7. Appendix : Propositions A, B, C.

We are interested in finding estimates for partial sums of coefficients, of the form

$$\Phi_w(N) := \sum_{k\leq N} c_k(w).$$

However, Perron's formula of order two provide estimates only for double sums,

$$\Psi_w(T) := \sum_{N\leq T} \Phi_w(N) = \sum_{N\leq T} \sum_{k\leq N} c_k(w)$$

of the form

$$\Psi_w(T) = \frac{R(w)}{2\tau(w) + 1} T^{2\tau(w)+1} \left[1 + O\left(A^2(w)T^{-2\sigma}\right)\right], \tag{63}$$

where the two conditions are fullfilled:

(i) the real σ belongs to $]0, 1/2[$ and the O-term is uniform with respect to $w \in \mathcal{W}$, as $T \to \infty$.

(ii) $\Re\tau(w) > 1/2$ and $\exists R_0 > 0$ for which $|R(w)| > R_0$ when $w \in \mathcal{W}$.

In this case, we say that Ψ_w satisfies (P) on $\mathcal{W}$ with the quadruple (τ, R, σ, A).

The main question is as follows:

From estimates on $\Psi_w(N)$, is it possible to deduce estimates for $\Phi_w(N)$?

It proves useful to introduce intermediate objects: for two indices N_- and N_+ which satisfy $N_- < N < N_+$, consider the two averages

$$\Phi^+_w(N) := \frac{1}{N_+ - N} \sum_{k=N+1}^{N_+} \Phi_w(k), \qquad \Phi^-_w(N) := \frac{1}{N - N_-} \sum_{k=N_-+1}^{N} \Phi_w(k).$$

The following of the appendix is devoted to the three main steps:

(*i*) It is always possible to deduce from (63) estimates for $\Phi_w^{\pm}(N)$, as soon as

$$N_- := N - \lfloor A(w)\, N^{1-\sigma} \rfloor, \quad N_+ = N + \lfloor A(w)\, N^{1-\sigma} \rfloor.$$

This is the aim of Proposition A.

(*ii*) Then, if the coefficients $c_n(w)$ are positive, these estimates can be tranferred into estimates for $\Phi_w(N)$. This is the aim of Proposition B.

(*iii*) Finally, if the coefficients $c_n(w)$ are dominated by $\widehat{c}_n(w)$ (i.e. $|c_n(w)| \leq \widehat{c}_n(w)$, and if estimates for $\widehat{\Psi}_w(T)$ of the same vein as $\Psi_w(T)$ hold, then it is possible to obtain estimates for $\Phi_w(N)$. This is the aim of Proposition C. Furthermore, this proposition naturally applies to a "moment generating function" setting.

7.1. **Statements of the propositions.** We now describe the three results in a more formal way.

Proposition A. [Basic Version] *Consider a sequence of functions $c_n : \mathcal{W} \to \mathbb{C}$ for which Ψ_w satisfies (P) on $\mathcal{W}$ with the quadruple (τ, R, σ, A). Then, for $N_- = N - \lfloor A(w)\, N^{1-\sigma} \rfloor$, the sum $\Phi_w^-(N)$ satisfies*

$$\Phi_w^-(N) := \frac{1}{N - N_-} \sum_{N_- < k \leq N} \phi_w(k) = R(w) N^{2\tau(w)} \cdot [1 + O(A(w)\, N^{-\sigma})],$$

where the constant in the O–term is uniform on $\mathcal{W}$. The same estimate holds for $\Phi_w^+(N)$.

Proposition B. [Positive coefficients] *Consider a sequence of functions $\widehat{c}_n : \mathcal{W} \to \mathbb{R}^+$ for which $\widehat{\Psi}_w$ satisfies (P) on $\mathcal{W}$ with the quadruple $(\widehat{\tau}, \widehat{R}, \widehat{\sigma}, \widehat{A})$. Then the sum $\widehat{\Phi}_w(N)$ satisfies*

$$\widehat{\Phi}_w(N) := \sum_{n \leq N} \widehat{c}_n(w) = \widehat{R}(w) N^{2\widehat{\tau}(w)} \cdot [1 + O(\widehat{A}(w)\, N^{-\widehat{\sigma}})],$$

where the constant in the O–term is uniform on $\mathcal{W}$.

Proposition C. [Domination] *Consider two sequences of functions $c_n : \mathcal{W} \to \mathbb{C}, \widehat{c}_n : \mathcal{W} \to \mathbb{R}^+$, for which the sums $\Psi_w, \widehat{\Psi}_w$ satisfy (P) on $\mathcal{W}$ with the respective quadruples (τ, R, σ, A) and $(\widehat{\tau}, \widehat{R}, \widehat{\sigma}, \widehat{A})$. Suppose furthermore that the following holds:*

(i) $\widehat{c}_n$ dominates c_n, i.e., $|c_n(w)| \leq \widehat{c}_n(w), \quad \forall w \in \mathcal{W}$.

(ii) The two functions $\tau(w), \widehat{\tau}(w)$ satisfy : $\exists \alpha < \widehat{\sigma}/2, \quad \forall w \in \mathcal{W}, |\Re\tau(w) - \Re\widehat{\tau}(w)| \leq \alpha,$.

Then, the sum $\Phi_w(T)$ satisfies, for any $w \in \mathcal{W}$,

$$\Phi_w(T) := \sum_{n \leq T} c_n(w) = R(w) T^{2\tau(w)} \cdot [1 + O(B(w) T^{-\beta})],$$

with $\beta := \min(\sigma, \widehat{\sigma} - 2\alpha), B := \max\{A, \widehat{A}\}$ and where the constant in the O–term is uniform on $\mathcal{W}$.

Proposition C' [Particular case of Proposition C.] *Consider the case when $\mathcal{W}$ is a neighborhood of 0, $c_n : \mathcal{W} \cap \mathbb{R} \to \mathbb{R}$, and $|c_n(w)| \leq c_n(\Re w)$. We let in this case $\widehat{c}_n(w) := c_n(\Re w)$. Then, if Ψ satisfies (P) with the quadruple (τ, R, σ, A), the function $\tau(w)$ is real as soon as w is real, and $\widehat{\Psi}_w$ satisfies (P) with the triple $(\tau(\Re w), R(\Re w), \sigma, A)$. If, moreover the function τ is continuous, then the difference $\Re\tau(w) - \widehat{\tau}(w) = \Re\tau(w) - \tau(\Re w)$ is less than $\sigma/4$ on a small enough neighborhood of $w = 0$. And, it is possible to apply Proposition C, with $\beta = \sigma/2$.*

This framework arises in a natural way when we study moment generating functions, since, in this case, the coefficient $c_n(w)$ is a sum of terms of the form $a_{j,n} exp[w b_{j,n}]$, with reals $a_{j,n}, b_{j,n}$.

7.2. **Proof of Proposition A.** We suppose that the following estimate holds:

$$\Psi_w(T) = F_w(T) \left[1 + O\left(A(w)\, T^{-2\sigma}\right)\right], \; T \to \infty, \qquad \text{with} \qquad F_w(T) = b(w)\, T^{a(w)},$$

where $b(w)$, $a(w)$ satisfy $\Re a(w) > 0$, $b(w) \neq 0$ on $\mathcal{W}$, and the O-error term is uniform for $w \in \mathcal{W}$. Denote by $T_- := T - \lfloor T^{1-\sigma} \rfloor$. The estimate of $\Psi_w(T)$ entails

$$\begin{aligned}\frac{1}{T-T_-}[\Psi_w(T) - \Psi_w(T_-)] \\ &= \frac{1}{T-T_-}\left(F_w(T) - F_w(T_-)\right) + \frac{A^2(w)}{T-T_-} O\left(F_w(T) T^{-2\sigma}\right) \\ &= F'_w(T)\left[1 + O\left((T-T_-)\frac{F''_w(T)}{F'_w(T)}, \frac{A^2(w)}{T-T_-}\frac{F_w(T)T^{-2\sigma}}{F'_w(T)}\right)\right].\end{aligned}$$

Then, if $a(w) \neq 1$, our assumptions on $F_w(T)$ and $a(w)$ imply

$$F'_w(T) = \Theta\left(T^{-1} F_w(T)\right), \qquad F''_w(T) = \Theta\left(T^{-2} F_w(T)\right),$$

with a uniform Θ. Therefore, we choose

$$T - T_- := \left(A^2(w)\frac{F_w(T)T^{-2\sigma}}{F''_w(T)}\right)^{1/2} = \Theta\left(A(w)\, T^{1-\sigma}\right),$$

and we take the same choice when $a(w) = 1$. If we wish to transfer theses estimates on integer parts, we need the condition $\sigma < 1/2$.

We now apply this result to our framework. We denote

$$N_- := N - \lfloor A(w)\, N^{1-\sigma} \rfloor, \qquad N_+ = N + \lfloor A(w)\, N^{1-\sigma} \rfloor,$$

and we consider

$$\Phi^+_w(N) := \frac{1}{N_+ - N}\sum_{k=N+1}^{N_+} \Phi_w(k), \qquad \Phi^-_w(N) := \frac{1}{N - N_-}\sum_{k=N_-+1}^{N} \Phi_w(k).$$

From (63), we have obtained estimates for $\Phi^\pm_w(N)$ of the form

$$\Phi^\pm_w(N) = R(w) N^{2\tau(w)}\left[1 + O\left(A(w)\, N^{-\sigma}\right)\right],$$

7.3. **Proof of Theorems B and C.** We compare now $\Phi^\pm_w(N)$ and $\Phi_w(N)$. We have

$$\Phi_w(N) = \Phi^-_w(N) + \frac{1}{N - N_-}\sum_{k=N_-+1}^{N} (\Phi_w(N) - \Phi_w(k))$$

$$\Phi_w(N) = \Phi^+_w(N) + \frac{1}{N_+ - N}\sum_{k=N+1}^{N_+} (\Phi_w(N) - \Phi_w(k))$$

If the coefficients $c_n(w)$ are real positive, the sequence $k \mapsto \Phi_w(k)$ is increasing, so that

$$\Phi^-_w(N) \leq \Phi_w(N) \leq \Phi^+_w(N)$$

and $\Phi_w(N)$ has the same estimate as $\Phi^\pm_w(N)$. This is true in particular for $\widehat{\Phi}_w(N)$ which has the same estimate as $\widehat{\Phi}^\pm(N)$, namely

$$\widehat{\Phi}_w(N) = \widehat{R}(w) N^{2\widehat{\tau}(w)}\left[1 + O\left(N^{-\widehat{\sigma}}\right)\right], \qquad |\widehat{\Phi}_w(N) - \widehat{\Phi}^-_w(N)| = O(\widehat{A}(w)\, N^{2\widehat{\tau}(w) - \widehat{\sigma}})$$

This provides the proof of Proposition B.

We now prove Theorem C. If the series has no longer positive coefficients, but is dominated, we observe that, for $k \leq N$,

$$|\Phi_w(N) - \Phi_w(k)| = \left|\sum_{n=k+1}^{N} c_n(w)\right| \leq \sum_{n=k+1}^{N} \widehat{c}_n(w) = \widehat{\Phi}_w(N) - \widehat{\Phi}_w(k),$$

which entails the inequality

$$|\Phi_w(N) - \Phi^-_w(N)| \leq |\widehat{\Phi}_w(N) - \widehat{\Phi}^-_w(N)|.$$

We apply the arguments of Proposition B which prove that

$$|\widehat{\Phi}_w(N) - \widehat{\Phi}^-_w(N)| = O(\widehat{A}(w) N^{2\widehat{\tau}(w) - \widehat{\sigma}}),$$

together with the estimate for $\Phi_w^-(N)$ obtained in Proposition A, and finally

$$\Phi_w(N) = R(w)N^{2\tau(w)}\left[1 + O\left(A(w)\,N^{-\sigma}\right) + O(\widehat{A}(w)\,N^{2\widehat{\tau}(w)-2\tau(w)-\widehat{\sigma}})\right]$$
$$= R(w)N^{2\tau(w)}\left[1 + O\left(B(w)\,N^{-\beta}\right)\right],$$

with $\beta := \min(\sigma, \widehat{\sigma} - 2\alpha), B := \max\{A, \widehat{A}\}$. This proves Proposition C.

8. Conclusion

This paper provides the first average-case analysis of a subquadratic gcd algorithm. We therefore extend the domain of applicability of dynamical analysis techniques, and show that such methods are also efficient for studying more complex Euclidean algorithms. The type of analysis performed here requires a precise study of the interrupted algorithms, and a precise description of the evolution of the distribution during the execution of the algorithm. This heavily uses the powerful tools of distributional analysis provided by [2, 21].

It would be also interesting to adapt the methodology developed here to other subquadratic gcd algorithms. We have in mind the algorithm recently designed by Stehlé and Zimmermann [25], based on a division using the least significant bits of the integers. The analysis of the plain gcd algorithm using this division is done in [8]. This is clearly a first step in that direction; however, a complete analysis of the SZ Algorithm would use Property *US*, and this Property is not known to hold in the context of the dynamical system related to this gcd using the least significant bits. Anyway, comparing the average–case behaviour of the SZ algorithm to other $\mathcal{HG}$ algorithms would be interesting since it would point out the influence of the division used, and explain experimental results observed in [25, 22].

References

[1] Baladi, V. and Vallée, B. Exponential Decay of Correlations for surface semi-flows without finite Markov partitions, Proceedings of the American Mathematical Society, 133 (3) pp 865-874, 2004.

[2] Baladi, V. and Vallée, B. Euclidean Algorithms are Gaussian, Journal of Number Theory, Volume 110, Issue 2 (2005) pp 331–386.

[3] Bedford, T., Keane, M., and Series, C., Eds. *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, Oxford University Press, 1991.

[4] Brent, R.P. Analysis of the Binary Euclidean algorithm, Algorithms and Complexity, New directions and recent results, ed. by J.F. Traub, Academic Press 1976, pp 321–355.

[5] Cesari, G. Parallel Implementation of Schönhage's Integer GCD Algorithm, Proceedings of ANTS-III, LNCS 1423, pp64-76.

[6] Dixon, J. D. The number of steps in the Euclidean algorithm, Journal of Number Theory 2 (1970), 414–422.

[7] Daireaux, B., and Vallée, B. Dynamical analysis of the parameterized Lehmer-Euclid Algorithm, Combinatorics, Probability, Computing, pp 499–536 (2004).

[8] Daireaux, B., Maume-Deschamps, V., Vallée, B. The Lyapounov Tortoise and the Dyadic hare, Discrete Mathematics and Theoretical Computer Science 2005, Proceedings of AofA'05, pp 71-94 (2005).

[9] Dolgopyat, D. On decay of correlations in Anosov flows, *Ann. of Math.* 147 (1998), pp 357–390.

[10] Ellison, W. and Ellison, F. *Prime Numbers,* Hermann, Paris, 1985.

[11] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics*, Book in preparation (1999), see also INRIA Research Reports 1888, 2026, 2376, 2956.

[12] Fürer, M. Faster Integer Multiplication, Proceedings of STOC'07, pp 57–66

[13] Heilbronn, H. On the average length of a class of continued fractions, Number Theory and Analysis, ed. by P. Turan, New-York, Plenum, 1969, pp 87-96.

[14] Hensley, D. The number of steps in the Euclidean algorithm, *Journal of Number Theory* 49, 2 (1994), 142–182.

[15] Hwang, H.-K. On convergence rates in the central limit theorems for combinatorial structures, *European Journal of Combinatorics* 19 (1998) pp 329–343.

[16] Jebelean, T. Practical Integer Division with Karatsuba Complexity, Proceedings of ISSAC'97.

[17] Jebelean, T. A Double-Digit Lehmer–Euclid Algorithm for finding the GCD of Long Integers. Journal of Symbolic Computation (1995) 19, pp 145–157.

[18] Knuth, D.E. *The art of Computer programming*, Volume 2, 3rd edition, Addison Wesley, Reading, Massachussets, 1998.

[19] Knuth, D.E. The analysis of algorithms, Actes du Congrès des Mathématiciens, Volume 3, pp 269-274, Gauthier-Villars 1971.

[20] Lehmer, D. H. Euclid's algorithm for large numbers, Am. Math. Mon. (1938) 45 pp 227–233.

[21] Lhote, L. and Vallée, B. Gaussian Laws for the main parameters of the Euclid Algorithm, to appear in Algorithmica (2007) –Short version : Sharp estimates for the main parameters of the Euclid Algorithm, Proceedings of LATIN'06, LNCS 3887, pp 689-702.

[22] Möller, N. On Schönhage's algorithm and subquadratic integer gcd computation, Mathematics of Computation, Volume 77, Number 261, January 2008, pp 589–607.
[23] Ruelle, D. *Thermodynamic formalism,* Addison Wesley (1978).
[24] Schönhage, A. Schnelle Berechnung von Kettenbruchentwicklungen, Acta Informatica pp 139–144 (1971)
[25] Stehlé, D. and Zimmermann, P. A Binary Recursive Gcd Algorithm, Proceedings of ANTS'04, LNCS 3076 (2004), pp 411-425.
[26] Vallée, B. Dynamical Analysis of a Class of Euclidean Algorithms, Theoretical Computer Science, vol 297/1-3 (2003) pp 447–486.
[27] Vallée, B. Euclidean Dynamics, [55 pages], Discrete and Continuous Dynamical Systems, 15 (1) May 2006, pp 281-352.
[28] Vallée, B. Digits and Continuants in Euclidean Algorithms. Ergodic Versus Tauberian Theorems, Journal de Théorie des Nombres de Bordeaux 12 (2000) pp 531-570.
[29] Yap, C.K. *Fundamental Problems in Algorithmic Algebra*, Princeton University Press (1996).

Eda Cesaratto: Facultad de Ingeniera, Universidad de Buenos Aires, Argentina and GREYC, UMR CNRS 6072, Université de Caen and ENSICAEN, F-14032 Caen, France. ecesara@fi.uba.ar

Julien Clément : GREYC, UMR CNRS 6072, Université de Caen and ENSICAEN, F-14032 Caen, France. julien.clement@info.unicaen.fr

Benoît Daireaux: IrisResearch Center, Stavanger, Norway. Benoit.Daireaux@irisresearch.no

Loïck Lhote: GREYC, UMR CNRS 6072, Université de Caen and ENSICAEN, F-14032 Caen, France. loick.lhote@info.unicaen.fr

Véronique Maume-Deschamps: Institut de Science Financière et d'Assurances - ISFA Université Lyon 1, France. veronique.maume@univ-lyon1.fr

Brigitte Vallée: GREYC, UMR CNRS 6072, Université de Caen and ENSICAEN, F-14032 Caen, France. brigitte.vallee@info.unicaen.fr