Open access • Journal Article • DOI:10.1007/S00224-015-9633-5

Towards a Realistic Analysis of the QuickSelect Algorithm — Source link

Julien Clément, James Allen Fill, Thu Hien Nguyen Thi, Brigitte Vallée

Institutions: University of Caen Lower Normandy, Johns Hopkins University

Published on: 01 May 2016 - Theory of Computing Systems \/ Mathematical Systems Theory (Springer US)

Topics: Quickselect

Towards a realistic analysis of the QuickSelect algorithm

Julien Clément, James Allen Fill, Thu Hien Nguyen Thi, Brigitte Vallée

Noname manuscript No.
(will be inserted by the editor)

Towards a realistic analysis of the QuickSelect algorithm

Julien Clément · James Allen Fill · Thu Hien Nguyen Thi · Brigitte Vallée

Received: date / Accepted: date

Dedicated to the memory of Philippe Flajolet

Abstract We revisit the analysis of the classical `QuickSelect` algorithm. Usually, the analysis deals with the mean number of key comparisons, but here we view keys as words produced by a source, and words are compared via their symbols in lexicographic order. Our probabilistic models belong to a broad category of information sources that encompasses memoryless (i.e., independent-symbols) and Markov sources, as well as many unbounded-correlation sources. The "realistic" cost of the algorithm is here the total number of symbol comparisons performed by the algorithm, and, in this context, the average-case analysis aims to provide estimates for the mean number of symbol comparisons. For the `QuickSort` algorithm, known average-case complexity results are of $\Theta(n \log n)$ in the case of key comparisons, and $\Theta(n \log^2 n)$ for symbol comparisons. For `QuickSelect` algorithms, and with respect to key comparisons, the average-case complexity is $\Theta(n)$. In this present article, we prove that, with respect to symbol comparisons, `QuickSelect`'s average-case complexity remains $\Theta(n)$. In each case, we provide explicit expressions for the dominant constants, closely related to the probabilistic behaviour of the source.

We began investigating this research topic with Philippe Flajolet, and the short version of the present paper (the ICALP'2009 paper) was written with him. As usual, Philippe played a central role, notably on the following points:

The first, third, and fourth authors thank two ANR Projects: ANR BOOLE (ANR 2009 BLAN 0011) and ANR MAGNUM (ANR 2010 BLAN 0204). Research for the second author was supported by The Johns Hopkins University's Acheson J. Duncan Fund for the Advancement of Research in Statistics.

J. Clément, T. H. Nguyen Thi and B. Vallée
GREYC, CNRS-UMR 6072, Université de Caen, ENSICAEN, 14032 Caen, France, E-mail: {Julien.clement, thu-hien.nguyen-thi, brigitte.vallee}@unicaen.fr.

J.A. Fill
Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD 21218-2682 (USA), E-mail: jimfill@jhu.edu.

introduction of the `QuickVal` *algorithm, tameness of sources, and use of the Rice's method. He also made many experiments exhibiting the asymptotic slope $\rho(\alpha)$ and plotted nice graphs, which are reproduced in this paper. Even though the extended abstract does not provide any proof of the analysis of the algorithm* `QuickQuant`*, Philippe also devised with us a precise plan for this proof which has now completely been written. For all these reasons, we could have added (and certainly would have liked to add) Philippe as a co-author of this paper. On the other hand, Philippe was extremely exacting of how his papers were to be written and organised, and we cannot be sure that he would have liked or validated our editing choices. In the end, this is why we have decided not to include him as a co-author, but instead, to dedicate, with deference and affection, this paper to his memory. Thank you, Philippe!*

Keywords Probabilistic analysis of algorithms; Sorting and searching algorithms; Selection problem; Pattern matching; Permutations; Information theory; Rice's method; Asymptotic estimates; Quickselect

Mathematics Subject Classification (2000) F2.2: Pattern matching, sorting and searching – G2.1: Generating functions, permutations – G4: Algorithm design and analysis – H1.1: Information theory – I1.2: Analysis of algorithms

Introduction

Every student of a basic algorithms course is taught that, on average, the complexity of `QuickSort` is $O(n \log n)$, that of binary search is $O(\log n)$, and that of radix-exchange sort is $O(n \log n)$; see for instance [16], [22]. Such statements are based on specific assumptions —that the comparison of data items (for the first two) and the comparison of symbols (for the third one) have unit cost— and they have the obvious merit of offering an easy-to-grasp picture of the complexity landscape. However, as noted by Sedgewick, these simplifying assumptions suffer from limitations: they do not make possible a precise assessment of the relative merits of algorithms and data structures that resort to different methods (e.g., comparison-based versus radix-based sorting) in a way that would satisfy the requirements of either information theory or algorithms engineering. Indeed, computation is not reduced to its simplest terms, namely, the manipulation of totally elementary symbols, such as bits, bytes, characters. Furthermore, such simplified analyses say little about a great many application contexts, in databases or natural language processing, for instance, where information is highly "non-atomic", in the sense that it does not plainly reduce to a single machine word.

First, we observe that, for commonly used data models, the mean costs $S(n)$ and $K(n)$ of *any* algorithm under the symbol comparison and the key comparison model, respectively, are connected by the universal relation $S(n) = K(n) \cdot O(\log n)$. This results from the fact that at most $O(\log n)$ symbols suffice, with high probability, to distinguish n keys; cf. the analysis of the height of digital trees, also known as "*tries*", in [2]. The surprise is that there are cases where this upper bound is tight, as in `QuickSort`; others where both costs are of the same order, as in `QuickSelect`. In previous works [26,3], we have already shown that the expected cost of `QuickSort` is $\Theta(n \log^2 n)$, *not* $\Theta(n \log n)$, when *all* elementary operations—symbol comparisons—are taken into account. By contrast, we prove here that the cost of `QuickSelect` turns out to be $\Theta(n)$, in both the old and the new world, albeit, of course, with different implied constants.

Main results. Our main object of study is the `QuickSelect` algorithm, when the keys are *words* that are drawn from a given source $\mathcal{S}$, and we deal with the mean number of *symbol comparisons* performed by the algorithm. This analysis is *more realistic* than the classical analyses which study the mean number of keys comparisons in two ways. First it studies a more realistic cost, that is the number of symbol comparisons instead of the number of key comparisons. Second it deals with more realistic inputs: words drawn from a specified source instead of keys. However, we do not claim that this is an actual realistic analysis: this is just a first step towards such an analysis.

When applied to n keys assumed to be *independently drawn* from the same source $\mathcal{S}$, the `QuickSelect`(m, n) algorithm selects the m-th smallest element. We shall mostly focus our attention on situations where the rank m is proportional to n, being of the form $m = \lfloor 1 + \alpha(n-1) \rfloor$, for some $\alpha \in [0, 1]$, so that the algorithm determines the α-th quantile; it will then be denoted by `QuickQuant`$_\alpha(n)$. For $\alpha = 0$, the rank m equals 1 and this is `QuickMin`(n). For $\alpha = 1$, the rank m equals n and this is `QuickMax`(n) (both `QuickMin` and `QuickMax` are particular cases of `QuickSelect`). We also consider the case where m is randomly chosen in $[1 \,..\, n]$, and this is `QuickRand`(n) which is clearly not a "real" algorithm and is introduced for modelling purposes.

Our main result describes the case of the `QuickSelect` (or more accurately `QuickQuant`$_\alpha$) algorithm with all its variants. We first show in Theorem 1 that, with respect to the number of symbol comparisons, and provided the source be *tame*[1], the mean cost of `QuickSelect` remains of order $\Theta(n)$. However, this cost involves various constants that depend on the source $\mathcal{S}$ (and possibly on the real α); these are precisely described in Theorem 2 and displayed in Figure 2. Here, for the `QuickSelect` algorithms, we prove all the results which were only stated in the extended abstract [26], and we exhibit the probabilistic features of the source which play a role in the analysis: each algorithm of interest is related to a particular constant depending on the source; this constant describes the interplay between the algorithm and the source and explains how the efficiency of the algorithm depends on the source.

General methodology and plan of the paper. Like many studies in analytic combinatorics, our analysis is divided into two main parts: an *algebraic*

[1] The word "tame" was proposed by Philippe Flajolet and used for the first time in [26]. Later on, most of the papers which deal with probabilistic sources use similar notions and the word "tame" is now largely used.

or *combinatorial* part and an *asymptotic* and *analytic* step. In analytic combinatorics, the first part computes a mathematical object (most of the times a generating function) with algebraic and exact tools, and the second part deals with the generating function as a function of the complex variable, studies its singularities, and transfers this knowledge about its singularities into an asymptotic expansion. We adopt and adapt the same philosophy here.

After a description of the main results done in Section 1, we first draw our general framework in Section 2, and introduce the two main objects: the algorithm and its inputs. We begin in Section 2.2 by a description of the algorithm `QuickSelect` and some particular cases of this algorithm, and, inside the classical model of permutations, we are interested in the analysis of the mean number of *local* key comparisons between a pair of keys of given ranks. Then, in Section 2.3, we model the inputs, and introduce the mechanism which emits the words (now the inputs of the algorithm). Such a probabilistic mechanism, called a source, is described by two generating functions, of Dirichlet type, which encapsulate the main probabilistic properties of the source. A source can be parameterized by the unit interval (see Section 2.6). Here, as we explain in Section 2.7, the source mainly intervenes via its *coincidence*, which measures the similarity between a pair of words with given parameters, and defines *fundamental triangles* which describe the location of pairs of words which share a common prefix.

We then perform our analysis, in the sequel of the paper. We compute in Section 3 the generating function related to our specific problem: this is the *mixed Dirichlet series* $\varpi(s)$, which encapsulates both the properties of the source and the characteristics of the algorithm. In the algebraic part, it yields the *exact value of the mean number of symbol comparisons* performed by each algorithm on words produced by a parameterized source. We then perform a second part of the analysis where we study the analytic properties of the mixed Dirichlet series $\varpi(s)$, namely the position and the nature of its singularities, and transfer this knowledge to obtain the *asymptotic* mean number of *symbol comparisons.* This transfer may be direct as in Section 4, or indirect as in Section 5.

Section 3 is devoted to the algebraic part of the analysis. We first compute the *density* of the algorithm. It uniquely depends on the algorithm and provides a *precise* measure of the mean number of *key comparisons* performed near *specific points.* As it is easier to compute this density in the *Poisson model*, where the number of keys instead of being fixed follows a Poisson distribution, this first step provides an expression of the Poissonized density relative to each algorithm. We use the following chain to get the result: We begin with the mean number of local key comparisons between two keys, obtained in Section 2.2. Viewing keys as words, we use the parameterization of sources described in Section 2.6. Together with basic properties of the Poisson model, we obtain a precise expression of the density, but not for all the variants of `QuickSelect`. This is why we introduce a new algorithm `QuickVal`, closely related to `QuickQuant`, for which it is possible to compute the density.

Then, the source intervenes via its coincidence and its *fundamental triangles*. With integrals on such triangles, we obtain an exact expression for the *mean number of symbol comparisons* in the Poisson model in Section 3.7. With an easy transfer from the Poisson model to the Bernoulli model (where the number n of keys is fixed), we deduce in Section 3.8 an exact expression of the mean number of symbol comparisons in the Bernoulli model. This expression is given as a binomial sum which deals with the *mixed Dirichlet series*, at least for the algorithms `QuickMin`, `QuickMax` and `QuickVal`. For the `QuickQuant` algorithm, there also exists an exact expression, of another type, which involve integral expressions and notably the density of the Beta distribution.

For the first three algorithms, we carry out in Section 4 the corresponding asymptotic analysis with Rice's method (described in Section 4.1) which transforms a binomial sum into an integral in the complex plane. To use this method, we need the mixed Dirichlet series to be *tame*. As this series $\varpi(s)$ is related to the generating functions of the source, it is possible to transfer the tameness of these generating functions to the series $\varpi(s)$. We prove that convenient tameness of (the generating functions of) the source indeed entails tameness of the mixed Dirichlet series in Section 4.3. Then applying the Rice's method leads to asymptotic estimates for the mean number of symbol comparisons of `QuickMin`, `QuickMax`, `QuickVal`; we also directly deduce the analysis of `QuickRand`.

The analysis of the `QuickQuant` algorithm, i.e., the quantile version of `QuickSelect`, is performed in an indirect way in Section 5, and uses the close connection between $\texttt{QuickVal}_\alpha$ and $\texttt{QuickQuant}_\alpha$. From the analytic point of view, this translates into an integral form which involves the density of the α-quantile. We then use Laplace's method (see Section 5.2) that leads to asymptotic estimates for the mean number of symbol comparisons of `QuickQuant`.

Finally, the last Section 6 is devoted to the precise study of the constants that are involved in the analysis, and will depend on the pair algorithm/source considered.

Relation between the four articles on related subjects. Our approach relies on methods from analytic combinatorics [10], as well as on information theory. It also relates to earlier analyses of digital trees ("tries") by Clément, Flajolet, and Vallée [2,25]. But we now discuss the relation between the present paper and three other studies on closely related subjects.

There are two extended abstracts: the paper [26], in the proceedings of ICALP 2009, and the paper [3], in the proceedings of STACS 2013. Then, there are two long papers which are journal versions of the previous ones: The paper [4] is the journal version of [3], whereas the present paper can be viewed as the journal version of [26].

The paper [26] was the first paper devoted to the general subject "analysis with respect to symbol comparisons"; this short paper just mentioned the main steps of the methodology: even if the algebraic steps are well described (as in our Section 3), the analytic steps (as in our Sections 4 and 5) are just mentioned. It then focused on the `QuickSort` and `QuickSelect` algorithms. It stated the results which are described here in the introduction, but did not provide any proofs: the present Section 4.3 is completely new.

The paper [3] was the second paper devoted to the subject, and it performed two tasks: it shows the generality of the method, and designs a quite general framework for the "analysis with respect to symbol comparisons". It provides a precise description of the algebraic steps, and also explains the analytic steps, with the introduction of the various notions of tameness. The second aim of [3] was to apply the method to five algorithms: amongst them, one again finds `QuickSort` and `QuickMin`, but also other classical algorithms, as *insertion sort*, *bubble sort* and *minimum selection*. Again, this is only a extended abstract in which most proofs are omitted.

The two journal versions provide all the details for the methodology, and precise analyses for algorithms of interest; each of the two journal versions is devoted to a class of algorithms. The paper [4] focuses on the three sorting algorithms `QuickSort`, `InsSort`, `BubSort`, whereas the present paper adapts the general method to the algorithms of the `QuickSelect` class, for which it provides a complete analysis. This analysis is particular and somehow indirect, as was already announced in [3]: it does not seem possible to directly analyse `QuickSelect`, and there is a "detour" via the `QuickVal` algorithm. Then, in the long present paper, we explain how to precisely return from `QuickVal` to `QuickSelect`.

Relation with other studies. Here, we restrict ourselves to a precise average-case analysis of the `QuickSelect` algorithm (with respect to the number of symbol comparisons). Other studies, done by the second author with collaborators, focus on distributional analyses in the same framework (see [8,6]).

1 Main results.

1.1 *Classical results on the* `QuickSelect` *algorithm*

We recall classical results on the average-case analysis of the `QuickSelect` algorithm. They involve the κ function, defined from the entropy function h,

$$\kappa(\alpha) = 2[1 + h(\alpha)], \qquad \text{with} \quad h(\alpha) = \alpha\,|\log\alpha| + (1-\alpha)\,|\log(1-\alpha)|\,, \quad (1)$$

that plays an important role here, notably in the analysis of $\texttt{QuickQuant}_\alpha$ (and where $|\cdot|$ denotes the absolute value). Its graph is represented in Figure 1. The exact mean number $K(m, n)$ of key comparisons performed by $\texttt{QuickSelect}(m, n)$ for any m was obtained by Knuth in [15]. It is given in the following proposition. The column 4 of Table 1 gives asymptotic estimates for some particular cases of interest.

Lemma 1 [Classical Average-case analysis] *The mean number of key comparisons performed by the algorithm* $\texttt{QuickSelect}(m, n)$ *in the uniform permutation model satisfies the following:*

(*a*) *For a general rank m, it involves the harmonic number H_k under the form*

$$K(m,n) = 2\left[(n+1)H_n - (n+3-m)H_{n+1-m} - (m+2)H_m + n + 3\right].$$

In the case of QuickMin ($m=1$) *and* QuickMax ($m=n$), *one has*

$$K(1,n) = K(n,n) = 2n - 2H_n.$$

(*b*) *For* $\texttt{QuickQuant}_\alpha$, *the mean number $K_\alpha(n)$ involves the κ function defined in* (1), *under the form*

$$K_\alpha(n) \sim \kappa(\alpha)n.$$

We have in the case of QuickMed *(selecting for the median),*

$$K_{\frac{1}{2}}(n) \sim \kappa(\tfrac{1}{2})\, n = 2(1+\log 2)n.$$

(*c*) *For* QuickRand(n), *one has $K(n) \sim 3n$.*

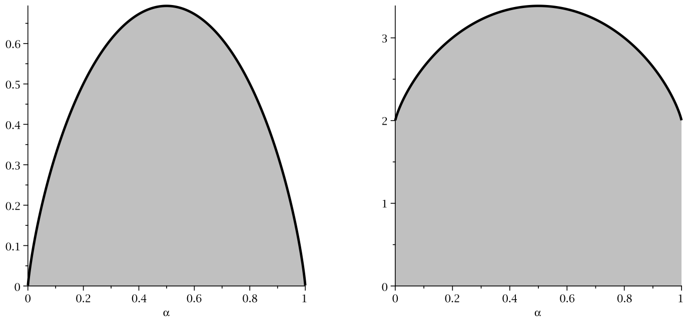


Fig. 1: The functions h and κ.

Special cases of QuickSelect(m,n)	Output	Value of m	Asymptotic estimate of $K(m,n)$
$\texttt{QuickQuant}_\alpha$	α-quantile	$m = \lfloor 1+\alpha(n-1)\rfloor$	$\kappa(\alpha)n$
QuickMed	median	$m = \lfloor (n+1)/2 \rfloor$	$2(1+\log 2)n$
QuickMin	minimum	$m = 1$	$2n$
QuickMax	maximum	$m = n$	$2n$
QuickRand	random	$m \in [1\,..\,n]_{\mathcal{R}}$	$3n$

Table 1: Asymptotic estimates for the mean number of key comparisons for QuickSelect and special cases of QuickSelect.

1.2 ***Tameness of sources.***

For a source $\mathcal{S}$ on the alphabet Σ, and for a prefix $w \in \Sigma^\star$, we define the *fundamental probability* p_w as the probability that a word emitted by the source $\mathcal{S}$ begins with the prefix w. We consider two generating functions of the source, of Dirichlet type, which are built with fundamental probabilities. They encapsulate the main probabilistic properties of the source and are central in our study.

Definition 1 (Dirichlet series of the source) There are two main types of Dirichlet series associated to a source $\mathcal{S}$.

(i) The Λ generating functions is defined as

$$\Lambda(s) = \sum_{w \in \Sigma^\star} p_w^s = \sum_{k \geqslant 0} \Lambda_k(s), \qquad \text{with} \quad \Lambda_k(s) = \sum_{w \in \Sigma^k} p_w^s.$$

(ii) The Π generating function is defined as

$$\Pi(s) = \sum_{k \geqslant 0} \pi_k^s, \quad \text{with} \quad \pi_k = \sup\{p_w, w \in \Sigma^k\}.$$

The Λ series was introduced for the first time in [25]. It plays a central role in many probabilistic studies about sources which adopt the point of view of analytic combinatorics. This series intervenes mainly in the analysis of sorting algorithms. The Π series was introduced for the first time in [26], and is suited to the analysis of selection algorithms (like `QuickSelect`). Here, these two series are important to us notably via the following definition which will be motivated later on in the paper.

Definition 2 (Tameness) There are two different notions for the tameness of the source:

(i) A source is weakly Λ-tame with width $\delta = \delta_\Lambda$ if the abscissa of absolute convergence of the Dirichlet series $\Lambda(s)$ is equal to $1 + \delta$.

(ii) A source is Π-tame with width $\delta = \delta_\Pi$ if the abscissa of absolute convergence of the Dirichlet series $\Pi(s)$ is equal to δ.

We now see that the tameness width which intervenes in most of our results is $\delta_0 := \max(2\delta_\Lambda, \delta_\pi)$, notably via its position with respect to 1.

1.3 ***General result.***

We provide general asymptotic estimates for the algorithm `QuickSelect` (or more accurately `QuickQuant`$_\alpha$), particular cases like `QuickMin` and `QuickMax`, as well as `QuickRand` (which models the average behaviour of `QuickSelect` for a random uniform rank).

Theorem 1 (General asymptotic estimates) *For a source $\mathcal{S}$ both Π-tame and weakly Λ-tame with a small enough width, the mean number of symbol comparisons of all the variants of* $\texttt{QuickSelect}(m, n)$ *is* $\Theta(n)$. *More precisely, the following holds:*

(*a*) *Consider a source which is both weakly Λ-tame with width δ_Λ and Π-tame with width δ_Π and assume that $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$ is strictly less than 1. Then the mean number of symbol comparisons $\texttt{QQ}_\alpha(n)$ performed by $\texttt{QuickQuant}_\alpha$ satisfies, for any $\delta \in](1+\delta_0)/2, 1[$ and any $\alpha \in [0,1]$,*

$$\texttt{QQ}_\alpha(n) = \rho_{\mathcal{S}}(\alpha)\, n + O(n^\delta).$$

Moreover the constant hidden in the O-term is uniform (with respect to α) for any interval of the form $[\alpha_1, 1-\alpha_1]$ with $\alpha_1 > 0$.

(*b*) *Consider a source which is both weakly Λ-tame with width δ_Λ and Π-tame with width δ_Π and assume that $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$ is strictly less than 1. Then, the mean number $\texttt{QR}(n)$ of symbol comparisons performed by $\texttt{QuickRand}(n)$ satisfies, for $\delta \in]\delta_0, 1[$,*

$$\texttt{QR}(n) = \gamma_{\mathcal{S}}\, n + O(n^\delta), \qquad \textit{with} \quad \gamma_{\mathcal{S}} = \int_0^1 \rho_{\mathcal{S}}(\alpha)\, d\alpha.$$

(*c*) *If the source is weakly Λ-tame with width $\delta_\Lambda < 1/2$, the mean number of symbol comparisons, $\texttt{QM}_{(-)}(n) = \texttt{QQ}_0(n)$ for $\texttt{QuickMin}(n)$ and $\texttt{QM}_{(+)}(n) = \texttt{QQ}_1(n)$ for $\texttt{QuickMax}(n)$, satisfies, for $\delta \in]2\delta_\Lambda, 1[$, with $\epsilon = \pm$,*

$$\texttt{QM}_{(\epsilon)}(n) = \rho_{\mathcal{S}}^{(\epsilon)} n + O(n^\delta), \qquad \textit{with} \quad \rho_{\mathcal{S}}^{(+)} = \rho_{\mathcal{S}}(1), \qquad \rho_{\mathcal{S}}^{(-)} = \rho_{\mathcal{S}}(0).$$

1.4 *Expressions for the constants.*

The following result provides precise expressions for the constants which appear in the dominant terms of the previous Theorem[2]:

Theorem 2 (Expressions for constants) *For a general source $\mathcal{S}$, where the probabilities $p_w^{(+)}$, $p_w^{(-)}$, and $p_w = 1 - p_w^{(+)} - p_w^{(-)}$ are defined as in Definition 3 and Eq. (5), the constants involved in Theorem 1 admit the following expressions:*

(*a*) *The quantile constant $\rho_{\mathcal{S}}(\alpha)$ is defined as*

$$\rho_{\mathcal{S}}(\alpha) = \sum_{w \in \Sigma^\star} p_w\, \kappa\left(\frac{\alpha - p_w^{(-)}}{p_w}\right),$$

and involves the function κ given by $\kappa(y) = 2[1 + H(y)]$, with

$$H(y) = \begin{cases} y \log\left(1 - \frac{1}{y}\right) & \textit{if } y < 0, \\ h(y) = -y \log y - (1-y)\log(1-y), & \textit{if } y \in [0,1], \\ (y-1)\log\left(1 - \frac{1}{y}\right) & \textit{if } y > 1. \end{cases}$$

[2] This corrects small errors in Figure 1 of [26].

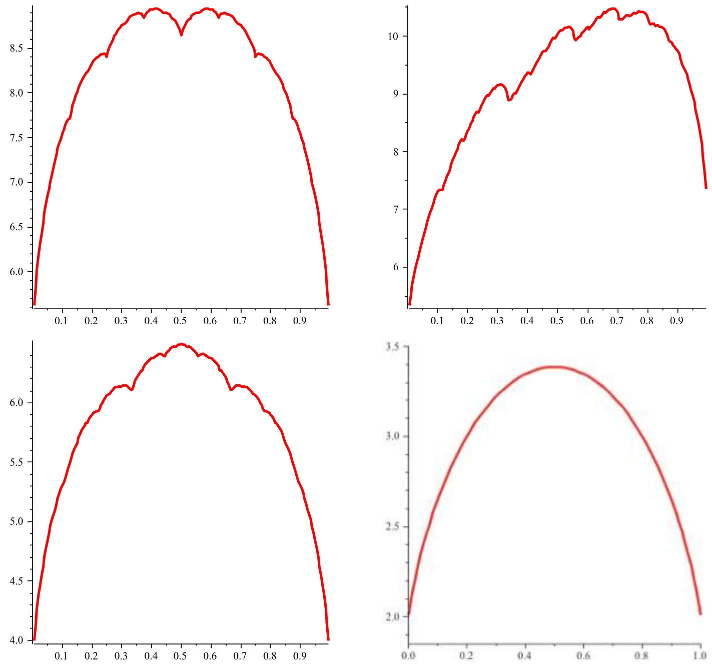


Fig. 2: Plots of the function $\rho_{\mathcal{S}}(\alpha)$ for $\alpha \in [0,1]$ and three sources: $\mathcal{B}(1/2,1/2)$, $\mathcal{B}(1/3,2/3)$, and $\mathcal{B}(1/3,1/3,1/3)$. The curves are deformations of the curve $\alpha \mapsto \kappa(\alpha)$ whose plot is on the bottom right corner. The plots illustrate the fractal character of the constants involved in `QuickQuant`.

(*b*) *The Random selection constant is*

$$\gamma_{\mathcal{S}} = \sum_{w \in \Sigma^\star} p_w^2 \left[2 + \frac{1}{p_w} + \sum_{\epsilon = \pm} \left[\log\left(1 + \frac{p_w^{(\epsilon)}}{p_w}\right) - \left(\frac{p_w^{(\epsilon)}}{p_w}\right)^2 \log\left(1 + \frac{p_w}{p_w^{(\epsilon)}}\right) \right] \right].$$

(*c*) *The Min/Max constants are*

$$\rho_{\mathcal{S}}^{(\epsilon)} = 2 \sum_{w \in \Sigma^\star} p_w \left[1 - \frac{p_w^{(\epsilon)}}{p_w} \log\left(1 + \frac{p_w}{p_w^{(\epsilon)}}\right) \right].$$

Note that the function H of Assertion (a) is a continuous extension, symmetric about $y = 1/2$, of the entropic function h to the whole real line. Note also that the κ function—at least its restriction to the interval $[0,1]$—already intervenes in the mean number $K(n)$ of key comparisons of the `QuickQuant` algorithm, as recalled in Lemma 1. Figure 2 shows the graphs of the function $\alpha \mapsto \rho(\alpha)$ for different Bernoulli sources, together with the graph of the κ function. It appears that the graph of ρ is a deformation of the graph of κ, and the plots illustrate the fractal character of the constants involved in `QuickQuant`, as the expression for ρ in terms of κ suggests. We see that the function ρ is not always maximised at $\alpha = 1/2$, not even for symmetric sources, even though κ is maximised at $1/2$. There is then a natural question to be asked: For which (symmetric) sources is the maximum for ρ attained at $\alpha = 1/2$? If the maximum is not attained at $\alpha = 1/2$, at which point is it attained? We return to this question in Section 6.

1.5 ***Particular case of the binary source.***

All these constants specialise nicely for the *standard binary source* $\mathcal{B}$ (under which keys are compared via their binary expansions, with uniform independent bits), in which case they admit pleasant expressions that simplify and extend those of Fill and Nakama [7] and Grabner and Prodinger [11] and lead themselves to precise numerical evaluations.

Proposition 1 *For a binary source, the constants admit the following expressions:*

$$\rho_{\mathcal{B}}(\alpha) = \sum_{\ell \geqslant 0} \frac{1}{2^\ell} \sum_{k=0}^{2^\ell - 1} \kappa(2^\ell \alpha - k)$$

$$\rho_{\mathcal{B}}^{(\epsilon)} = 4 + 2 \sum_{\ell \geqslant 0} \frac{1}{2^\ell} \sum_{k=1}^{2^\ell - 1} \left[1 - k \log\left(1 + \frac{1}{k}\right)\right] \doteq 5.27937\,82410\,80958,$$

and also

$$\gamma_{\mathcal{B}} = \frac{14}{3} + 2 \sum_{\ell=0}^{\infty} \frac{1}{2^{2\ell}} \sum_{k=1}^{2^\ell - 1} \left[k + 1 + \log(k+1) - k^2 \log\left(1 + \frac{1}{k}\right)\right] \doteq 8.20730\,88638.$$

Here, the function κ is first defined in (1) *on the interval* $[0, 1]$ *and extended to the whole real line in Theorem 2(a).*

2 General framework.

Here, we draw our general framework, already provided in [26] and [3]. We first characterise in Section 2.1 the strategy of each algorithm (which keys are compared? with which probability?), then we describe the model of source in Section 2.3, together with the particular cases of "simple" sources. We then describe the modelling of a parameterized source in Section 2.6, and introduce the central notion of coincidence in Section 2.7.

2.1 ***The classical probabilistic model based on permutations.***

Consider a totally ordered set of keys $\mathcal{U} = \{U_1 < U_2 < \cdots < U_n\}$ and any algorithm $\mathcal{A}$ which only performs comparisons and exchanges between keys. The initial input is the sequence $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ defined from $\mathcal{U}$ by a permutation $\sigma \in \mathfrak{S}_n$ via the equalities $V_i = U_{\sigma(i)}$. The execution of the algorithm does not actually depend on the input sequence, but only on the permutation σ which defines the input sequence from the final (ordered) sequence. Thus the permutation σ is the actual input of the algorithm and the set of all possible inputs is the set $\mathfrak{S}_n$.

The classical average-case analysis deals with the (total) mean number of key comparisons performed by the algorithm. Here, we fix a pair of keys U_i and U_j and we are interested by the mean number of comparisons performed by the algorithm between the two keys U_i and U_j, which will be denoted by $\pi_n(i, j)$. The computation of $\pi_n(i, j)$ is not usual in classical analysis of algorithms, and this is the first step of our method. Indeed, when keys become words, their relative positions will have an influence on the "cost" needed for distinguishing them.

When we take the sum over all the pairs (i, j) of the probabilities $\pi_n(i, j)$, (with $1 \leqslant i < j \leqslant n$), we obtain the mean number of key comparisons performed by each algorithm. We thus obtain here a new derivation of well-known results for popular algorithms in the uniform permutation model.

2.2 *The algorithms* `QuickSort`*,* `QuickSelect` *and* `QuickVal`*.*

These algorithms are based on the "Divide and Conquer" principle. All other keys are compared to the first key of the array, which is used as a pivot. During the `Partition` stage, the keys that are smaller than the pivot are placed on its left in the array, whereas the keys that are greater are placed on its right. After this partitioning, the pivot is at the right place. In the sequel, the sets $\mathcal{U}_{[i,j]}$ formed by all the keys U_ℓ with $\ell \in [i, j]$ play an important rôle in the description of the algorithms.

`QuickSort`. The algorithm was introduced in 1962 by C.A.R Hoare in his original papers [12,13,14] and improved by several authors, including Knuth [16], van Emden [27], and Sedgewick [21]. The `QuickSort`(n) algorithm recursively sorts the two sub-arrays. While the pivot does not belong to the subset $\mathcal{U}_{[i,j]}$, this set is not separated by the pivot. When the pivot belongs to the subset $\mathcal{U}_{[i,j]}$, the keys U_i and U_j are compared if and only if U_i or U_j is a pivot. This event coincides with the event "U_i or U_j is the first key among those in the subset $\mathcal{U}_{[i,j]}$". After such a comparison, the keys are separated and no longer compared. Thus, the mean number of key comparisons between U_i and U_j is, for $i < j$, equal to

$$\pi_n(i, j) = \frac{2}{\operatorname{card} \mathcal{U}_{[i,j]}} = \frac{2}{j - i + 1}.$$

```
Algorithm Partition(V, left, right)
/* Partition ($V$, left, right) rearranges the sub-array $V[\text{left}\,..\,\text{right}]$ according
   to its first element $V[\text{left}]$, called the pivot, and returns the position
   of the pivot after partitioning                                              */
$v \leftarrow V[\text{left}]$;
$i \leftarrow \text{left}$;
$j \leftarrow \text{right}$;
repeat
    repeat
        $i \leftarrow i + 1$;
    until $V[i] \geqslant v$;
    repeat
        $j \leftarrow j - 1$;
    until $V[j] \leqslant v$;
    Swap($V[i], V[j]$);
until $j \leqslant i$;
Swap($V[i], V[\text{left}]$);
```

Fig. 3: Partition Algorithm.

```
Algorithm QuickSort(V, left, right)
/* Sorts the sub-array $V[\text{left}\,..\,\text{right}]$.                              */
/* Recursive function for an array $V[1\,..\,n]$: QuickSort ($V, 1, n$)              */
if left < right then
    $k \leftarrow$ Partition($V$, left, right);
    QuickSort ($V$, left, $k - 1$);
    QuickSort ($V$, $k + 1$, right);
```

Fig. 4: QuickSort Algorithm.

QuickSelect. The QuickSelect(m, n) algorithm described in Figure 5 returns the key of rank m in an array of size n. As previously, it uses the first key of the array as a pivot and performs the partition operation. If the rank k of the pivot equals m, the algorithm returns the pivot. If $k > m$, the algorithm continues with the left sub-array; otherwise, it continues with the right sub-array.

In the rest of this paper, we use here the notation

$$x \wedge y = \min(x, y), \qquad x \vee y = \max(x, y). \tag{2}$$

In the QuickSelect algorithm, as in the QuickSort algorithm, the keys U_i and U_j are compared if and only if U_i or U_j is the first key among the subset $\mathcal{U}_{[k,\ell]}$, with $k = i \wedge m$, $\ell = j \vee m$. Then, the mean number of key comparisons between U_i and U_j performed by QuickSelect(m, n) is, for $i < j$, equal to

$$\pi_n^{(m)}(i, j) = \frac{2}{\operatorname{card} \mathcal{U}_{[k,\ell]}} = \frac{2}{\ell - k + 1}, \qquad k = i \wedge m, \quad \ell = j \vee m.$$

```
Algorithm QuickSel(V, left, right, m)
/* Returns the element of rank m of the array V[left..right]           */
/* Recursive function for an array V[1..n]: QuickSel (V, 1, n, m)       */
k ← Partition(V, left, right);
if k > m then
    return QuickSel (V, left, k − 1, m);
else
    if k < m then
        return QuickSel (V, k, right, m − k);
    else
        return V[k];
```

```
Algorithm QuickVal(V, left, right, v)
/* Returns the rank of the element v in the array V[left..right].       */
/* Recursive function for an array V[1..n]: QuickVal (V, 1, n, v)       */
k ← Partition(V, left, right);
if V[k] > v then
    return QuickVal (V, left, k − 1, v);
else
    if V[k] < v then
        return QuickVal (V, k, right, v);
    else
        return k;
```

Fig. 5: Algorithms QuickSelect and QuickVal. Algorithms QuickMin and QuickMax correspond to the QuickSelect algorithm with particular ranks $m = 1$ and $m = n$ for an array of size n.

Observe that this probability depends on both i and j, except when $m = 1$ (case of QuickMin) or $m = n$ (case of QuickMax), where it satisfies

$$\pi_n^{(1)}(i,j) = \frac{2}{\operatorname{card}\mathcal{U}_{[1,j]}} = \frac{2}{j}, \qquad \pi_n^{(n)}(i,j) = \frac{2}{\operatorname{card}\mathcal{U}_{[i,n]}} = \frac{2}{n-i+1}.$$

We are also interested in two “variants” of the QuickSelect(m,n) algorithm:

(a) The QuickQuant$_\alpha(n)$ algorithm outputs the α-quantile, namely, the key of rank $\lfloor 1 + \alpha(n-1) \rfloor$. With this definition, we remark that for $\alpha = 0$, QuickQuant$_\alpha$ coincides with QuickMin, whereas for $\alpha = 1$, QuickQuant$_\alpha$ coincides with QuickMax.

(b) The QuickRand(n) algorithm outputs a key whose rank is chosen uniformly at random in the interval $[1\,..\,n]$.

QuickVal. We will be led (see Section 3.3) to consider another algorithm, named QuickVal, which may be of independent interest. Given an array of size n and a key which belongs to the array, it returns the rank of this key. This algorithm has the same structure as QuickSelect: it uses the first key of the array as a pivot and performs the partition operation. If the value of the pivot is equal to the input key, the algorithm returns the rank of the pivot.

If the value of the pivot is larger than the input key, the algorithm continues with the left sub-array; otherwise, it continues with the right sub-array. It is important to note that this algorithm rearrange the array exactly in the same way `QuickSelect` (using the partition algorithm).

About the `Partition` *process.* They are various possibilities for implementing this process. It is first important to remark that all the partition processes share the common following features: each element of the array $V[\text{left} \,..\, \text{right}]$ with $i \neq \text{left}$ is compared to the pivot $V[\text{left}]$ (and thus the number of comparisons is always greater or equal to $n-1$). At the end of the procedure, the pivot is at the right place, the keys that are smaller than the pivot are placed on its left, and keys that are greater than the pivot are placed on its right. However, the number of exchanges, and the final positions of the elements inside the two sub-arrays, the left one and the right, may depend on the precise implementation of the `Partition` procedure.

We now describe the implementation described in Figure 3, see also [23]. We choose the first element of the array $V[\text{left} \,..\, \text{right}]$ as the pivot, and the variable v holds the value of the pivot $V[\text{left}]$. We use two scan pointers i and j, the left pointer i scans from the left, and the right pointer j scans from the right. Each time we find a pair (i, j) such that $V[i] \geqslant v$ and $V[j] \leqslant v$, we exchange $V[i]$ and $V[j]$. When the pointers cross each other, an extra exchange of $V[i]$ and $V[j]$ is done with $j < i$ just after the pointers cross (but before the crossing is detected) and the outer **repeat** loop exited. The last three assignments implement the exchange between $V[i]$ and $V[j]$ (to undo the extra exchange), and between $V[i]$ and $V[\text{left}]$ to put the pivot into position.

2.3 *Sources and their Dirichlet series*

Here, we consider that the keys are words produced by a general source on a finite alphabet Σ. A general source $\mathcal{S}$ built on the alphabet Σ produces at each discrete time $t = 0, 1, \ldots$ a symbol from Σ. If X_n is the symbol emitted at time $t = n$, a source produces the infinite word $(X_0, X_1, \ldots, X_n, \ldots)$. For any finite prefix $w \in \Sigma^\star$, the probability p_w that a word produced by the source $\mathcal{S}$ begins with the finite prefix w is called the fundamental probability of prefix w. The set of probabilities p_w (for $w \in \Sigma^\star$) completely defines the source $\mathcal{S}$, and the two equalities hold

$$\text{for } k \geqslant 0 \sum_{w \in \Sigma^k} p_w = 1, \text{and for any } w \in \Sigma^\star \sum_{i \in \Sigma} p_{w \cdot i} = p_w.$$

Definition 3 Let Σ be a totally ordered alphabet of cardinality r. A source over the alphabet Σ produces infinite words of $\Sigma^{\mathbb{N}}$, and is specified by the *fundamental probabilities* p_w, $w \in \Sigma^\star$, where p_w is the probability that an infinite word begins with the finite prefix w. When the two following properties hold:

(i) $p_w > 0$ for any $w \in \Sigma^\star$,

(ii) $\pi_k = \max\{p_w : w \in \Sigma^k\}$ tends to 0 as $k \to \infty$,

the source is said to be *non-ambiguous.*

In the sequel, all the sources are assumed to be non-ambiguous.

As already mentioned in Section 1.2, our analyses deal with the two Dirichlet series of fundamental probabilities $\Lambda(s)$ and $\Pi(s)$ introduced in Definition 1.

2.4 *Instances of classical sources and their Dirichlet series.*

We describe here four models of sources, and begin with two models of simple sources, namely memoryless sources and Markov chains. We also give expressions for Dirichlet series $\Lambda(s)$ and $\Pi(s)$ in each case.

Memoryless sources. A *memoryless source* associated to the alphabet Σ is a source where the symbols $X_i \in \Sigma$ are independent and identically distributed. It is defined by the set $(p_j)_{j\in\Sigma}$ of probabilities, with $p_j = \mathbb{P}[X_k = j]$ for any $k \geqslant 0$. The Dirichlet series Λ and Λ_k can be expressed in terms of

$$\lambda(s) = \sum_{i\in\Sigma} p_i^s \qquad \text{as} \qquad \Lambda_k(s) = \lambda(s)^k, \quad \Lambda(s) = \frac{1}{1-\lambda(s)}. \tag{3}$$

The maximum probability π_k satisfies $\pi_k = \rho^k$ where $\rho = \max\{p_i : i \in \Sigma\}$, and the equality $\Pi(s) = 1/(1-\rho^s)$ holds.

Markov chains. A Markov chain associated to the finite alphabet Σ is defined by the (column) vector R of initial probabilities $(r_i)_{i\in\Sigma}$ together with the transition matrix $P = (p_{i,j})_{(i,j)\in\Sigma\times\Sigma}$, with each row summing to 1. Here one has $r_i = \mathbb{P}[X_0 = i]$ and $p_{i,j} = \mathbb{P}[X_{k+1} = j \mid X_k = i]$ for any $i, j \in \Sigma$ and $k \geqslant 0$. We denote by $P(s)$ the matrix with entry $p_{i,j}^s$ for general (i,j), and by $R(s)$ the vector with entries r_i^s. Then (with Q^T denoting the transpose of matrix Q)

$$\Lambda_k(s) = R(s)^T P(s)^{k-1}\mathbf{1} \text{ for } k \geqslant 1, \qquad \Lambda(s) = 1 + R(s)^T(I - P(s))^{-1}\mathbf{1}. \tag{4}$$

If, moreover, the matrix P is irreducible and aperiodic, then, for any real $s > 0$, the matrix $P(s)$ has a unique dominant eigenvalue $\lambda(s)$. The maximum probability π_k satisfies $\pi_k \leqslant \rho^k$ where ρ is the maximum of all the entries in R and P, and the inequality $|\Pi(s)| \leqslant 1/(1-\rho^\sigma)$ holds with $\sigma = \Re s$.

Other instances of "simple" sources: intermittent sources. Intermittent sources are an interesting particular case of a source of VLMC type (Variable Length Markov Chain), where the dependency from the past is unbounded (see [1]). An intermittent source has two regimes, depending on whether it emits a particular symbol $\underline{\sigma} \in \Sigma$ or not. Consider a source with an alphabet of finite cardinality $r \geqslant 2$. The source is *intermittent of exponent* $a > 0$ *with respect to* $\underline{\sigma}$ if one has the following conditional probability distribution for the emission of each symbol in the word given the prefix preceding it. Let us define the event $\mathcal{S}_k$ as

$\mathcal{S}_k = \{$the prefix finishes with a sequence of exactly k occurrences of $\underline{\sigma}\}$.

Then the conditional distribution of the next symbol emitted depends on the length k; more precisely, one has $\mathbb{P}[\sigma \mid \mathcal{S}_0] = 1/r$ and, for $k \geqslant 1$,

$$\mathbb{P}[\underline{\sigma} \mid \mathcal{S}_k] = \left(1 - \frac{1}{k+1}\right)^a, \quad \mathbb{P}[\sigma \mid \mathcal{S}_k] = \left(1 - \left(1 - \frac{1}{k+1}\right)^a\right) \frac{1}{r-1} \text{ for } \sigma \neq \underline{\sigma}.$$

Then, in the case of a binary alphabet $\Sigma = \{0, 1\}$, when the source is intermittent with respect to 0, the probability of the prefixes 0^k and 0^k1 are respectively equal to

$$p_{0^k} = \frac{1}{2}\left(\frac{1}{k^a}\right), \qquad p_{0^k1} = \frac{1}{2}\left(\frac{1}{k^a} - \frac{1}{(k+1)^a}\right),$$

and, with the equality $\{0,1\}^\star = (0^\star 1)^\star 0^\star$, the series $\Lambda(s)$ admits the expression

$$\Lambda(s) = \frac{1 + 2^{-s}\zeta(as)}{1 - 2^{-s}[1 + \Sigma_a(s)]} \qquad \text{with} \quad \Sigma_a(s) = \sum_{k \geqslant 1} \left[\frac{1}{k^a} - \frac{1}{(k+1)^a}\right]^s.$$

(Here, $\zeta(\cdot)$ is the Riemann zeta function.) Moreover, for $a \geqslant 1$, and $k \geqslant 1$, the maximum probability π_k is attained for the word 0^k; it satisfies $\pi_0 = 1$, and

$$\pi_k = \frac{1}{2}\left(\frac{1}{k^a}\right) \text{ for } k \geqslant 1, \qquad \text{so that} \quad \Pi(s) = 1 + \left(\frac{1}{2}\right)^s \zeta(as).$$

Dynamical sources. An important subclass is formed by *dynamical sources*, which are closely related to dynamical systems on the interval; see [25]. One starts with a partition $\{\mathcal{I}_\sigma\}$ indexed by symbols $\sigma \in \Sigma$, a coding map $\tau : \mathcal{I} \to \Sigma$ which equals σ on $\mathcal{I}_\sigma$, and a shift map $T : \mathcal{I} \to \mathcal{I}$ whose restriction to each $\mathcal{I}_\sigma$ is increasing, invertible, and of class $\mathcal{C}^2$. Then the word $M(u)$ is the word that encodes the trajectory $(u, Tu, T^2u \ldots)$ via the coding map τ, namely, $M(u) = (\tau(u), \tau(Tu), \tau(T^2u), \ldots)$. All memoryless (Bernoulli) sources and all Markov chain sources can be viewed as dynamical sources: they correspond to a piecewise linear shift. For instance, the standard binary system is obtained by $T(x) = \{2x\}$ ($\{\cdot\}$ is the fractional part). Dynamical sources with a non-linear shift allow for correlations that depend on the entire past (e.g., sources related to continued fractions obtained by $T(x) = \{1/x\}$). Moreover, the Λ-series can be expressed as the quasi-inverse of a variant of the transfer operator of the system [25], and this gives a precise tool for investigating the tameness of such a model of sources.

2.5 *Return to tameness of sources.*

We recall that tameness of sources is described via the abscissa of absolute convergence of series $\Lambda(s)$ and $\Pi(s)$. The series $\Lambda(s)$ always satisfies $\Lambda_k(1) = 1$ and thus $\Lambda(1) = +\infty$; its abscissa of convergence is then at least 1, and we define it as $1 + \delta_\Lambda$ where $\delta_\Lambda \geqslant 0$ is the width of Λ. In the same vein, the series $\Pi(s)$ always satisfies $\pi_k^0 = 1$ and thus $\Pi(0) = +\infty$; its abscissa of convergence is then at least 0. and we define it as δ_Π where $\delta_\Pi \geqslant 0$ is the width of Π. Furthermore, the inequality $\Lambda(s) \leqslant \Pi(s-1)$ for real $s \geqslant 1$ entails the inequality $\delta_\Pi \geqslant \delta_\Lambda$ between the two widths.

In general, the simple sources (memoryless sources, aperiodic Markov chains or good dynamical sources) are both weakly Λ-tame and Π-tame with both widths $\delta_\Pi = \delta_\Lambda = 0$. However, there exist “natural” sources, for example intermittent sources described in Section 2.4 which are Π-tame with width $\delta_\Pi = 1/a$ and weakly Λ-tame with $\delta_\Lambda = 0$ as soon as $a > 1$. In the following, the condition under which our results hold is

$$\max(2\delta_\Lambda, \delta_\Pi) < 1,$$

and it is satisfied for all the examples of sources which are described in the previous section, *except* for the *intermittent* source with $a = 1$.

2.6 *Parameterization of a source.*

The parameterization of a source is based on the same principles as those used for arithmetic coding in compression frameworks. It also aims to extend what we have done in the case of a dynamical system. We have associated to a real $u \in \mathcal{I}$ a word $M(u)$ that is the encoding of the trajectory $\mathcal{T}(u)$. We wish also to build a family of fundamental intervals $\mathcal{I}_w$ whose measures are the fundamental probabilities of the source. Then, for each depth k, the fundamental intervals $\mathcal{I}_w$ will form a topological partition of the interval $\mathcal{I}$.

For any prefix $w \in \Sigma^\star$, we denote by $|w|$ the length of w (i.e., the number of the symbols that it contains) and $p_w^{(-)}$, $p_w^{(+)}$, p_w the probabilities that a word produced by the source begins with a prefix u of the same length as w, which satisfies $u \prec w$, $u \succ w$, or $u = w$, meaning

$$p_w^{(-)} = \sum_{\substack{u:\ |u|=|w|,\\ u\prec w}} p_u, \qquad p_w^{(+)} = \sum_{\substack{u:\ |u|=|w|,\\ u\succ w}} p_u, \qquad p_w = 1 - \left(p_w^{(+)} + p_w^{(-)}\right). \quad (5)$$

We also use alternative notations for these probabilities, especially when we represent them on the interval $[0,1]$, and we let

$$a_w = p_w^{(-)} = \sum_{\substack{u:\ |u|=|w|\\ u\prec w}} p_u, \qquad b_w = a_w + p_w = \sum_{\substack{u:\ |u|=|w|,\\ u\preceq w}} p_u = 1 - p_w^{(+)}. \quad (6)$$

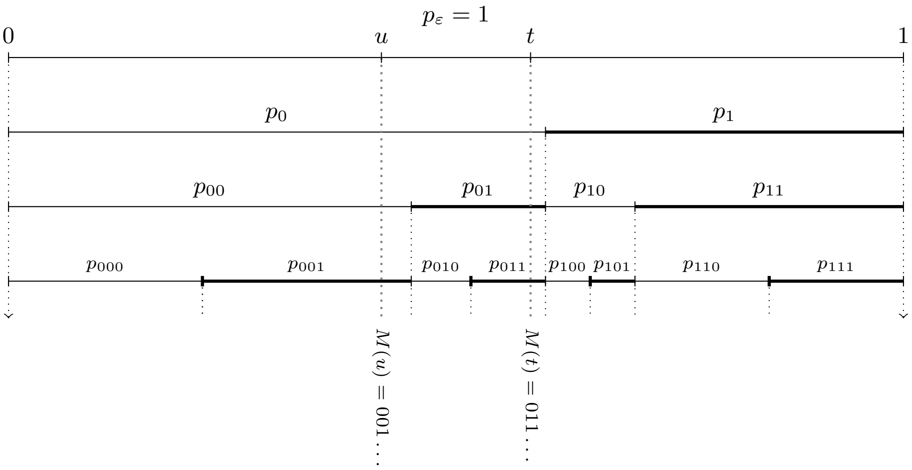


Fig. 6: The parameterization of a source.

Thus, for a given k, when the prefix w varies in Σ^k, this gives rise to a partition of the unit interval with sub-intervals of length p_w. When the prefixes $w \in \Sigma^k$ are ordered in increasing lexicographic order, and the sub-intervals are arranged from left to right, then, the sub-interval corresponding to prefix w has left (respectively, right) endpoint equal to a_w (resp., b_w). See Figure 6.

Consider the set $\Sigma^{\mathbb{N}}$ of (infinite) words produced by the source $\mathcal{S}$, ordered via the lexicographic order. Given an infinite word $X \in \Sigma^{\mathbb{N}}$, denote by w_k its prefix of length k. The sequence $(a_{w_k})_{k \geqslant 0}$ is increasing, the sequence $(b_{w_k})_{k \geqslant 0}$ is decreasing, and $b_{w_k} - a_{w_k} = p_{w_k}$ tends to 0 when k tends to infinity. Thus a unique real $P(X) \in [0,1]$ is defined as the common limit of (a_{w_k}) and (b_{w_k}), and $P(X)$ is simply the probability that an infinite word Y generated by the source is smaller than X. The mapping $P : \Sigma^{\mathbb{N}} \to [0,1]$ is surjective and strictly increasing outside the exceptional set formed with words of $\Sigma^{\mathbb{N}}$ which end with an infinite sequence of the smallest symbol or with an infinite sequence of the largest symbol.

Conversely, almost everywhere (except on the set $\{a_w, w \in \Sigma^\star\}$), there is a mapping M which associates to a number u from the interval $\mathcal{I} = [0,1]$ a word $M(u) \in \Sigma^{\mathbb{N}}$. Hence, the probability that a word Y be smaller than $M(u)$ equals u. The lexicographic order on words ('$\prec$') is then compatible with the natural order on the interval $\mathcal{I}$, namely, $M(t) \prec M(u)$ if and only if $t < u$. The interval $\mathcal{I}_w = [a_w, b_w]$, of length p_w, gathers (up to a denumerable set) all the reals u for which $M(u)$ begins with the finite prefix w. This is the fundamental interval of the prefix w.

Reflected source. Such a source will be used in the proofs of Propositions 8 and 9. The reflected source $\check{\mathcal{S}}$ of a source $\mathcal{S}$ is defined on the same alphabet $\Sigma := [0\,..\,r-1]$. However, this alphabet is ordered with the "reflected" order; namely, we consider the mapping $\Sigma \to \Sigma$ which maps the symbol σ to the symbol

$\breve{\sigma} = (r-1) - \sigma$. This mapping is then extended to $\Sigma^\star$ using concatenation: for any $w \in \Sigma^\star$, of the form $w = w_1 \cdot w_2 \dots w_k$, the word $\breve{w}$ is defined as $\breve{w} = \breve{w}_1 \cdot \breve{w}_2 \dots \breve{w}_k$. This is further extended to $\Sigma^{\mathbb{N}}$ in a similar way, and gives rises to a parameterization $\breve{M}$ for the source $\breve{\mathcal{S}}$ which is closely related to the initial parameterization of the source $\mathcal{S}$ via the equality $\breve{M}(u) = M(1-u)$. Note that the source $\mathcal{S}$ and its reflected source $\breve{\mathcal{S}}$ have, for each k, the same set $\{p_w : w \in \Sigma^k\}$ of fundamental probabilities for prefixes of length k, and thus have the same Dirichlet series Λ and Π.

2.7 *Geometry of the source and coincidence*

We are interested in a more realistic cost related to the number of symbol comparisons performed by these algorithms, when the keys are words independently produced by the same source. The words are ordered with respect to the lexicographic order, and the cost for comparing two words (measured as the number of symbol comparisons needed) is closely related to the coincidence, defined as follows.

Definition 4 The *coincidence function* $\gamma(u,t) : [0,1] \times [0,1] \to \mathbb{N} \cup \{+\infty\}$ is the length of the largest common prefix of words $M(u)$ and $M(t)$.

More precisely, the realistic cost of the comparison between $M(u)$ and $M(t)$ equals $\gamma(u,t) + 1$.

We represent the pair of words $(M(u), M(t))$ with $u \leqslant t$ by the point (u,t) of the triangle $\mathcal{T} = \{(u,t) : 0 \leqslant u \leqslant t \leqslant 1\}$, and the *fundamental triangles*

$$\mathcal{T}_w = (\mathcal{I}_w \times \mathcal{I}_w) \cap \mathcal{T} = \{(u,t) : a_w \leqslant u \leqslant t \leqslant b_w\} \tag{7}$$

define the *level sets* of the function γ. Indeed, the coincidence $\gamma(u,t)$ is at least ℓ if and only if $M(u)$ and $M(t)$ have the same common prefix w of length ℓ, so that the parameters u and t belong to the same fundamental interval $\mathcal{I}_w$ relative to a prefix w of length ℓ. Then, the two relations

$$\mathcal{T} \cap [\gamma \geqslant \ell] = \bigcup_{w \in \Sigma^\ell} \mathcal{T}_w, \qquad \sum_{\ell \geqslant 0} \mathbf{1}_{[\gamma \geqslant \ell]} = \sum_{\ell \geqslant 0} (\ell+1) \mathbf{1}_{[\gamma = \ell]},$$

entail the following equality which deals with the functional $\mathcal{J}$ and holds for any integrable function g on the unit triangle $\mathcal{T}$,

$$\mathcal{J}[g] = \int_{\mathcal{T}} [\gamma(u,t)+1] g(u,t)\, du\, dt = \sum_{w \in \Sigma^\star} \int_{\mathcal{T}_w} g(u,t)\, du\, dt. \tag{8}$$

This functional $\mathcal{J}$ will be extensively used in the sequel.

The family of triangles $\mathcal{T}_w$, defines the "geometry" of the source. Figure 7 represents the geometry of two memoryless sources.

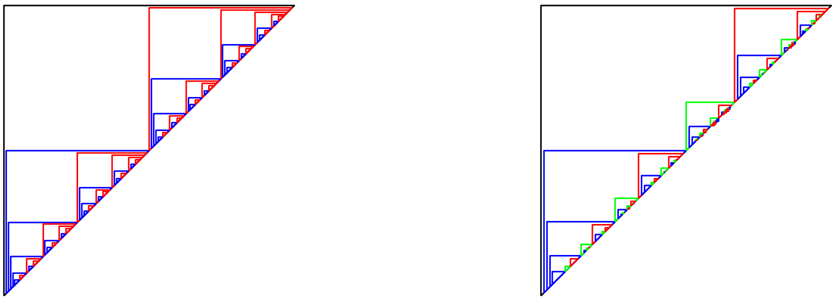


Fig. 7: The geometry of two memoryless sources. On the left, the case of $\Sigma = \{a, b\}$ with $p_a = p_b = 1/2$. On the right, the case of $\Sigma = \{a, b, c\}$ with $p_a = 1/2, p_b = 1/6, p_c = 1/3$.

3 Algebraic analysis.

We are now ready to begin our probabilistic analysis. The set of inputs is the set $\mathcal{M}^\star$ of all the sequences of words emitted by the source, and all the costs of interest (for instance the number of key or symbol comparisons) are defined on $\mathcal{M}^\star$. The purpose of average-case analysis of structures (or algorithms) is to characterise the mean value of their "costs" under a well-defined probabilistic model that describes the initial distribution of its inputs.

3.1 ***Two probabilistic models.***

The two models. Here, the set of inputs for the algorithm is the set of all the sequences of words $\mathcal{V}$ emitted by the source, denoted here by $\mathcal{M}^\star$, and all the costs are defined on $\mathcal{M}^\star$. For such a cost R, we denote by $R[\mathcal{V}]$ its value on the sequence $\mathcal{V}$.

We consider two distinct probabilistic models on the set $\mathcal{M}^\star$, each of them being of interest. When the cardinality n of $\mathcal{V}$ is fixed, and words $V_i \in \mathcal{V}$ are independently emitted by the source $\mathcal{S}$, this is the Bernoulli model denoted by $(\mathcal{B}_n, \mathcal{S})$. There is also another model, where the sequence $\mathcal{V}$ has a variable number N of elements that obeys a Poisson law of rate Z,

$$\mathbb{P}[N = k] = e^{-Z} \frac{Z^k}{k!}.$$

In this model, called the Poisson model of rate Z, the rate Z plays a role quite similar to the cardinality of $\mathcal{V}$. When it is relative to probabilistic source $\mathcal{S}$, the model, denoted by $(\mathcal{P}_Z, \mathcal{S})$, is composed with two main steps:

(a) The number N of words is drawn according to the Poisson law of rate Z;
(b) Then the N words are independently drawn from the source $\mathcal{S}$.

Random variables N_I. Thanks to the mapping M, we can associate with a sequence $\mathcal{V} = (V_1, \ldots, V_n)$ of infinite words a sequence of real points $(v_1, \ldots, v_n)$

in $[0,1]$ such that $V_i = M(v_i)$ for $i \in [1\,..\,n]$. This correspondence exists for almost all sequences in $\mathcal{M}^\star$ (up to a set of sequences of zero measure). For $I \subset [0,1]$, consider the variable N_I, defined on $\mathcal{M}^\star$ by the equality

$$N_I[\mathcal{V}] = \text{the number of words in } \mathcal{V} \text{ whose parameter lies in the interval } I.$$

We have of course $N := N_{[0,1]}$. The random variables N_I are easier to deal with in the Poisson model, due to very useful independence properties. In the Poisson model of rate Z, for two disjoint intervals I and J in $[0,1]$, the random variables N_I and N_J are both independent Poisson variables with adequate rates: For any interval $I \subset [0,1]$ of measure λ, N_I obeys a Poisson law of rate λZ.

This is very often the case in our study: *the Poisson model is easier to deal with, but the Bernoulli model is more natural.* So, we begin the analysis in the Poisson model, and then we wish to return to the Bernoulli model.

Notation. For a random variable R defined on the set $\mathcal{M}^\star$ and a sequence $\mathcal{V} \in \mathcal{M}^\star$, we denote by $R[\mathcal{V}]$ the value of R on $\mathcal{V}$ and by $R\langle\mathcal{V}\rangle$ the average of the cost R on all sequences equivalent to $\mathcal{V}$ up to a permutation

$$R\langle\mathcal{V}\rangle = \frac{1}{|\mathcal{V}|!} \sum_{\mathcal{V}' \equiv \mathcal{V}} R[\mathcal{V}'],$$

where, again, the equivalence '$\equiv$' relation means up to a permutation. We note that, when R depends only on the set underlying the sequence $\mathcal{V}$, the equality $R[\mathcal{V}] = R\langle\mathcal{V}\rangle$ holds. This is the case for the variables N_I aforementioned.

Relation between the two models. Consider a random variable R defined on the set $\mathcal{M}^\star$. Denote by $R(n)$ its expectation in the Bernoulli model $(\mathcal{B}_n, \mathcal{S})$, and by R_Z its expectation in the Poisson model $(\mathcal{P}_Z, \mathcal{S})$ of rate Z. Using conditional probabilities, the following equalities relate the two types of expectations,

$$R_Z = \sum_{n \geqslant 0} \mathbb{E}_Z[R \mid N = n]\, \mathbb{P}_Z[N = n] = e^{-Z} \sum_{n \geqslant 0} \mathbb{E}_{[n]}[R] \frac{Z^n}{n!} = e^{-Z} \sum_{n \geqslant 0} R(n) \frac{Z^n}{n!},$$

and the Poisson expectation R_Z can be expressed with the exponential generating function $G(Z)$ of the sequence $R(n)$ of the Bernoulli expectations via the relation

$$R_Z = e^{-Z} G(Z), \qquad \text{with} \quad G(Z) = \sum_{n \geqslant 0} R(n) \frac{Z^n}{n!}. \tag{9}$$

Assume now that the expectation R_Z of the random variable R is known in the Poisson model, via its coefficients $\varphi(n)$ in its series expansion,

$$\varphi(n) = (-1)^n [Z^n] R_Z, \quad \text{with the expansion} \quad R_Z = \sum_{n \geqslant 0} (-1)^n \varphi(n) \frac{Z^n}{n!}. \tag{10}$$

The equality (9) between the series R_Z and $G(Z)$ is also written as $G(Z) = e^Z R_Z$, and entails the following *binomial relation* between the coefficients $R(n)$ and $\varphi(n)$:

$$R(n) = \sum_{k \leqslant n} \binom{n}{k} (-1)^k \varphi(k). \tag{11}$$

It is then possible to return to the Bernoulli model $(\mathcal{B}_n, \mathcal{S})$ and obtain the expectation $R(n)$ in the Bernoulli model, that is expressed as a binomial recurrence involving the coefficients $\varphi(k)$ of the Poisson expectation R_Z.

3.2 ***When the keys are words.***

Now, the keys are words, and we denote by $\mathcal{M} := \Sigma^{\mathbb{N}}$ the set of all the words emitted by the source. The set of inputs for the algorithm is thus the set $\mathcal{M}^\star$ which gathers all the possible sequences of words. Consider a sequence $\mathcal{V} \in \mathcal{M}^\star$ that contains two given words $M(u)$ and $M(t)$. The following equalities hold:

$$\mathrm{Rank}_{\mathcal{V}}\, M(u) = N_{[0,u[}[\mathcal{V}] + 1, \qquad \mathrm{Rank}_{\mathcal{V}}\, M(t) = N_{[0,u[}[\mathcal{V}] + N_{]u,t[}[\mathcal{V}] + 2,$$

where the respective translations of 1 and 2 express that $M(u)$ and $M(t)$ belong to $\mathcal{V}$.

We denote by $\Phi(u,t)[\mathcal{V}]$ the number of key comparisons performed by the algorithm on the input sequence $\mathcal{V}$ between the keys $M(u)$ and $M(t)$. This defines a cost on $\mathcal{M}^\star$. We also consider the *mean* $\Phi(u,t)\langle\mathcal{V}\rangle$ of the cost taken with respect to all the permutations of $\mathcal{V}$. There is a relation between $\Phi(u,t)\langle\mathcal{V}\rangle$ and the mean number $\pi_N(i,j)$ defined in Section 2.1. More precisely, the following equality holds:

$$\Phi(u,t)\langle\mathcal{V}\rangle = \pi_{|\mathcal{V}|}(N_{[0,u[}[\mathcal{V}] + 1, N_{[0,u[}[\mathcal{V}] + N_{]u,t[}[\mathcal{V}] + 2). \tag{12}$$

We remark that the right member of (12) is defined on the whole set $\mathcal{M}^\star$ (even for sequences which do not contain $M(u)$ or $M(t)$). Setting $N = N_{[0,1]}$, the function

$$\widehat{\pi}(u,t) := \pi_N(N_{[0,u[} + 1, N_{[0,u[} + N_{]u,t[} + 2) \tag{13}$$

provides thus an extension of the function $\mathcal{V} \mapsto \Phi(u,t)\langle\mathcal{V}\rangle$ on the set $\mathcal{M}^\star$. This expression plays a central role in our analyses.

Applying these principles, with expressions of $\pi_n(i,j)$ obtained in Section 2.2, we obtain an expression for random variable $\widehat{\pi}_\pm(u,t)$ associated with `QuickMin` and `QuickMax`, namely

$$\widehat{\pi}_-(u,t) = \frac{2}{N_{[0,t]}}, \qquad \widehat{\pi}_+(u,t) = \frac{2}{N_{[u,1]}}$$

$$\widehat{\pi}_-(u,t) = \frac{2}{N_{[0,u[} + N_{]u,t[} + 2}, \qquad \widehat{\pi}_+(u,t) = \frac{2}{N_{]u,t[} + N_{]t,1]} + 2} \tag{14}$$

For the algorithm `QuickSelect` the mean number of comparisons $\widehat{\pi}_\alpha(u,t)$ performed by the `QuickSelect` algorithm between the words $M(u)$ and $M(t)$ *chosen as keys* when it searches the word $M(\alpha)$ is

$$\widehat{\pi}_\alpha(u,t) = \frac{2}{N_{[x,y]}} \qquad \text{with} \quad x := u \wedge \alpha, \quad y := t \vee \alpha,$$

with the usual notation of (2). We denote by z the "middle point" (i.e., $x \leqslant z \leqslant y$), so that the equality $\{x, y, z\} = \{\alpha, u, t\}$ holds.

(i) In the case when $\alpha \neq u$ and $\alpha \neq t$, the three words $M(x)$, $M(y)$, and $M(z)$ are distinct and, in the same vein as previously, $N_{[x,y]} = N_{]x,z[} + N_{]z,y[} + 3$ where the translation of 3 expresses that the three words $M(x), M(y), M(z)$ belong to $\mathcal{V}$.

(ii) In the case when $\alpha = u$ or $\alpha = t$, there are only two distinct words amongst $M(x)$, $M(y)$, and $M(z)$.

In fact, we will be interested only in the first case (i), as the second case arises when (u,t) belongs to the union of the two lines $u = \alpha$ and $t = \alpha$, which will be of measure 0 in the triangle $\mathcal{T}$. This is why we focus only on the case (i) where, with $x := u \wedge \alpha$, $y := t \vee \alpha$, $\{x, y, z\} = \{\alpha, u, t\}$, we obtain for $\widehat{\pi}_\alpha(u,t)$,

$$\widehat{\pi}_\alpha(u,t) = \frac{2}{N_{]x,z[} + N_{]z,y[} + 3}, \qquad \text{when } u \neq \alpha \text{ and } t \neq \alpha \tag{15}$$

3.3 *The* `QuickVal` *algorithm*

We remark that Eq. (15) does not provide a clear way for computing the function $\widehat{\pi}(u,t)$ of the `QuickSelect`(m,n) algorithm, for a general rank m. This is due to the fact that the interval $[x,y]$ depends on the parameter α of the word of rank m, which is not directly related to the rank m. In fact, the real α is mostly the output of the `QuickSelect` algorithm.

This is why we consider an algorithm (already described in Figure 5) that is the counterpart of `QuickSelect`: it takes as input a set of words $\mathcal{V}$ and a given *word* $V \in \mathcal{V}$, and returns the *rank* of V inside the set $\mathcal{V}$. This algorithm is of independent interest and is easily implemented as a variant of `QuickSelect` by resorting to the usual partitioning phase, then doing a comparison between the *value* of the pivot and the input *word* V (rather than a comparison between their ranks). We call this algorithm `QuickVal`$_\alpha$ when it is used to seek the rank of the word $M(\alpha)$. By definition, the two algorithms `QuickSelect`(m,n) and `QuickVal`$_\alpha$ behave exactly in the same way when the rank of the word $M(\alpha)$ equals m. Now, with (15), one computes the mean number of key comparisons $\widehat{\pi}_\alpha(u,t)$ performed by `QuickVal`$_\alpha$ between two words $M(u)$ and $M(t)$.

If we consider the algorithm `QuickQuant`$_\alpha(n)$ which outputs the key of rank $\lfloor 1 + (n-1)\alpha \rfloor = \lfloor n\alpha + 1 - \alpha \rfloor$, then, for $\alpha = 0$, we recover `QuickMin`, and, for $\alpha = 1$, we recover `QuickMax`. For a general $\alpha \in]0,1[$, the behaviors of the two algorithms of $\texttt{QuickVal}_\alpha(n)$ and $\texttt{QuickQuant}_\alpha(n)$ should be asymptotically similar. Indeed, the parameter of the α-*quantile* of a random set of words of large enough cardinality must be, with high probability, close to α. It is proven in [17] that this parameter follows asymptotically a Gaussian law with mean α and variance $\alpha(1-\alpha)/n$. We will return to this question in Section 5.

3.4 *Plan of our approach.*

We deal with the random variable S counting the number of symbol comparisons which is defined on the set $\mathcal{M}^\star$: for each $\mathcal{V} \in \mathcal{M}^\star$, $S[\mathcal{V}]$ is the number of symbol comparisons performed by the algorithm on the input $\mathcal{V}$. We are interested in its expectation $S(n)$ in the Bernoulli model, but it proves convenient to first perform the study in the Poisson model, and we denote by S_Z the expectation in the Poisson model. This study has four main steps:

(a) We begin by studying the *density* of the algorithm in the Poisson model, which deals with the *mean number of local key comparisons* performed by the algorithm on pairs of words whose parameters are close to (u, t).

(b) With the *coincidence* of the source described by $\gamma(u, t)$, we obtain information about the *mean number of local symbol comparisons* on pairs of words whose parameters are close to (u, t).

(c) Integrating on all possible points (u, t) of triangle $\mathcal{T}$ gives rise to the mean number S_Z.

(d) We return to $S(n)$ using (11). The binomial expression involves the coefficients $\varphi(k)$ of the series S_Z, and the mixed Dirichlet series $\varpi(s)$ will be defined as the analytic lifting of the sequence $\varphi(k)$.

For reference, the main objects used in the analysis are gathered in Table 2.

3.5 *Density: a first expression in the Poisson model.*

The density of an algorithm is the expectation of the random variable $\Phi(u, t)$, defined in Section 3.2, which associates with a random sequence $\mathcal{V} \in \mathcal{M}^\star$ the number of key comparisons $\Phi(u, t)[\mathcal{V}]$ performed by the algorithm on the input sequence $\mathcal{V}$ between the keys $M(u)$ and $M(t)$.

Definition 5 (Poissonized density) For a given algorithm, and a parameterized source $\mathcal{S}$, the Poissonized density $\Phi_Z(u, t)$ is the expectation of the variable $\Phi(u, t)$ in the Poisson model $(\mathcal{P}_Z, \mathcal{S})$.

We remark that the density $\Phi_Z(u, t)$ can be also defined by the relation

> $\Phi_Z(u, t)\, du\, dt$ is the mean number of key comparisons in the Poisson model performed by the algorithm between two words $M(u')$ and $M(t')$ with parameters $u' \in [u - du, u]$ and $t' \in [t, t + dt]$.

Objects of the analysis	Definition	Computation
$\pi_n(i,j)$	Mean number of key comparisons between the keys of rank i and j	Sec. 2.1
$\Phi(u,t)$	Number of key comparisons between the words of parameters u and t	Eq. (12)
$\widehat{\pi}(u,t)$	Extension of the random variable $\Phi(u,t)$	Eq. (13)
$\Phi_Z(u,t)$	Density of the algorithm at (u,t) in the model $\mathcal{P}_Z$	Eq. (16)
$\varphi(k,u,t)$	k-th coefficient in the series expansion of $\Phi_Z(u,t)$ i.e., $\varphi(k,u,t) = (-1)^k k![Z^k]\Phi_Z(u,t)$	Prop. 2
S_Z	Mean number of symbol comparisons in the model $\mathcal{P}_Z$	Eq. (21)
$\varphi(k)$	k-th Coefficient of $\widehat{S}_Z$, i.e., $\varphi(k) = (-1)^k k![Z^k]S_Z$	Eq. (23)
$S(n)$	Exact mean number of symbol comparisons in the model $\mathcal{B}_n$	Eq. (24)
$\varpi(s)$	Analytic lifting of the sequence $\varphi(k)$, also called the mixed Dirichlet series (algorithm/source)	Sec. 4.2

Table 2: The main objects used in the analysis, with their names, their definitions, and their computations.

It is then possible to relate the Poissonized density and the expectation of the random variable $\widehat{\pi}(u,t)$ defined in (13) which provides an extension of the variable $\Phi(u,t)$. This is due to the nice properties of this model, and this is why we choose the Poisson model as our (first) probabilistic model.

Lemma 2 *The Poissonized density $\Phi_Z(u,t)$ satisfies*

$$\Phi_Z(u,t) = Z^2 \cdot \mathbb{E}_Z\left[\widehat{\pi}(u,t)\right] \tag{16}$$

and involves the random variable $\widehat{\pi}(u,t)$ defined as

$$\widehat{\pi}(u,t) = \pi_N(N_{[0,u[} + 1, N_{[0,u[} + N_{]u,t[} + 2)\,.$$

Proof First, we recall a notation which has been already used in Section 3.2. For a random variable R defined on the set $\mathcal{M}^\star$ and a sequence $\mathcal{V} \in \mathcal{M}^\star$, we denote by $R[\mathcal{V}]$ the value of R on $\mathcal{V}$ and by $R\langle\mathcal{V}\rangle$ the average of the cost R on all sequences equivalent to $\mathcal{V}$ up to a permutation.

We consider the "interesting" sequences $K \in \mathcal{M}^\star$ for which

$$K \cap M[u-du,u] \neq \varnothing, \qquad K \cap M[t,t+dt] \neq \varnothing,$$

and the subset $\mathcal{K} \subset \mathcal{M}^\star$ which gathers such sequences, that contain a pair $(M(u'),M(t'))$ with $u' \in [u-du,u]$ and $t' \in [t,t+dt]$. Gathering all the sequences $\mathcal{V}$ corresponding to the same set gives rise to the expression

$$\Phi_Z(u,t) = \int_{\mathcal{M}^\star} \Phi(u,t)\langle\mathcal{V}\rangle\, d_Z\mathcal{V}\,.$$

Now, as the variable $\widehat{\pi}(u,t)$ provides an extension of the random variable $\mathcal{V} \mapsto \Phi(u,t)\,\langle\mathcal{V}\rangle$ outside $\mathcal{K}$ which only depends on the underlying set, the following equalities hold:

$$\Phi(u,t)\langle\mathcal{V}\rangle\,du\,dt = \mathbf{1}_{\mathcal{K}}\langle\mathcal{V}\rangle \cdot \widehat{\pi}(u,t)\langle\mathcal{V}\rangle = \mathbf{1}_{\mathcal{K}}[\mathcal{V}] \cdot \widehat{\pi}(u,t)[\mathcal{V}],$$

where the last equality comes from the fact that $\mathbf{1}_{\mathcal{K}}[\mathcal{V}]$ and $\widehat{\pi}(u,t)[\mathcal{V}]$ depend only on the underlying set of the sequence $\mathcal{V}$. This yields another expression for $\Phi_Z(u,t)\,d\,dt$, namely,

$$\Phi_Z(u,t)\,du\,dt = \int_{\mathcal{M}^\star} \mathbf{1}_{\mathcal{K}}[\mathcal{V}] \cdot \widehat{\pi}(u,t)[\mathcal{V}]\,.$$

Observe that the variable $\widehat{\pi}(u,t)$ depends only on the intervals $[0,u[$ and $]u,t[$, whereas $\mathbf{1}_{\mathcal{K}}$ depends only on the intervals $[u-du,u]$ and $[t,t+dt]$. Then, in the Poisson model, the two functions are independent, and we finally get

$$\Phi_Z(u,t)\,du\,dt = \left(\int_{\mathcal{M}^\star} \mathbf{1}_{\mathcal{K}}[\mathcal{V}]\,d_Z\mathcal{V}\right) \cdot \left(\int_{\mathcal{M}^\star} \widehat{\pi}(u,t)[\mathcal{V}]d_Z\mathcal{V}\right).$$

As the two intervals $[u-du,u]$ and $[t,t+dt]$ are disjoint, by the independence property of the Poisson model the first integral equals

$$\mathbb{E}_Z[\mathbf{1}_{\mathcal{K}}] = \mathbb{P}_Z[\mathcal{K}] = Z\,du \cdot Z\,dt = Z^2\,du\,dt,$$

and the lemma is proven. □

3.6 *Density: a second expression in the Poisson model.*

In order to return to the Bernoulli model, we are interested in the coefficients of the series expansion of $\Phi_Z(u,t)$ defined as

$$\Phi_Z(u,t) = \sum_{k\geqslant 2}(-1)^k \frac{Z^k}{k!}\varphi(k,u,t), \qquad \varphi(k,u,t) := (-1)^k k![Z^k]\Phi_Z(u,t) \quad (17)$$

and, with (16), the coefficients $\varphi(k,u,t)$ satisfy $\varphi(k,u,t)=0$ for $k=0,1$.

The following Lemma provides nice expressions for these coefficients for three algorithms:

Lemma 3 *For each algorithm* `QuickMin`, `QuickMax` *and* `QuickVal`$_\alpha$, *the sequence of functions* $\varphi(k,u,t)$ *satisfies, for* $k \geqslant 2$,

$$\varphi_-(k,u,t) = 2t^{k-2}, \quad \varphi_+(k,u,t) = 2(1-u)^{k-2},$$

and

$$\varphi_\alpha(k,u,t) = \frac{4}{k+1}(t \vee \alpha - u \wedge \alpha)^{k-2} \quad \textit{if } u \neq \alpha \textit{ and } t \neq \alpha\,.$$

The proof of Lemma 3 is based on two assertions:

(i) For any of the three algorithms, the expressions obtained in (14) and (15) show that the functions $\widehat{\pi}(u,t)$ that arise is our framework are always expressed as a linear combination of basic random variables $g_m(N_I)$ with

$$g_m(X) := \frac{1}{(X+1)(X+2)\dots(X+m)} \quad (m \geqslant 1),$$

and I an interval $I \subset [0,1]$. Here, as before, N_I is a variable defined on $\mathcal{M}^\star$ as the number of words whose parameter belongs to I.

(ii) It is easy to compute the coefficients of the expectation $\mathbb{E}_Z[g_m(N_{\mathcal{J}})]$, as the following proposition shows.

Proposition 2 *Consider an integer $m \geqslant 1$ and an interval $I \subset \mathcal{I} = [0,1]$. The expectation in the Poisson model $(\mathcal{P}_Z, \mathcal{S})$ of the random variable $g_m(N_I)$ only depends on m and λZ where λ is the length of the interval I. It is denoted by $F_m(\lambda Z)$. Moreover, for $k \geqslant 2$, the sequence*

$$\beta_m(k,\lambda) = (-1)^k k! [Z^k] \left(Z^2\, F_m(\lambda Z)\right)$$

admits the following expression:

$$\beta_m(k,\lambda) = \frac{1}{(m-1)!} \frac{k(k-1)}{k+m-2} \lambda^{k-2}. \tag{18}$$

Proof Using the independence property of the Poisson model, we know that N_I follows a Poisson law of parameter λZ. Then, $\mathbb{E}_Z[g_m(N_I)]$ only depends on m and λZ, and

$$\mathbb{E}_Z[g_m(N_I)] = \sum_{k\geqslant 0} g_m(k)\, \mathbb{P}_Z[N_I = k] = e^{-\lambda Z} \sum_{k \geqslant 0} g_m(k) \frac{(\lambda Z)^k}{k!} := F_m(\lambda Z).$$

The coefficients $\alpha_m(k,\lambda)$ in the series expansion of $F_m(Z)$ satisfy

$$\alpha_m(k) := (-1)^k k! [Z^k] F_m(Z) = \frac{1}{(m-1)!} \frac{1}{k+m}. \tag{19}$$

Then, the coefficients $\beta_m(k,\lambda)$ are related to $\alpha_m(k)$ for $k \geqslant 2$:

$$\beta_m(k,\lambda) = k(k-1)\, \lambda^{k-2} \alpha_m(k-2),$$

this proves, with the help of (19), the expressions in Eq. (18) and finally the result. □

We now use Proposition 2 to prove Lemma 3.

Proof First, we observe the decompositions

$$\frac{1}{X+2} = g_1(X) - g_2(X), \qquad \frac{1}{X+3} = 2g_3(X) - 2g_2(X) + g_1(X).$$

Then, with $\lambda = t$, this gives the density for `QuickMin`; and for $\lambda = 1 - u$, this gives the following density for `QuickMax`:

$$\Phi_Z(u,t) = \begin{cases} 2Z^2\ [F_1(Zt) - F_2(Zt)] & (\texttt{QuickMin}), \\ 2Z^2\ [F_1(Z(1-u)) - F_2(Z(1-u))] & (\texttt{QuickMax}). \end{cases}$$

For $\texttt{QuickVal}_\alpha$, we pose $x = u \wedge \alpha$, $y = t \vee \alpha$, and we obtain, in the case when $t \neq \alpha$ and $u \neq \alpha$,

$$\Phi_Z(u,t) = 2Z^2\left[2F_3(Z(y-x)) - 2F_2(Z(y-x)) + F_1(Z(y-x))\right]. \qquad (20)$$

Applying the previous proposition then ends the proof of Lemma. □

3.7 ***Exact mean number of comparisons in the Poisson model.***

In the model $(\mathcal{P}_Z, \mathcal{S})$, the density Φ_Z is a main tool for computing not only the mean number of key comparisons K_Z performed by the algorithm, but also the mean number of symbol comparisons S_Z.

With the definition of density $\Phi_Z(u,t)$ given in Definition 5, the mean number of key comparisons K_Z is obtained via the integral

$$K_Z = \mathcal{L}[\Phi_Z], \qquad (21)$$

where $\mathcal{L}$ is the linear functional defined for a function $\Phi : \mathcal{T} \to \mathbb{R}$ as

$$\mathcal{L}[\Phi] = \int_{\mathcal{T}} \Phi(u,t)\, du\, dt.$$

The definitions of the density $\Phi_Z(u,t)$ and the coincidence given in Definition 4 enable us to define *symbol-density* as the product $[\gamma(u,t)+1]\,\Phi_Z(u,t)$. This is the mean number of *symbol comparisons* (in the Poisson model $\mathcal{P}_Z$) between two words $M(u')$ and $M(t')$ for (u',t') close to (u,t). Then, the mean number of symbol comparisons S_Z is obtained via the formula

$$S_Z = \mathcal{J}[\Phi_Z] \quad \text{with} \quad \mathcal{J}[\Phi] = \int_{\mathcal{T}} [\gamma(u,t)+1]\,\Phi(u,t)\, du\, dt.$$

As we have already seen in Section 2.3 (also Eq. (8)), the functional $\mathcal{J}$ admits an alternative expression which involves the fundamental triangles $\mathcal{T}_w$,

$$\mathcal{J}[\Phi_Z] := \int_{\mathcal{T}} (\gamma(u,t)+1)\,\Phi_Z(u,t)\, du\, dt = \sum_{w \in \Sigma^\star} \int_{\mathcal{T}_w} \Phi_Z(u,t)\, du\, dt. \qquad (22)$$

This is a general phenomenon: formulas for the mean number of key comparisons or symbol comparisons are similar. When considering symbol comparisons, the functional $\mathcal{J}$ replaces the simple integral $\mathcal{L}$ used for key comparisons.

Coefficients of the mean number of symbol comparisons in the Poisson model. The expectation S_Z is then computed via its series expansion

$$S_Z = \sum_{k \geqslant 2} (-1)^k \frac{Z^k}{k!} \varphi(k).$$

Using the series expansion stated in (17), the expressions of S_Z as integrals given in (22), and the linearity of such integrals, the sequence $\varphi(k)$ is now defined, for any $k \geqslant 2$, in terms of integrals,

$$\varphi(k) := \int_{\mathcal{T}} (\gamma(u,t) + 1)\, \varphi(k,u,t)\, du\, dt = \sum_{w \in \Sigma^\star} \int_{\mathcal{T}_w} \varphi(k,u,t)\, du\, dt. \qquad (23)$$

This sequence $\varphi(k)$ depends both on the algorithm (via the sequence of functions $\varphi(k,u,t)$) and the source (via the fundamental triangles $\mathcal{T}_w$). In Lemma 3, for the `QuickVal` algorithm, we have only given the expression for $\varphi_\alpha(k,u,t)$ which holds almost everywhere on the triangle $\mathcal{T}$, except on the union of the two lines $u = \alpha$ and $t = \alpha$. In the integral, it is only this expression which intervenes.

Lemma 4 *For each algorithm* `QuickMin`*,* `QuickMax`*, and* `QuickVal`$_\alpha$*, the sequence $\varphi(k)$ satisfies, for $k \geqslant 2$,*

$$\varphi_-(k) = 2\mathcal{J}[t^{k-2}], \quad \varphi_+(k) = 2\mathcal{J}[(1-u)^{k-2}],$$

and

$$\varphi_\alpha(k) = \frac{4}{k+1} \mathcal{J}[(t \vee \alpha - u \wedge \alpha)^{k-2}].$$

3.8 *Exact expression of the mean number of symbol comparisons in the Bernoulli model.*

We now wish to return to the Bernoulli model $(\mathcal{B}_n, \mathcal{S})$, where the number of keys is fixed and equal to n. The mean number $S(n)$ of symbol comparisons used by the algorithm when it deals with n words independently drawn from the same source is related to S_Z and then to the sequence $\varphi(k)$ by the equation

$$S(n) = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \varphi(k)\,. \qquad (24)$$

Case of the first three algorithms. Lemma 4 provides in these three cases an exact formula for the mean number $S(n)$ of symbol comparisons:

Proposition 3 *The mean number of symbol comparisons, namely* $\mathtt{QM}_\pm(n)$ *for* `QuickMin` *and* `QuickMax`*, and* $\mathtt{QV}_\alpha(n)$ *for* $\mathtt{QuickVal}_\alpha(n)$*, admits an exact expression which involves the functional* $\mathcal{J}$ *of the source:*

$$\mathtt{QM}_-(n) = 2\sum_{k=2}^{n}(-1)^k\binom{n}{k}\mathcal{J}[t^{k-2}],$$

$$\mathtt{QM}_+(n) = 2\sum_{k=2}^{n}(-1)^k\binom{n}{k}\mathcal{J}[(1-u)^{k-2}],$$

$$\mathtt{QV}_\alpha(n) = 4\sum_{k=2}^{n}\frac{(-1)^k}{k+1}\binom{n}{k}\mathcal{J}[(t\vee\alpha - u\wedge\alpha)^{k-2}].$$

Note that our principle applies and, when we replace the functional $\mathcal{J}$ by the integral $\mathcal{L}$ on the triangle $\mathcal{T}$, we obtain a formula for the number of key comparisons for each algorithm.

Case of $\mathtt{QuickQuant}_\alpha$ *and* `QuickRand`*.* We now relate the mean number of symbol comparisons $\mathtt{QQ}_\alpha(n)$ of the algorithm $\mathtt{QuickQuant}_\alpha(n)$ to the mean number of symbol comparisons $\mathtt{QV}_\alpha(n)$ of the algorithm $\mathtt{QuickVal}_\alpha(n)$.

For a sequence $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{I}^n$, the algorithm $\mathtt{QuickQuant}(\alpha, n, \mathbf{x})$ outputs the word whose parameter is the real of rank $\lfloor 1 + \alpha(n-1)\rfloor = \lfloor \alpha n + 1 - \alpha\rfloor$ in the sequence $\mathbf{x}$, denoted by $\text{Rank}(\alpha, n, \mathbf{x})$. The two algorithms $\mathtt{QuickQuant}(\alpha, n, \mathbf{x})$ and $\mathtt{QuickVal}(\text{Rank}(\alpha, n, \mathbf{x}), n, \mathbf{x})$ behave exactly in the same way. This is why we have to study, for each α and n fixed, the distribution of the random variable $[0,1]^n \to [0,1]$, defined as $\mathbf{x} \mapsto \text{Rank}(\alpha, n, \mathbf{x})$. This variable is just the order statistic of rank $m = \lfloor \alpha n + 1 - \alpha\rfloor$, and its density $f_n(\alpha, v)$ is the Beta density with parameters $(m, n-m+1)$, namely

$$f_n(\alpha, v) = \frac{n!}{(m-1)!(n-m)!}v^{m-1}(1-v)^{n-m}, \quad \text{with} \quad m = \lfloor \alpha n + 1 - \alpha\rfloor. \tag{25}$$

This means that $f_n(\alpha, v)\,dv$ is the probability that $\text{Rank}(\alpha, n, \mathbf{x})$ belongs to the interval $[v, v+dv]$, and the following relation holds:

$$\mathtt{QQ}(\alpha, n, \mathbf{x}) = \int_{\mathcal{I}} \mathtt{QV}(v, n, \mathbf{x})\, f_n(\alpha, v)\, dv. \tag{26}$$

On the other hand, if we denote by $\mathtt{QQ}(\alpha, n, \mathbf{x})$ and $\mathtt{QV}(\alpha, n, \mathbf{x})$ the number of symbol comparisons of each algorithm when it is given the sequence $\mathbf{x}$, the two mean numbers $\mathtt{QQ}_\alpha(n)$, $\mathtt{QV}_\alpha(n)$ are given by

$$\mathtt{QQ}_\alpha(n) = \int_{\mathcal{I}^n} \mathtt{QQ}(\alpha, n, \mathbf{x})\, d\mathbf{x}, \qquad \mathtt{QV}_\alpha(n) = \int_{\mathcal{I}^n} \mathtt{QV}(\alpha, n, \mathbf{x})\, d\mathbf{x}.$$

Then, using (26), there is an exact integral formula for $\mathtt{QQ}_\alpha(n)$, namely,

$$\begin{aligned} \mathtt{QQ}_\alpha(n) &= \int_{\mathcal{I}^n} \mathtt{QQ}(\alpha, n, \mathbf{x})\, d\mathbf{x} \\ &= \int_{\mathcal{I}^n \times \mathcal{I}} \mathtt{QV}(v, n, \mathbf{x})\, f_n(\alpha, v)\, dv\, d\mathbf{x} \\ &= \int_{\mathcal{I}} \mathtt{QV}_v(n)\, f_n(\alpha, v)\, dv. \end{aligned} \tag{27}$$

In the same vein, the mean number $\mathtt{QS}(m, n)$ of symbol comparisons of $\mathtt{QuickSelect}(m, n)$ is then related to the Beta distribution $\beta_{m,n-m+1}$ via the equality

$$\begin{aligned} \mathtt{QS}(m, n) &= \int_{\mathcal{I}} \mathtt{QV}_v(n)\, \beta_{m,n-m+1}(v)\, dv \\ &= \frac{n!}{(m-1)!(n-m)!} \int_{\mathcal{I}} \mathtt{QV}_v(n) v^{m-1}(1-v)^{n-m}\, dv. \end{aligned}$$

The mean number of symbol comparisons $\mathtt{QR}(n)$ performed by the $\mathtt{QuickRand}$ algorithm on n words is equal to

$$\mathtt{QR}(n) = \frac{1}{n} \sum_{m=1}^{n} \mathtt{QS}(m, n),$$

and involves the mean of the densities $\beta_{m,n-m+1}$, which can be written as a binomial sum:

$$\frac{1}{n} \sum_{m=1}^{n} \beta_{m,n-m+1}(v) = \sum_{m=1}^{n} \frac{(n-1)!}{(m-1)!(n-m)!} v^{m-1}(1-v)^{n-m} = 1.$$

This proves the equality

$$\mathtt{QR}(n) = \int_{\mathcal{I}} \mathtt{QV}_v(n)\, dv.$$

We have thus obtained an exact formula for the mean number $S(n)$ for the other two variants $\mathtt{QuickQuant}_\alpha$ and $\mathtt{QuickRand}$.

Proposition 4 *The mean numbers of symbol comparisons* $\mathtt{QQ}_\alpha(n)$ *for the algorithm* $\mathtt{QuickQuant}_\alpha$ *and* $\mathtt{QR}(n)$ *for* $\mathtt{QuickRand}$ *each admit exact expression as a function of* $\mathtt{QV}_\alpha(n)$*:*

$$\mathtt{QQ}_\alpha(n) = \int_{\mathcal{I}} f_n(\alpha, v)\, \mathtt{QV}_v(n)\, dv, \qquad \mathtt{QR}(n) = \int_{\mathcal{I}} \mathtt{QV}_v(n)\, dv.$$

Towards the analytic part of the analysis. We have now obtained the exact expression of $S(n)$, and this concludes the algebraic part of the analysis. However, the formulae obtained in Propositions 3 and 4 do not give an easy or straightforward access to the asymptotic behaviour of $S(n)$ (when $n \to \infty$). In order to obtain asymptotic estimates, we have now to conduct the asymptotic part of the analysis, which deals with analytic tools. This will be accomplished in the following two sections, as there are indeed two different cases:

(a) For the algorithms `QuickVal` (and its variants), the Rice's method will be applied to expressions obtained in Proposition 3, and this is done in Section 4.

(b) For the $\texttt{QuickQuant}_\alpha(n)$ algorithm, we operate in an indirect way, and we deal with $\texttt{QQ}_\alpha(n)$ by *comparing* it to $\texttt{QV}_\alpha(n)$, with Proposition 4. This is done in Section 5, with a strong use of Laplace's method.

4 Asymptotic Analysis of `QuickMin`, `QuickMax`, `QuickVal`, and `QuickRand`.

In order to obtain asymptotic estimates of the mean $S(n)$, we now use analytic tools, namely the Rice's method, which transforms a binomial recurrence into an integral in the complex plane. We are led to deal with the mixed Dirichlet series $\varpi(s)$, which will be the analytic lifting of the sequence $\varphi(n)$. We then need tameness properties of the source, which entail tameness for the series $\varpi(s)$, so that the asymptotic behaviour can be estimated.

4.1 ***Rice's method.***

First step: An integral form. The first step of the Rice's method transforms a binomial recurrence of type (11) into an integral form (see [19,20]).

Proposition 5 *Let $S(n)$ be a numerical sequence which can be written as*

$$S(n) = \sum_{k=2}^{n} \binom{n}{k} (-1)^k \varphi(k).$$

Assume that the sequence $(\varphi(k))$ admits an analytic lifting $\varpi(s)$ in the half-plane $\Re s > 1$, which is there of polynomial growth with order at most r. Then the sequence $S(n)$ admits a Nörlund-Rice representation, for $n > r + 1$ and any d with $d \in]1, 2[$, namely,

$$S(n) = -\frac{1}{2i\pi} \int_{d-i\infty}^{d+i\infty} L_n(s) \cdot \varpi(s) \quad \textit{with } L_n(s) := \frac{(-1)^n \, n!}{s(s-1)\dots(s-n)}. \quad (28)$$

The function $L_n(s)$ is called the *Rice kernel*.

Second step: shifting to the left. Then, along general principles in analytic combinatorics as explained in [9,10], the integration line can be pushed to the left, as soon as $\varpi(s)$ has good analytic properties.

Proposition 6 *Assume that the lifting $\varpi(s)$ is meromorphic in a region $\mathcal{R}$ on the left of $\Re s = 1$ and of polynomial growth there (for $|\Im s| \to \infty$). Then*

$$S(n) = -\left[\sum_k \mathrm{Res}[L_n(s)\varpi(s); s = s_k] + \frac{1}{2i\pi}\int_{\mathcal{C}} L_n(s)\varpi(s)\, ds\right], \qquad (29)$$

where $\mathcal{C}$ is a curve (oriented from the south to the north) of class $\mathcal{C}^1$ included in $\mathcal{R}$ and the sum is extended to all poles s_k of $L_n(s)$ inside the domain $\mathcal{D}$ delimited by the vertical line $\Re s = d$ and the curve $\mathcal{C}$.

The dominant singularities of $L_n(s) \cdot \varpi(s)$ provide the asymptotic behaviour of $S(n)$, and the remainder integral is estimated using the polynomial growth of $L_n(s) \cdot \varpi(s)$ when $|\Im(s)| \to \infty$.

A particular case. We thus need a region $\mathcal{R}$ on the left of $\Re s = 1$, where $\varpi(s)$ is of polynomial growth (for $|\Im s| \to \infty$) and meromorphic. We need also a good knowledge of its poles. The functions $\varpi(s)$ involved in this paper fulfil nice properties, and the general framework for applying the Rice's method can be described as follows:

Proposition 7 *If the Dirichlet series $\varpi(s)$ has an abscissa of absolute convergence equal to δ with $\delta \in]0, 1[$, then, for any $\delta' \in]\delta, 1[$, the following asymptotics holds for the sequence $S(n)$:*

$$S(n) = \varpi(1) \cdot n + O(n^{\delta'}).$$

Here the constant of the O–term only depends on δ'.

4.2 ***Expression of the analytic lifting.***

We need first the analytic lifting $s \mapsto \varpi(s)$ of coefficients $\varphi(k)$ computed in Lemma 4 to exist in the half-plane $\Re s > 1$.

Lemma 5 *For each of the algorithms* `QuickMin`, `QuickMax`, *and* `QuickVal`$_\alpha$, *the analytic lifting $s \mapsto \varpi(s)$ of coefficients $\varphi(k)$ computed in Lemma 4 exists in the half-plane $\Re s > 1$, and satisfies*

$$\varpi_-(s) = 2\mathcal{J}[t^{s-2}], \qquad \varpi_+(s) = 2\mathcal{J}[(1-u)^{s-2}],$$

and

$$\varpi(\alpha, s) = \frac{4}{s+1}\mathcal{J}[(t \vee \alpha - u \wedge \alpha)^{s-2}].$$

As we wish to apply Rice's Method, we need stronger properties for the series $\varpi(s)$. We will prove in the following Proposition 8 that, under tameness properties of Dirichlet series of the source, then the Dirichlet series $\varpi(s)$ will satisfy hypotheses of Proposition 7.

4.3 ***Tameness of the mixed series.***

We will show the following result, which relates the tameness of the mixed series $\varpi(\alpha, s)$ to the tameness of the source.

Proposition 8 *Consider a source, its mixed series $\varpi_-(s)$, $\varpi_+(s)$ relative to* `QuickMin`, `QuickMax`, *and, for $\alpha \in [0,1]$, the series $\varpi(\alpha, s)$ relative to the* $\texttt{QuickVal}_\alpha$ *algorithm. Then the following holds:*

(*a*) *If the source is weakly Λ-tame with a width $\delta_\Lambda < 1/2$, let $\delta_0 := 2\delta_\Lambda$. Then the abscissa of absolute convergence of the mixed series $\varpi_-(s)$, $\varpi_+(s)$, $\varpi(0,s)$, $\varpi(1,s)$ is at most equal to δ_0.*

(*b*) *If the source is weakly Λ-tame with a width δ_Λ and Π-tame with a width δ_Π, let $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$. Then, for any $\alpha \in [0,1]$, the abscissa of absolute convergence of the mixed series $\varpi(\alpha, s)$ is at most equal to δ_0.*

We defer the proof of this proposition to Sections 4.5 through 4.8. We recall that the relation between the series Π and Λ entails the inequality $\delta_\Lambda \leqslant \delta_\Pi$.

4.4 ***Final asymptotic estimates.***

Now Rice's method, with Lemma 5 and Propositions 7 and 8, leads to the following result:

Theorem 3 *The following holds:*

(*a*) *Consider a source which is weakly Λ-tame with a width δ_Λ, and Π-tame with a width δ_Π. Assume that $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$ is strictly less than 1. Then the mean number of symbol comparisons for the* $\texttt{QuickVal}_\alpha(n)$ *algorithm satisfies*

$$\texttt{QV}_\alpha(n) = n\rho(\alpha) + O(n^\delta), \qquad \text{with } \delta \in]\delta_0, 1[.$$

The constant $\rho(\alpha)$ is the value of the mixed series $\varpi(\alpha, s)$ at $s = 1$, and the constant hidden in the O-term is uniform for $\alpha \in [0,1]$. The function $\alpha \mapsto \rho(\alpha)$ is called the asymptotic slope of the `QuickVal` *algorithm.*

(*b*) *Consider a source which is weakly Λ-tame with a width δ_Λ, and Π-tame with a width δ_Π. Assume that $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$ is strictly less than 1. Then, the mean number of symbol comparisons for the* $\texttt{QuickRand}(n)$ *algorithm satisfies*

$$\texttt{QR}(n) = n \left[\int_{\mathcal{I}} \rho(\alpha) d\alpha \right] + O(n^\delta), \qquad \delta \in]\delta_0, 1[.$$

(*c*) *Consider a weakly Λ-tame source with width $\delta_\Lambda < 1/2$. Then, the mean number of symbol comparisons for the* $\texttt{QuickMin}(n)$ *and* $\texttt{QuickMax}(n)$ *algorithm satisfy*

$$\texttt{QM}_\pm(n) = n\rho_\pm + O(n^\delta), \qquad \text{with } \delta \in]2\delta_\Lambda, 1[.$$

The constant $\rho_\pm$ is the value of the mixed series $\varpi_\pm(s)$ at $s = 1$. The equalities $\rho_- = \rho(0)$ and $\rho_+ = \rho(1)$ hold.

Proof This is a clear application of the Rice's method, as stated in Proposition 7, with together with Proposition 8. □

It then remains to study the tameness of $\varpi(s)$ and prove Proposition 8.

4.5 ***Principles of the proof of Proposition 8.***

We recall that the analysis of the `QuickMin` or `QuickMax` algorithms is based on the study of the mixed series

$$\varpi_\pm(s) = 2\mathcal{J}[\phi_\pm^{s-2}], \quad \text{with } \phi_-(u,t) = t \text{ and } \phi_+(u,t) = 1-u,$$

whereas the analysis of the algorithm $\texttt{QuickVal}_\alpha$ is based on the study of the mixed series

$$\varpi(\alpha, s) := \frac{4}{s+1}\underline{\varpi}(\alpha, s), \qquad \text{with} \quad \underline{\varpi}(\alpha, s) = \mathcal{J}[\phi_\alpha^{s-2}] \tag{30}$$

which involves the function

$$\phi_\alpha(u,t) := t \vee \alpha - u \wedge \alpha. \tag{31}$$

For $\alpha = 0$, the function $\phi_0(u,t)$ coincides with the function $\phi_-(u,t) = t$ of `QuickMin` on the whole triangle $\mathcal{T}$, whereas, for $\alpha = 1$, the function $\phi_1(u,t)$ coincides with the function $\phi_+(u,t) = 1-u$ of `QuickMax` on the triangle $\mathcal{T}$.

For $\alpha \in]0,1[$, there are three domains for the function ϕ_α, shown in Figure 8:

- The rectangle $\mathcal{R} = [0,\alpha] \times [\alpha, 0]$ where $\phi_\alpha(u,t)$ equals $t-u$.
- The lower triangle $\mathcal{T}_- = \mathcal{T}_{[0,\alpha]}$, where $\phi_\alpha(u,t)$ equals $\alpha - u$.
- The upper triangle $\mathcal{T}_+ = \mathcal{T}_{[\alpha,1]}$, where $\phi_\alpha(u,t)$ equals $t - \alpha$.

There is only one domain of interest (the whole triangle $\mathcal{T}$) when α equals 0 or 1, and

$$\varpi_-(s) = \frac{s+1}{2}\,\varpi(0,s), \qquad \varpi_+(s) = \frac{s+1}{2}\,\varpi(1,s).$$

The important behaviour of the mixed series $\varpi(s)$ is around $\Re s = 1$ and more precisely at $s = 1$, where the factor $(s+1)/2$ equals 1. This explains why the behaviours of `QuickMin` and $\texttt{QuickVal}_0$ on the one hand, and the behaviours of `QuickMax` and $\texttt{QuickVal}_1$ on the other hand, are asymptotically equivalent.

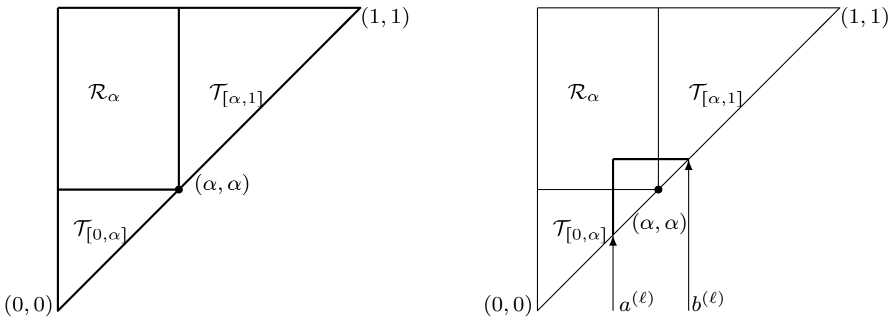


Fig. 8: The three domains useful for the study of $\varpi(\alpha, s)$ (left) and the straddling triangle (right).

Scheme of the proof of Proposition 8. For a domain $\mathcal{U} \subset \mathcal{T}$ and any integer $\ell \geqslant 2$, and any function $\phi \in L^1(\mathcal{T})$ we define

$$\mathcal{J}^{(\mathcal{U})}[\phi] := \int_{\mathcal{U}} (\gamma(u,t)+1)\phi(u,t)du\, dt, \qquad \mathcal{J}_\ell^{(\mathcal{U})}[\phi] := \int_{\mathcal{T}^{(\ell)}\cap\mathcal{U}} \phi(u,t)\, du\, dt,$$

where $\mathcal{T}^{(\ell)}$ is the union of all triangles $\mathcal{T}_w$ with $w \in \Sigma^\ell$. Then the following decomposition holds:

$$\mathcal{J}^{(\mathcal{U})}[\phi] = \sum_{\ell \geqslant 0} \mathcal{J}_\ell^{(\mathcal{U})}[\phi].$$

In the case of general α, the integral $\varpi(\alpha, s)$ is the sum of three terms. Each term is of the form

$$\varpi^{(\mathcal{U})}(\alpha, s) = \frac{4}{s+1}\underline{\varpi}^{(\mathcal{U})}(\alpha, s), \qquad \text{with} \quad \underline{\varpi}^{(\mathcal{U})}(\alpha, s) := \mathcal{J}^{(\mathcal{U})}[\phi_\alpha^{s-2}],$$

and $\mathcal{U}$ is any subset $\mathcal{R}_\alpha, \mathcal{T}_{[0,\alpha]}$ or $\mathcal{T}_{[\alpha,1]}$. Furthermore, the following decomposition holds:

$$\underline{\varpi}^{(\mathcal{U})}(\alpha, s) = \sum_{\ell \geqslant 0} \underline{\varpi}_\ell^{(\mathcal{U})}(\alpha, s), \qquad \text{with} \quad \underline{\varpi}_\ell^{(\mathcal{U})}(\alpha, s) := \mathcal{J}_\ell^{(\mathcal{U})}[\phi_\alpha^{s-2}]. \tag{32}$$

Relation between $\varpi(s)$ and Dirichlet series $\Lambda(s)$, $\Pi(s)$. This study aims at relating the series $\varpi(s)$ to (one of) the series $\Lambda(s)$ or $\Pi(s)$. This entails a relation between their abscissae of absolute convergence. We deal with the functions $\underline{\varpi}_\ell^{(\mathcal{U})}(s)$ and with the inequality

$$\left|\underline{\varpi}_\ell^{(\mathcal{U})}(s)\right| \leqslant \underline{\varpi}_\ell^{(\mathcal{U})}(\sigma) \qquad \text{for } \sigma = \Re s,$$

it is sufficient to deal with real values of parameters σ, and obtain inequalities of the kind

$$\underline{\varpi}_\ell^{(\mathcal{U})}(\sigma) \leqslant C_1\, \Lambda_\ell(\sigma_0) \qquad \text{or} \qquad \underline{\varpi}_\ell^{(\mathcal{U})}(\sigma) \leqslant C_2\, \Pi_\ell(\sigma_1) \tag{33}$$

for some constants C_1 and C_2 and some abscissae σ_0 or σ_1 related to σ. When now one of the inequalities $\sigma_0 > 1 + \delta_\Lambda$ or $\sigma_1 > \delta_\Pi$ holds, then weak Λ-tameness or Π–tameness entails that the abscissa of absolute convergence of the series $\underline{\varpi}^{(\mathcal{U})}(s)$ is at most σ.

We first consider the integral on the two triangles, then on the (possible) rectangle, and aim to obtain bounds of the kind (33). We will see that Lemmas 6 or 7 indeed provide such bounds.

4.6 ***Study of integrals over triangles.***

We first consider integrals over triangles $\mathcal{T}_{[0,\alpha]}, \mathcal{T}_{[\alpha,1]}$ and we prove the following estimates:

Lemma 6 *For any $\sigma_1 > 0$, there exists a constant C, such that, for any $\alpha \in [0,1]$, and for $\mathcal{U} \in \{\mathcal{T}_{[0,\alpha]}, \mathcal{T}_{[\alpha,1]}\}$, the function $\varpi^{(\mathcal{U})}(\alpha, s)$ satisfies the inequality, for $\sigma := \Re s \geqslant \sigma_1$,*

$$\left|\varpi^{(\mathcal{U})}(\alpha, s)\right| \leqslant C\Lambda\left(1 + \frac{\sigma}{2}\right).$$

If the source is weakly Λ-tame with width δ_Λ, then the abscissa of absolute convergence of the series $s \mapsto \varpi^{(\mathcal{U})}(\alpha, s)$ is at most $2\delta_\Lambda$.

Proof This proof is an easy extension of the proof which is done in [4] for the `QuickMin` algorithm. Using the reflected source $\check{\mathcal{S}}$ described in Section 2.6, it is sufficient to deal with one of the two cases for $\mathcal{U}$. Indeed, the mixed Dirichlet series $\varpi(\alpha, s)$ relative to a source $\mathcal{S}$ and the mixed Dirichlet series $\check{\varpi}(\alpha, s)$ relative to the reflected source $\check{\mathcal{S}}$ satisfy the identity

$$\check{\varpi}^{(\mathcal{T}_{[0,\alpha]})}(\alpha, s) = \varpi^{(\mathcal{T}_{[1-\alpha,1]})}(1-\alpha, s).$$

Since the two sources have the same Dirichlet series $\Pi(s) := \check{\Pi}(s)$, they share the same tameness properties.

We consider the function on the upper triangle, and we study the "underlined" functions, namely

$$\underline{\varpi}^{(\mathcal{T}_{[\alpha,1]})}(\alpha, s) = \sum_{\ell \geqslant 0} \underline{\varpi}_\ell^{(\mathcal{T}_{[\alpha,1]})}(\alpha, s)$$
$$\underline{\varpi}_\ell^{(\mathcal{T}_{[\alpha,1]})}(\alpha, s) = \sum_{w \in \Sigma^\ell} \int_{\mathcal{T}_w \cap \mathcal{T}_{[\alpha,1]}} (t-\alpha)^{s-2} du\, dt.$$

For each ℓ, we consider some real $A \in [\alpha, 1]$ (to be fixed later as a function of σ and ℓ) and split the sum into three sums, each of them relative to a subset of prefixes: the prefixes w for which the right end b_w belongs to $[\alpha, A[$, the prefixes w for which the left end a_w belongs to $[A, 1]$, and finally the unique prefix τ for which $A \in [a_\tau, b_\tau]$. We omit the reference to the real α, and the three sums are respectively denoted by $\underline{\varpi}_\ell^{(+)}(\sigma), \underline{\varpi}_\ell^{(-)}(\sigma), \underline{\varpi}_\ell^{(=)}(\sigma)$.

For $b_w \in [\alpha, A]$, we use the inequality

$$\int_{a_w}^{b_w} (t-a_w)(t-\alpha)^{\sigma-2} dt \leqslant \int_{a_w}^{b_w} (t-\alpha)^{\sigma-1} dt,$$

yielding

$$\underline{\varpi}_\ell^{(-)}(\sigma) \leqslant \int_\alpha^A (t-\alpha)^{\sigma-1} dt = \frac{1}{\sigma}(A-\alpha)^\sigma.$$

For $a_w \in [A, 1]$, we observe the inequality

$$\int_{a_w}^{b_w} (t-a_w)(t-\alpha)^{\sigma-2} dt \leqslant \frac{1}{2}(A-\alpha)^{\sigma-2} p_w^2,$$

which entails

$$\underline{\varpi}_\ell^{(+)}(\sigma) \leqslant \frac{1}{2}(A-\alpha)^{\sigma-2} \Lambda_\ell^{[\alpha,1]}(2).$$

We now choose A such that the two previous bounds are equal, namely,

$$A - \alpha = \left(\frac{\sigma}{2}\, \Lambda_\ell^{[\alpha,1]}(2)\right)^{1/2} \leqslant 1 - \alpha,$$

so that A belongs to the interval $[\alpha, 1]$. Then

$$\underline{\varpi}_\ell^{(+)}(\sigma) + \underline{\varpi}_\ell^{(-)}(\sigma) \leqslant C_1(\sigma) \left[\Lambda_\ell^{[\alpha,1]}(2)\right]^{\sigma/2},$$

where $C_1(\sigma)$ is bounded for $\sigma \geqslant \sigma_1$ (for any $\sigma_1 > 0$).
The middle part $\underline{\varpi}_\ell^{(=)}(\sigma)$ corresponds to the fundamental interval $[a_\tau, b_\tau]$ of length $p_\tau \leqslant \Lambda_\ell^{[\alpha,1]}(2)^{1/2}$, and

$$\underline{\varpi}_\ell^{(=)}(\sigma) \leqslant \int_{a_\tau}^{b_\tau} (t-\alpha)^{\sigma-1} dt \leqslant \frac{1}{\sigma}(A-\alpha+p_\tau)^\sigma \leqslant C_2(\sigma) \left[\Lambda_\ell^{[\alpha,1]}(2)\right]^{\sigma/2},$$

where $C_2(\sigma)$ is bounded for $\sigma \geqslant \sigma_1$ (for any $\sigma_1 > 0$).

Finally, we use the inequality $\Lambda_\ell^{[\alpha,1]}(2) \leqslant \Lambda_\ell(2)$, the log-convexity of $\Lambda_\ell(2)$, and the equality $\Lambda_\ell(1) = 1$ which together entail the inequality, for $\sigma \leqslant 2$,

$$\Lambda_\ell(2)^{\sigma/2} = \Lambda_\ell(2)^{\sigma/2} \Lambda_\ell(1)^{1-\sigma/2} \leqslant \Lambda_\ell\left(2\frac{\sigma}{2} + 1 - \frac{\sigma}{2}\right) = \Lambda_\ell\left(1 + \frac{\sigma}{2}\right),$$

and, taking the sum over integers ℓ, we have obtained the bound for the integral over the upper triangle,

$$\left|\varpi^{(\mathcal{T}_{[1,\alpha]})}(\alpha, s)\right| \leqslant C\, \Lambda\left(1 + \frac{\sigma}{2}\right).$$

This is the same bound for the integral on the lower triangle. □

4.7 ***Study of the integral over the rectangle.***

We prove the following:

Lemma 7 *For any $\sigma_1 > 0$, there exists a constant C for which the integral $\varpi^{(\mathcal{R})}(\alpha, s)$ satisfies*

$$\left|\varpi^{(\mathcal{R})}(\alpha, s)\right| \leqslant C\Pi(\sigma) \qquad \text{for any } \sigma \geqslant \sigma_1.$$

If the source is Π-tame with width δ_Π, then the abscissa of convergence of all the series $s \mapsto \varpi^{(\mathcal{R})}(\alpha, s)$ is at most δ_Π.

Proof The integral $\underline{\varpi}^{(\mathcal{R})}(\alpha, s)$ is the sum of the integrals $\underline{\varpi}_\ell^{(\mathcal{R})}(\alpha, s)$ defined as

$$\underline{\varpi}_\ell^{(\mathcal{R})}(\alpha, s) := \int_{\mathcal{T}^{(\ell)} \cap \mathcal{R}} (t-u)^{s-2}\, du\, dt.$$

For each depth ℓ, and for each α, there exists a unique word $w^{(\ell)}$ whose fundamental interval of the form $[a^{(\ell)}, b^{(\ell)}]$ contains α. Then, the intersection $\mathcal{T}^{(\ell)} \cap \mathcal{R}$ is the rectangle $[a^{(\ell)}, \alpha] \times [\alpha, b^{(\ell)}]$, and

$$\begin{aligned}\underline{\varpi}_\ell^{(\mathcal{R})}(\alpha, \sigma) &= \int_\alpha^{b^{(\ell)}} dt \int_{a^{(\ell)}}^{\alpha} (t-u)^{\sigma-2} du \\ &= \frac{1}{\sigma(\sigma-1)} \left[(b^{(\ell)} - a^{(\ell)})^\sigma - (b^{(\ell)} - \alpha)^\sigma - (\alpha - a^{(\ell)})^\sigma\right].\end{aligned}$$

If we let

$$c = \frac{b^{(\ell)} - \alpha}{b^{(\ell)} - a^{(\ell)}}, \qquad d = \frac{\alpha - a^{(\ell)}}{b^{(\ell)} - a^{(\ell)}}, \quad \text{with} \quad c + d = 1,$$

then the integral is written as

$$\underline{\varpi}_\ell^{(\mathcal{R})}(\alpha, \sigma) = \frac{1}{\sigma(\sigma-1)} (b^{(\ell)} - a^{(\ell)})^\sigma \left[c(1 - c^{\sigma-1}) + d(1 - d^{\sigma-1})\right].$$

There are two cases for $\sigma := \Re s$. Consider any $\rho_0 > 0$. The (easy) first case arises for $\sigma \geqslant 1 + \rho_0 > 1$, where

$$\underline{\varpi}_\ell^{\mathcal{R})}(\sigma) \leqslant \frac{2}{\rho_0(1+\rho_0)}\, (b^{(\ell)} - a^{(\ell)})^\sigma \leqslant \frac{2}{\rho_0(1+\rho_0)} \Pi_\ell(\sigma).$$

Consider now the case when $|\sigma - 1| \leqslant \rho_0$, and let $\rho := \sigma - 1$. The series expansion

$$x^\rho - 1 = \sum_{k \geqslant 1} \frac{\rho^k}{k!} \log^k x = \rho \log x + \rho^2 \left|\log x\right|^2 \sum_{k \geqslant 0} \frac{\rho^k}{(k+2)!} \log^k x$$

of $x \mapsto x^\rho$ provides the estimate

$$|x^\rho - 1 - \rho \log x| \leqslant \rho^2 \left|\log x\right|^2 e^{|\rho \log x|}.$$

For $x \in [0,1]$, one has $|\log x| = -\log x$ and $e^{|\rho||\log x|} = e^{-|\rho|\log x} = x^{-|\rho|}$, so that

$$\left|\frac{x(1-x^\rho)}{\rho} + x\log x\right| \leqslant \rho\, x \log^2 x\, e^{|\rho||\log x|} = \rho x^{1-|\rho|}\log^2 x.$$

If now $|\rho| \leqslant \rho_0$, with $\rho_0 < 1$, there exists C_1 (which depends on ρ_0) for which the inequality $x^{1-|\rho|}\log^2 x \leqslant C_1$ holds for any $x \in [0,1]$ and any ρ with $|\rho| \leqslant \rho_0$. Furthermore, the entropic sum $h(x) = -x\log x - (1-x)\log(1-x)$ belongs to $[0, \log 2]$, and finally

$$\underline{\varpi}_\ell^{(\mathcal{R})}(\alpha,\sigma) = \frac{2}{\sigma}(b^{(\ell)} - a^{(\ell)})^\sigma \left[h(c) + |\sigma - 1|\ \ O(1)\right].$$

Returning to $\varpi(\alpha, s)$ itself and taking the sum over all the integers ℓ gives the final bound

$$\varpi^{(\mathcal{R})}(\alpha,\sigma) \leqslant C(\sigma)\varPi(\sigma), \qquad \text{for any } \sigma \geqslant \sigma_0 > 0.$$

□

4.8 ***End of the proof.***

With the two Lemmas, and the formula

$$\varpi(\alpha,s) = \varpi^{(\mathcal{T}_-)}(\alpha,s) + \varpi^{(\mathcal{T}_+)}(\alpha,s) + \varpi^{(\mathcal{R})}(\alpha,s),$$

we obtain the following bound

$$|\varpi(\alpha,s)| \leqslant C\left[\varLambda\left(1+\frac{\sigma}{2}\right) + \varPi(\sigma)\right] \qquad \text{for some constant } C.$$

This entails that the series $\varpi(\alpha, s)$ is (absolutely) convergent as soon as $\sigma \geqslant \max(2\delta_\varLambda, \delta_\pi)$, and thus the result for $\texttt{QuickVal}_\alpha$. Observe that, in the cases $\alpha = 0$ or $\alpha = 1$, there is only one triangle (either $\mathcal{T}_-$ or $\mathcal{T}_+$) and no rectangle, and the weak $\varLambda$-tameness is then sufficient to conclude.

5 Asymptotic estimates for the $\texttt{QuickQuant}_\alpha$ Algorithm.

We return to the analysis of the `QuickQuant` algorithm with Proposition 4 and Theorem 3, and will prove the following result that exactly entails Assertion (a) of Theorem 1.

Theorem 4 *Consider a source which is weakly $\varLambda$-tame with a width $\delta_\varLambda$, and $\varPi$-tame with a width $\delta_\varPi$. Assume that $\delta_0 := \max(2\delta_\varLambda, \delta_\varPi)$ is strictly less than 1. Then, for any $\alpha \in [0,1]$, the mean number of symbols comparisons performed by the $\texttt{QuickQuant}_\alpha$ algorithm satisfies*

$$\texttt{QQ}_\alpha(n) = n\rho(\alpha) + O(n^\delta), \qquad \textit{with } \delta \in](1+\delta_0)/2, 1[.$$

Moreover, the constant hidden in the O-term is uniform on any interval of the form $[\alpha_1, 1-\alpha_1]$ with $\alpha_1 > 0$.

We first remark that the theorem is already proven for $\alpha = 0$ or $\alpha = 1$. Indeed, for $\alpha = 0$, the $\texttt{QuickQuant}_\alpha$ algorithm coincides with `QuickMin` and for $\alpha = 1$, with `QuickMax`. In these cases, this is just an application of Theorem 3, which provides an exponent $\delta \in]\delta_0, 1[$ (a priori smaller than the exponent of the present Theorem 4).

We now focus on the case when α belongs to the open interval $]0, 1[$. The relation between the costs $\texttt{QQ}_\alpha(n)$ and $\texttt{QV}_v(n)$ given in Proposition 4 together with the estimate for $\texttt{QV}_v(n)$ obtained in Theorem 3 leads us to the following estimate for $\texttt{QQ}_\alpha(n)$, with $\delta \in]\delta_0, 1[$:

$$\texttt{QQ}_\alpha(n) = \int_0^1 f_n(\alpha, v)\, \texttt{QV}_v(n)\, dv = \int_0^1 f_n(\alpha, v) \left[n\rho(v) + O(n^\delta)\right] dv.$$

Since the constant of the O-term is uniform with respect to v, we obtain

$$\texttt{QQ}_\alpha(n) = n \cdot I_n(\alpha) + O(n^\delta) \qquad \text{with} \quad I_n(\alpha) := \int_0^1 f_n(\alpha, v)\rho(v)\, dv. \tag{34}$$

It remains to study the integral $I_n(\alpha)$, and compare it to the value $\rho(\alpha)$. As the density $f_n(\alpha, v)$ can be compared to a large power of a function f_α, we are then led to apply Laplace's method (see for instance [10] or [5]), which first needs some regularity properties of the function ρ that we now establish.

5.1 *Regularity of the asymptotic slope of* `QuickVal`.

In order to compare the behaviors of the algorithms `QuickVal` and `QuickQuant`, and use Laplace's method, we need the asymptotic slope to be regular enough.

Proposition 9 *Consider a source which is weakly Λ-tame with a width δ_Λ, and Π-tame with a width δ_Π. Assume that $\delta_0 := \max(2\delta_\Lambda, \delta_\Pi)$ is strictly less than 1. Then, the asymptotic slope $\alpha \mapsto \rho(\alpha)$ of the* `QuickVal` *algorithm is Hölder with exponent η with $\eta < 1 - \delta_0$: for any η with $0 < \eta < 1 - \delta_0$, there exists a constant C_η for which*

$$|\rho(\beta) - \rho(\alpha)| \leqslant C_\eta\, |\beta - \alpha|^\eta \qquad \forall (\alpha, \beta) \in [0, 1]^2.$$

Proof We assume that the inequality $\alpha \leqslant \beta$ holds and begin with the inequality

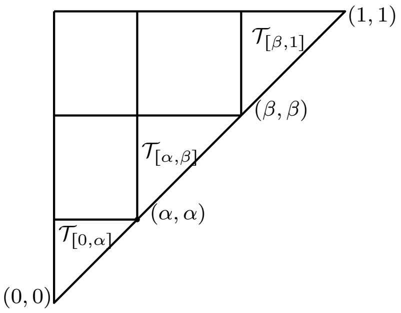
$$|\rho(\beta) - \rho(\alpha)| \leqslant 2\mathcal{J}\left[\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right|\right].$$

There are six domains, described in Figure 9. On each domain, we obtain an upper bound for the function $\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right|$, of the form

$$\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right| \leqslant |\beta - \alpha|^\eta \left(\phi_\alpha^{-1-\eta} + \phi_\beta^{-1-\eta}\right), \qquad \text{for any } \eta \in [0, 1].$$

We do not directly obtain such an upper bound: on each domain we get a "natural" upper bound of the form

$$\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right| \leqslant |\beta - \alpha|^\eta\, F_{\alpha,\beta}, \tag{35}$$


$(1,1)$
$\mathcal{T}_{[\beta,1]}$
(β,β)
$\mathcal{T}_{[\alpha,\beta]}$
(α,α)
$\mathcal{T}_{[0,\alpha]}$
$(0,0)$

Domain $\mathcal{U}$	$F_{\alpha,\beta}(u,t)$ for $(u,t) \in \mathcal{U}$
$\mathcal{T}_{[0,\alpha]}$	$(\alpha-u)^{-1-\eta} = \phi_\alpha^{-1-\eta}(u,t)$
$\mathcal{T}_{[\beta,1]}$	$(t-\beta)^{-1-\eta} = \phi_\beta^{-1-\eta}(u,t)$
$\mathcal{T}_{[\alpha,\beta]}$	$(t-\alpha)^{-1-\eta} + (\beta-u)^{-1-\eta}$ $= \phi_\alpha^{-1-\eta}(u,t) + \phi_\beta^{-1-\eta}(u,t)$
$[\alpha,\beta] \times [\beta,1]$	$(t-u)^{-1-\eta} = \phi_\beta^{-1-\eta}(u,t)$
$[0,\alpha] \times [\alpha,\beta]$	$(t-u)^{-1-\eta} = \phi_\alpha^{-1-\eta}(u,t)$
$[0,\alpha] \times [\beta,1]$	0

Fig. 9: The six domains for the study of $|\rho(\alpha) - \rho(\beta)|$ for any $\eta \in [0,1]$. The second column of the table on the right describes the "natural" intermediary function $F_{\alpha,\beta}$ mentioned in (35); it is expressed as a function of ϕ_α and ϕ_β.

and, on each domain, the function $F_{\alpha,\beta}$ is related to the functions ϕ_α or ϕ_β or both. This approach is summarised in Figure 9. The exponent η will be chosen later on.

We first explain how to obtain the results described in the table of Figure 9. Using the reflected source (described in Section 2.6), it is enough to study the function $|\phi_\alpha^{-1} - \phi_\beta^{-1}|$ on three domains: two triangles, $\mathcal{T}_{[0,\alpha]}$ and $\mathcal{T}_{[\alpha,\beta]}$, and one rectangle, $[\alpha,\beta] \times [\beta,1]$.

On the triangle $\mathcal{T}_{[0,\alpha]}$, with the inequalities $\beta - \alpha \leqslant \beta - u$ and $\alpha - u \leqslant \beta - u$ we have

$$\left|\phi_\beta^{-1}(u,t) - \phi_\alpha^{-1}(u,t)\right| = \frac{(\beta-\alpha)}{(\beta-u)(\alpha-u)} \leqslant \frac{(\beta-\alpha)^\eta}{(\beta-u)^\eta(\alpha-u)} \leqslant \frac{(\beta-\alpha)^\eta}{(\alpha-u)^{1+\eta}}.$$

On the rectangle $[\alpha,\beta] \times [\beta,1]$, using the inequalities $u - \alpha \leqslant \beta - \alpha \leqslant t - \alpha$ and $t - u \leqslant t - \alpha$ we have

$$\left|\phi_\beta^{-1}(u,t) - \phi_\alpha^{-1}(u,t)\right| = \frac{(u-\alpha)}{(t-\alpha)(t-u)} \leqslant (\beta-\alpha)^\eta \frac{(t-\alpha)^{1-\eta}}{(t-\alpha)(t-u)} \leqslant \frac{(\beta-\alpha)^\eta}{(t-u)^{1+\eta}}.$$

On the triangle $\mathcal{T}_{[\alpha,\beta]}$, with the inequalities $\beta - \alpha \geqslant t - \alpha$ and $\beta - \alpha \geqslant \beta - u$, we obtain

$$\left|\phi_\beta^{-1}(u,t) - \phi_\alpha^{-1}(u,t)\right| \leqslant \frac{1}{t-\alpha} + \frac{1}{\beta-u} \leqslant (\beta-\alpha)^\eta \left(\frac{1}{(t-\alpha)^{1+\eta}} + \frac{1}{(\beta-u)^{1+\eta}}\right).$$

Now, Figure 9 states six domains partitioning the unit triangle and gives in a table expressions for $F_{\alpha,\beta}(u,t)$ in each of these domains. For each domain $\mathcal{U}$ this yields

$$\mathcal{J}^{(\mathcal{U})}\left[\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right|\right] \leqslant |\beta-\alpha|^\eta \left(\mathcal{J}^{(\mathcal{U})}\left[|\phi_\alpha|^{-1-\eta}\right] + \mathcal{J}^{(\mathcal{U})}\left[|\phi_\beta|^{-1-\eta}\right]\right).$$

We take the sum over the six domains $\mathcal{U}$, and we obtain

$$\mathcal{J}\left[|\phi_\alpha|^{-1} - |\phi_\beta|^{-1}\right] \leqslant |\beta-\alpha|^\eta \left(\varpi(\alpha, 1-\eta) + \varpi(\beta, 1-\eta)\right).$$

We now use the results of Proposition 8 which proves that the functions $s \mapsto \varpi(\alpha, s)$ are uniformly bounded on the half-plane $\Re s \geqslant 1-\delta$, with $\delta < 1-2\delta_0$. Then, choosing any η with $0 < \eta < \delta$, we obtain, with

$$C_\eta := 2 \sup \{\varpi(\alpha, 1-\eta),\ \alpha \in [0,1]\},$$

the inequality

$$\mathcal{J}\left[\left|\phi_\alpha^{-1} - \phi_\beta^{-1}\right|\right] \leqslant |\beta-\alpha|^\eta \left(\varpi(\alpha, 1-\eta) + \varpi(\beta, 1-\eta)\right) \leqslant C_\eta\, |\beta-\alpha|^\eta\,.$$

□

5.2 *A tailored version of Laplace's method.*

Since we can only guarantee the function ρ to be Hölder, we need, and now state, a version of Laplace's method suited to this situation.

Theorem 5 (Laplace's method) *Let ρ and f be positive functions defined over the unit interval $\mathcal{I} = [0,1]$ of the real line, and consider the integrals*

$$J_n := \int_{\mathcal{I}} \rho(v)\, f(v)^n\, dv, \qquad K_n := \int_{\mathcal{I}} f(v)^n\, dv.$$

Assume the following:

(*i*) *ρ is Hölder with exponent η;*
(*ii*) *f is continuous on $[0,1]$ and infinitely differentiable on $]0,1[$;*
(*iii*) *f attains its maximum at a unique point $v_0 \in]0,1[$;*
(*iv*) *The three real numbers $\rho(v_0)$, $f(v_0)$, $f''(v_0)$ are non zero.*

Then, the integrals J_n and K_n satisfy

$$\frac{J_n}{K_n} = \rho(v_0)\big[1 + O(n^{-\eta'/2})\big], \qquad \textit{with} \quad \eta' < \eta.$$

Proof [Sketch of proof.] In Laplace's method, the contribution of a small interval around v_0 gives the main asymptotic term. More precisely, one considers the two integrals

$$J_n^{(1)} := \int_{A_n} \rho(v) f(v)^n\, dv, \quad K_n^{(1)} := \int_{A_n} f(v)^n\, dv,$$

with $A_n := [v_0 - \log n/\sqrt{n}, v_0 + \log n/\sqrt{n}]$.

First, the integral $K_n^{(1)}$ is proven to give the main term for K_n in Laplace's method, and, moreover, the Hölder condition for ρ is enough to prove that $J_n^{(1)}$ gives the main term for J_n in Laplace's method. More precisely, one has

$$\frac{J_n^{(1)}}{J_n} = 1 + O\left(\frac{1}{n}\right), \quad \frac{K_n^{(1)}}{K_n} = 1 + O\left(\frac{1}{n}\right).$$

Second, on the interval A_n, the following estimate holds for ρ:

$$\rho(v) = \rho(v_0) + O\left(\frac{\log n}{\sqrt{n}}\right)^{\eta} = \rho(v_0)\left[1 + O(n^{-\eta'/2})\right], \qquad \text{with} \quad \eta' < \eta.$$

Since $\eta' < 1$, it implies the estimates

$$J_n^{(1)} = \rho(v_0)\, K_n^{(1)}\, [1 + O(n^{-\eta'/2})], \quad \text{and thus} \quad J_n = \rho(v_0)\, K_n\, [1 + O(n^{-\eta'/2})].$$

□

We now study the integral $I_n(\alpha)$ defined in (34),

$$I_n(\alpha) = \int_{\mathcal{I}} \rho(v)\, f_n(\alpha, v)\, dv,$$

which involves the beta density $f_n(\alpha, v)$ with parameters $(\lfloor \alpha n + (1-\alpha)\rfloor, n - \lfloor \alpha n + (1-\alpha)\rfloor + 1)$. We follow the principles of Laplace's method, and we perform two steps.

5.3 ***Proof of Theorem 4. First step.***

In the first step, we "forget" the integer parts in the expression of $f_n(\alpha, v)$, and we deal with the Beta density $\widehat{f}_n(\alpha, v)$ of parameters $(\alpha n + 1 - \alpha, n - \alpha n + \alpha)$, and we wish to study the integral

$$\widehat{I}_n(\alpha) := \int_{\mathcal{I}} \rho(v) \widehat{f}_n(\alpha, v) dv.$$

The function $\widehat{f}_n(\alpha, v)$ is written as a product

$$\widehat{f}_n(\alpha, v) = \widehat{C}(\alpha, n)\, f_\alpha(v)^{n-1},$$

$$\text{with} \qquad \widehat{C}(\alpha, n) := \frac{\Gamma(n+1)}{\Gamma(\alpha n + 1 - \alpha)\Gamma(n - \alpha n + \alpha)}, \quad f_\alpha(v) := v^\alpha (1-v)^{(1-\alpha)}.$$

We first deal with the integrals *without* the Gamma-factor $\widehat{C}(\alpha, n)$, namely

$$J_n(\alpha) := \int_{\mathcal{I}} \rho(v) f_\alpha(v)^{n-1}\, dv, \quad K_n(\alpha) := \int_{\mathcal{I}} f_\alpha(v)^{n-1}\, dv,$$

and use Laplace's method. We first check the hypotheses for the function f_α. For $\alpha \in]0, 1[$, the function $f_\alpha(v)$ satisfies

$$f'_\alpha(v) = v^{\alpha-1}(1-v)^{-\alpha}\,(\alpha - v) \qquad \text{so that } f'_\alpha(\alpha) = 0,$$

$$f''_\alpha(\alpha) = -\frac{1}{\alpha(1-\alpha)} f_\alpha(\alpha) \qquad \text{so that } f''_\alpha(\alpha) \neq 0,$$

and $f_\alpha(v)$ attains its maximum at $v = \alpha$. Then, Laplace's method can be applied for any $\alpha \in]0,1[$, and provides the estimate

$$J_n(\alpha) = \rho(\alpha) \cdot K_n(\alpha) \cdot [1 + O(n^{-\eta'/2})].$$

Now, the initial integral of interest is

$$\widehat{I}_n(\alpha) := \int_{\mathcal{I}} \rho(v) \widehat{f}_n(\alpha, v)\, dv = \widehat{C}(\alpha, n) J_n(\alpha),$$

whereas the product $\widehat{C}(\alpha, n) K_n(\alpha)$ equals 1, since $\widehat{f}_n(\alpha, v)$ is a density. This leads to the estimate

$$\widehat{I}_n(\alpha) := \int_{\mathcal{I}} \rho(v) \widehat{f}_n(\alpha, v)\, dv = \rho(\alpha) \cdot [1 + O(n^{-\eta'/2})].$$

5.4 ***Proof of Theorem 4. Second step.***

We now take into account the "integer part", and we have to estimate the difference

$$\left| I_n(\alpha) - \widehat{I}_n(\alpha) \right| \leqslant \int_{\mathcal{I}} \rho(v) \left| f_n(\alpha, v) - \widehat{f}_n(\alpha, v) \right| \, dv \leqslant K R_n$$

with $R_n := \int_{\mathcal{I}} \left| f_n(\alpha, v) - \widehat{f}_n(\alpha, v) \right| \, dv$. We compare R_n to the integrals

$$K_n(\alpha) := \int_0^1 f_n(\alpha, v)\, dv, \qquad \widehat{K}_n(\alpha) := \int_0^1 \widehat{f}_n(\alpha, v)\, dv, \tag{36}$$

which are both equal to 1. We will show that R_n is "negligible" with respect to these integrals. Here, U_n is said to be negligible with respect to V_n if U_n/V_n is $O(1/n)$.

The two Beta densities $f_n(\alpha, v)$ and $\widehat{f}_n(\alpha, v)$ have the same shape: they are both unimodal, attain their maxima respectively at $v = \alpha$ and $v = \lfloor \alpha n - \alpha \rfloor / n$, and are equal to 0 at $v = 0$ and $v = 1$. We write the difference $|f_n(\alpha, v) - \widehat{f}_n(\alpha, v)|$ as

$$\left| f_n(\alpha, v) - \widehat{f}_n(\alpha, v) \right| = \left| \widehat{f}_n(\alpha, v) \right| \left| 1 - \frac{f_n(\alpha, v)}{\widehat{f}_n(\alpha, v)} \right|;$$

and the ratio $f_n(\alpha, v)/\widehat{f}_n(\alpha, v)$ is a product of two ratios, namely, the ratio between the Gamma factors and the ratio between the functions. First, the ratio between the Gamma factors, namely,

$$\frac{C(\alpha, n)}{\widehat{C}(\alpha, n)} = \frac{\Gamma(\alpha n + 1 - \alpha) \Gamma(n - \alpha n + \alpha)}{\Gamma(\lfloor \alpha n - \alpha \rfloor + 1) \Gamma(n - \lfloor \alpha n - \alpha \rfloor)},$$

is evaluated thanks to Stirling's formula for $x \to \infty$ and $a \in [0,1]$:

$$\frac{\Gamma(x+a)}{\Gamma(x)} = x^a \left[1 + O\left(\frac{1}{x}\right)\right] \quad \text{and thus} \quad \frac{C(\alpha,n)}{\widehat{C}(\alpha,n)} = \left(\frac{\alpha}{1-\alpha}\right)^{\{\alpha n - \alpha\}} + O\left(\frac{1}{n}\right).$$

On any interval $[\alpha_1, 1-\alpha_1]$, with $\alpha_1 > 0$, the ratio $\alpha/(1-\alpha)$ is bounded both from above and below, and the fractional part belongs to $[0,1]$ so that

$$\frac{C(\alpha,n)}{\widehat{C}(\alpha,n)} = \left(\frac{\alpha}{1-\alpha}\right)^{\{\alpha n - \alpha\}} \left[1 + O\left(\frac{1}{n}\right)\right]. \tag{37}$$

Second, it is possible to choose an interval A_n for which the following holds:

(a) The interval A_n contains the two points α and $\lfloor \alpha n - \alpha \rfloor / n = \alpha + O(1/n)$;

(b) The two integrals of $f_n(\alpha,v)$ and $\widehat{f}_n(\alpha,v)$ on the two intervals which form $I \backslash A_n$ are negligible with respect to the integral $K_n(\alpha)$ or $\widehat{K}_n(\alpha)$. It is then the same for the integral of the difference $|f_n(\alpha,v) - \widehat{f}_n(\alpha;v)|$;

(c) On the interval A_n, the ratio $(1-v)/v$ is uniformly bounded (with respect to n) both from below and above, and the ratio $f_n/\widehat{f}_n$ satisfies

$$\begin{aligned}\frac{f_n(\alpha,v)}{\widehat{f}_n(\alpha,v)} &= \left(\frac{1-v}{v}\right)^{\{\alpha n - \alpha\}} \cdot \frac{C(\alpha,n)}{\widehat{C}(\alpha,n)} \\ &= \left(\frac{1-v}{v}\right)^{\{\alpha n - \alpha\}} \left(\frac{\alpha}{1-\alpha}\right)^{\{\alpha n - \alpha\}} \left[1 + O\left(\frac{1}{n}\right)\right].\end{aligned}$$

These three properties for the interval A_n, entail the estimate, for $v \in A_n$,

$$\left|1 - \frac{f_n(\alpha,v)}{\widehat{f}_n(\alpha,v)}\right| = g_n(v) + O\left(\frac{1}{n}\right), \quad \text{with } g_n(v) = \left|1 - \left(\frac{1-v}{v}\frac{\alpha}{1-\alpha}\right)^{\{\alpha n - \alpha\}}\right|,$$

and thus the function $\left|f_n(\alpha,v) - \widehat{f}_n(\alpha,v)\right|$ satisfies

$$\left|f_n(\alpha,v) - \widehat{f}_n(\alpha,v)\right| = \widehat{f}_n(\alpha,v) \cdot \left|1 - \frac{f_n(\alpha,v)}{\widehat{f}_n(\alpha,v)}\right| = \widehat{f}_n(\alpha,v)\left(O\left(\frac{1}{n}\right) + g_n(v)\right).$$

We observe the upper bound

$$g_n(v) \leqslant g(v) \quad \text{with} \quad g(v) = \left|1 - \left(\frac{1-v}{v}\frac{\alpha}{1-\alpha}\right)\right| = \left|\frac{v-\alpha}{v(1-\alpha)}\right|.$$

As the last function is zero at $v = \alpha$, Laplace's method applies and shows that the integral

$$\int_{A_n} \widehat{f}_n(\alpha,v) g_n(v)\, dv$$

is negligible with respect to the integral $\widehat{K}_n(\alpha)$. And finally, the integral of $|f_n(\alpha, v) - \widehat{f}_n(\alpha, v)|$ on the interval A_n is negligible with respect to the integral $\widehat{K}_n(\alpha)$. We have proven in this section the estimate

$$I_n(\alpha) := \int_0^1 \rho(v) f_n(\alpha, v)\, dv = \widehat{I}_n(\alpha) + O(1/n),$$

and, together with the estimate of the previous section,

$$I_n(\alpha) = \rho(\alpha)[1 + O(n^{-\eta'/2})] + O(1/n) = \rho(\alpha) + O(n^{-\eta'/2}).$$

5.5 *Conclusion of the proof of Theorem 4.*

First we recall that, for $\alpha = 0$ the $\texttt{QuickQuant}_\alpha$ algorithm coincides with $\texttt{QuickMin}$, and for $\alpha = 1$ the $\texttt{QuickQuant}_\alpha$ algorithm coincides with $\texttt{QuickMax}$. Then, the theorem holds for $\alpha = 0$ and $\alpha = 1$. For $\alpha \in]0, 1[$, we begin with (34), and, with $\delta > \delta_0$ and $\delta_0 = \max(2\delta_\Lambda, \delta_\pi)$, we relate $\texttt{QQ}_\alpha(n)$ to the integral $I_n(\alpha)$:

$$\texttt{QQ}_\alpha(n) = n I_n(\alpha) + O(n^\delta).$$

Moreover, we have obtained an estimate for $I_n(\alpha)$ which involves an exponent $\eta' < \eta < 1 - \delta_0$:

$$I_n(\alpha) = \rho(\alpha) + O(n^{-\eta'/2}).$$

We finally obtain

$$\texttt{QQ}_\alpha(n) = n \cdot \rho(\alpha) + O(n^{1-(\eta'/2)}) + O(n^\delta).$$

As $1 - (\eta'/2) > 1 - (\eta/2) > 1 - (1 - \delta_0)/2 = (1 + \delta_0)/2 > \delta_0$, we obtain a remainder term of order $O(n^{\delta'})$ with $\delta' > (1 + \delta_0)/2$. This proves Assertion (a) of Theorem 1 for any $\alpha \in]0, 1[$. Note that on the interval $[\alpha_1, 1 - \alpha_1]$, with $\alpha_1 > 0$, the constant hidden in the O-term is uniform with respect to α.

Remark. It is probably possible to use the asymptotic normality of the α-quantile (as proven by Mosteller in [17]) for comparing more directly $\texttt{QV}_\alpha(n)$ and $\texttt{QQ}_\alpha(n)$. As we wish to obtain the precise remainder terms, we prefer to give the proof described here.

Now, we have proven all the assertions of Theorem 1. We now focus on the explicit expressions of the constants and prove Theorem 2.

6 Study of the asymptotic slope $\rho(\alpha)$.

We first provide an alternative expression for the asymptotic slope $\alpha \mapsto \rho(\alpha)$. Then we use it to study the maximality of the function $\alpha \mapsto \rho(\alpha)$ at $\alpha = 1/2$ when the source is unbiased memoryless, with an alphabet of odd size.

6.1 *An alternative expression for the asymptotic slope of* `QuickQuant`.

Let $\alpha \in [0,1]$. The asymptotic slope $\kappa(\alpha)$ for the mean number of key comparisons and the asymptotic slope $\rho(\alpha)$ for the mean number of symbol comparisons are given by

$$\begin{aligned}\kappa(\alpha) &= 2\int_{\mathcal{T}} \phi_\alpha(u,t)^{-1}\,du\,dt,\\ \rho(\alpha) &= 2\mathcal{J}[\phi_\alpha^{-1}] = \int_{\mathcal{T}} [\gamma(u,t)+1]\phi_\alpha(u,t)^{-1}\,du\,dt, \qquad (38)\end{aligned}$$

where ϕ_α is defined as $\phi_\alpha(u,t) := t \vee \alpha - u \wedge \alpha$ for a parameter $\alpha \in [0,1]$.

We are interested in giving a "short" expression for the asymptotic slope $\rho(\alpha)$. We first extend the definition of ϕ_α for any $\alpha \in \mathbb{R}$, with

$$\phi_\alpha(u,t) = t - \alpha \quad (\text{for } \alpha \leqslant 0), \quad \phi_\alpha(u,t) = \alpha - u \quad (\text{for } \alpha \geqslant 1).$$

This leads to an explicit expression for κ, now defined on the whole real line, which involves an extension H of the entropy function h, usually only defined on the interval $[0,1]$. More precisely, we write $\kappa(y) = 2[1+H(y)]$ where $H(y)$ is an extension of the entropic function h, defined as

$$H(y) := \begin{cases} y\log\left(1-\dfrac{1}{y}\right) & \text{if } y<0\\ h(y) = -y\log y-(1-y)\log(1-y), & \text{if } y\in[0,1]\\ (y-1)\log\left(1-\dfrac{1}{y}\right) & \text{if } y>1. \end{cases} \qquad (39)$$

With such an extension of the function κ, we obtain the following "short" expression for the asymptotic slope.

Proposition 10 *For a source $\mathcal{S}$ with fundamental intervals $[a_w, b_w]$, the asymptotic slope of the* `QuickVal` *algorithm satisfies, for any $\alpha \in [0,1]$,*

$$\rho(\alpha) = \sum_{w\in\Sigma^\star} (b_w - a_w)\,\kappa\left(\frac{\alpha - a_w}{b_w - a_w}\right).$$

Proof Starting with the expression of $\rho(\alpha) := \varpi(\alpha,1)$, namely,

$$\frac{1}{2}\,\rho(\alpha) = \sum_{w\in\Sigma^\star}\int_{\mathcal{T}_w} [(\alpha\vee t)-(\alpha\wedge u)]^{-1}\,du\,dt,$$

we use the simple change of variables which maps $\mathcal{T}_w$ onto $\mathcal{T}$, of the form $t = a_w + p_w t'$, $u = a_w + p_w u'$, and the relations

$$\begin{aligned}(\alpha\vee t) &= a_w + [(\alpha - a_w)\vee(p_w t')] = a_w p_w + p_w\left[\left(\frac{\alpha - a_w}{p_w}\right)\vee t'\right],\\ (\alpha\wedge u) &= a_w + [(\alpha - a_w)\wedge(p_w u')] = a_w p_w + p_w\left[\left(\frac{\alpha - a_w}{p_w}\right)\wedge u'\right].\end{aligned}$$

These entail the equality

$$\frac{1}{p_w}(\alpha \vee t) - (\alpha \wedge u) = \left[\left(\frac{\alpha - a_w}{p_w}\right) \vee t'\right] - \left[\left(\frac{\alpha - a_w}{p_w}) \wedge u'\right)\right],$$

yielding the final result. □

6.2 ***Maximality of the asymptotic slope of the*** `QuickVal` ***algorithm at*** $\alpha = 1/2$.

It is not always true that the asymptotic slope $\rho(\alpha)$ is maximal at $\rho = 1/2$, even for a symmetric source. A clear counterexample is shown for the simplest source, the binary source, in Figure 2. The following result shows that this is true for any unbiased memoryless source over an alphabet of *odd* size.

Proposition 11 *For an unbiased memoryless source over an alphabet of odd size, the constant $\rho(\alpha)$ is maximised by $\alpha = 1/2$.*

Proof For an unbiased memoryless source over an alphabet of finite size r, we can express $\rho(\alpha)$ in terms of κ, for any $\alpha \in \mathbb{R}$, as

$$\rho(\alpha) = \sum_{\ell=0}^{\infty} r^{-\ell}\rho_\ell(\alpha) \qquad \text{with} \quad \rho_\ell(\alpha) := \sum_{k=0}^{r^\ell-1} \kappa(r^\ell\alpha - k).$$

We will prove that each $\rho_\ell(\alpha)$ is maximum at $\alpha = 1/2$. When r is odd, we let $r^\ell - 1 = 2J$, and we write the index $k \in [0\,..\,2J]$ as $k = -j + J$ with $j \in [-J, J]$, together with $r^\ell\alpha - k = (2J+1)\alpha - J + j$. Now, for integer $J \geqslant 0$ we define

$$f_J(x) := \sum_{j=-J}^{J} \kappa(j + x), \quad x \in \mathbb{R}, \tag{40}$$

so that

$$\rho_\ell(\alpha) = f_J(x) \quad \text{with} \quad x := (2J+1)\alpha - J.$$

It is then sufficient to study f_J on the interval $[-J, J+1]$ and prove that f_J is maximal at $x = 1/2$. This will prove that ρ_ℓ is maximal at $\alpha = 1/2$.

Fix $J \geqslant 1$ and write f as shorthand for f_J. First observe that

$$\sup_{x\in\mathbb{R}} f(x) = \sup_{x\in[0,1]} \sup_{k\in\mathbb{Z}} f(k + x). \tag{41}$$

We will prove the following two facts:

(i) For any $x \in [0,1]$ the expression $f(j + x)$ is non-increasing in integer $j \geqslant 0$.
(ii) The expression $f(x)$ is maximised over $x \in [0,1]$ by $x = 1/2$.

With (i), and the symmetry of $f(x)$ with respect to $x = 1/2$, it follows, for any $x \in [0,1]$, that $f(j+x)$ is unimodal in $j \in \mathbb{Z}$ with maximum at $j = 0$. The lemma then follows from (ii) and (41).

To establish (i), let $j \geqslant 0$ and observe

$$f(j+1+x) - f(j+x) = \kappa(j+J+1+x) - \kappa(j-J+x). \tag{42}$$

If $j-J+x \geqslant 1/2$, then the difference (42) is $\leqslant 0$ because $\kappa(x)$ is non-increasing for $x \geqslant 1/2$. On the other hand, if $j-J+x < 1/2$, then the difference (42) is upper-bounded by

$$\kappa(j+J+1) - \kappa(j-J) = \kappa(-j-J) - \kappa(j-J) \leqslant 0.$$

To establish (ii), we begin by calculating $f(x)$ explicitly when $x \in [0,1]$. First, using the first line of (39), we find

$$\frac{1}{2}\sum_{j=1}^{J} \kappa(j+x) = J + \sum_{j=1}^{J} (1-j-x)[\log(j+x) - \log(j-1+x)].$$

Using Abel's transform, and introducing the function $g(y) := (1-y)\log y$, we obtain

$$\frac{1}{2}\sum_{j=1}^{J} \kappa(j+x) = J + g(J+x) + x\log x + \sum_{j=1}^{J-1} \log(j+x).$$

Next, using the symmetry of $\kappa(x)$ about $x = 1/2$, we observe

$$\begin{aligned}\frac{1}{2}\sum_{k=-J}^{-1} \kappa(j+x) &= \frac{1}{2}\sum_{j=1}^{J} \kappa(j+1-x) \\ &= J + g(J+1-x) + (1-x)\log(1-x) + \sum_{j=1}^{J-1} \log(j+1-x).\end{aligned}$$

Summing these last two expressions together with the expression for $\kappa(x)$ in the second line of (39) we arrive at

$$\frac{1}{2}f(x) = (2J+1) + [g(J+x) + g(J+1-x)] + \sum_{j=1}^{J-1} \log[(j+x)(j+1-x)]. \tag{43}$$

Of the three terms here, the first is constant, the third is unimodal with maximum at $x = 1/2$ (since each term in the sum has this property), and the second is maximised at $x = 1/2$ because $g(y)$ is concave for $y \geqslant 1$. (In fact, it is concave for all $y > 0$.) Thus (ii) is established, as is the proposition. □

Some experiments performed by Philippe Flajolet showed that this property does not hold for every memoryless source, even if it is symmetric and built on an alphabet of odd size. However, we conjecture the following:

Conjecture. Consider a symmetric memoryless source over an ordered alphabet Σ of odd size denoted by $\{0, 1, \ldots, r-1\}$ and assume that the middle symbol $(r-1)/2$ is the most probable, namely $p_{(r-1)/2} \geqslant p_i$ for all $i \in \Sigma$. Then the constant $\rho(\alpha)$ is maximised by $\alpha = 1/2$.

Conclusion

Our study shows that the QuickSelect algorithm uses a linear number of comparisons (on average) for a wide variety of models, considering either key or symbol comparisons, or even changing the type of sources for symbol comparisons. We exhibit a striking fact: the constants change but the complexity remains linear. The situation is completely different for sorting algorithms: in [4], we study for instance the QuickSort algorithm, and we show that the mean number of comparisons goes from $\Theta(n \log n)$ to $\Theta(n \log^2 n)$ depending if we consider key or symbol comparisons on a "nice" source. Moreover, there are intermittent sources (of type described in Section 2.4) on which the complexity of QuickSort becomes $\Theta(n^\beta)$ for some $\beta > 1$.

Then, there is a strong contrast between the behaviour of QuickSelect and QuickSort, and the QuickSelect algorithm seems to be more "robust" to the change of model, and notably the quality of the source. Our analysis exhibits such a phenomenon, via the tameness properties of the source which are needed for the mixed series to be itself tame. The two kinds of tameness are directed related for sorting algorithms, whereas this relation is less tight for selection algorithms and gives more latitude to the source. It would be thus interesting to find (or design) sources which could change the symbol complexity of QuickSelect; such a candidate could be the intermittent sources of Section 2.4 with exponent $a = 1$, or their extension presented in [1]. This will help to better understand the extent of our framework.

It would be also interesting to investigate the lower bound of the selection problem in our symbol comparison model. We have conducted such a study in [4] for the mean number of comparisons performed by sorting algorithms: we start with the decision tree used in the classical key comparison model, and we "mix" it with a trie, along the approach described by Seidel [24]. However, for the selection problem, the lower bound is already more intricate in the classical key comparison model, at least for a general rank m, even though there are some results given in [16], for instance. It would be natural to begin with the minimum problem, where the lower bound is n in the usual key comparison model. In [18], we have analysed, in our symbol comparison model, another algorithm for the selection of the minimum (the first step of the sorting Selection algorithm), and proved its complexity to be $\Theta(n)$, with another constant. It would be interesting to exhibit the constant for the lower bound of the minimum selection problem in our symbol comparison model.

Acknowledgements. The authors wish to thank an anonymous reviewer for constructive comments, and also Jérémie Lumbroso and Bruno Salvy for their wise advice concerning the last part of the abstract.

References

1. P. Cénac, B. Chauvin, F. Paccaut, and N. Pouyanne. Uncommon suffix tries. *Random Structures & Algorithms*, 2013.

2. J. Clément, P. Flajolet, and B. Vallée. Dynamical sources in information theory: A general analysis of trie structures. *Algorithmica*, 29(1):307–369, 2001.
3. J. Clément, T. H. Nguyen Thi, and B. Vallée. A general framework for the realistic analysis of sorting and searching algorithms. Application to some popular algorithms. In N. Portier and T. Wilke, editors, *30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013)*, volume 20 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 598–609, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
4. J. Clément, T. H. Nguyen Thi, and B. Vallée. Towards a realistic analysis of some popular sorting algorithms. *Special issue of CPC (Combinatorics, Probability and Computing) dedicated to Philippe Flajolet*, 24(01):104–144, 2014.
5. N. De Bruijn. *Asymptotic methods in Analysis.* Dover, 1981.
6. J. Fill and J. Matterer. QuickSelect tree process convergence, with an application to distributional convergence for the number of symbol comparisons used by worst-case Find. *To appear in the special issue of CPC dedicated to Philippe Flajolet*, 2014.
7. J. A. Fill and T. Nakama. Analysis of the expected number of bit comparisons required by QuickSelect. *Algorithmica*, 58(3):730–769, 2010.
8. J. A. Fill and T. Nakama. Distributional convergence for the number of symbol comparisons used by QuickSelect. *Advances in Applied Probability*, 45:425–450, 2013.
9. P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theor. Comput. Sci.*, 144(1&2):101–124, 1995.
10. P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.
11. P. J. Grabner and H. Prodinger. On a constant arising in the analysis of bit comparisons in Quickselect. *Quaestiones Mathematicae*, 31(4):303–306, 2008.
12. C. A. R. Hoare. Algorithm 63: partition. *Commun. ACM*, 4(7):321, 1961.
13. C. A. R. Hoare. Algorithm 64: Quicksort. *Commun. ACM*, 4(7):321, 1961.
14. C. A. R. Hoare. Quicksort. *Comput. J.*, 5(1):10–15, 1962.
15. D. E. Knuth. Mathematical Analysis of Algorithms. In *IFIP Congress (1)*, pages 19–27, 1971.
16. D. E. Knuth. *The Art of Computer Programming, Volume 3: (2Nd Ed.) Sorting and Searching.* Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.
17. F. Mosteller. On some useful inefficient statistics. *Ann. Math. Statist.*, 17(4):377–408, 1946.
18. T. H. Nguyen Thi. *Towards a realistic analysis of sorting and searching algorithms.* PhD thesis, Université de Caen Basse Normandie, Dec 2014.
19. N. E. Nörlund. Leçons sur les équations linéaires aux différences finies. In *Collection de monographies sur la théorie des fonctions.* Gauthier-Villars, Paris, 1929.
20. N. E. Nörlund. *Vorlesungen über Differenzenrechnung.* Chelsea Publishing Company, New York, 1954.
21. R. Sedgewick. *Quicksort.* Outstanding Dissertations in the Computer Sciences. Garland Publishing, New York, 1975.
22. R. Sedgewick. *Quicksort.* Garland Pub. Co., New York,, 1980.
23. R. Sedgewick and P. Flajolet. *An introduction to the analysis of algorithms.* Addison-Wesley-Longman, 1996.
24. R. Seidel. Data-specific analysis of string sorting. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 1278–1286, 2010.
25. B. Vallée. Dynamical sources in information theory: Fundamental intervals and word prefixes. *Algorithmica*, 29(1/2):262–306, 2001.
26. B. Vallée, J. Clément, J. A. Fill, and P. Flajolet. The number of symbol comparisons in QuickSort and QuickSelect. In S. Albers *et al.*, editor, *Proceedings of ICALP 2009, Part I*, volume 5555 of *Lecture Notes in Computer Science*, pages 750–763. Springer-Verlag, 2009.
27. M. H. van Emden. Increasing the efficiency of quicksort. *Commun. ACM*, 13(9):563–567, 1970.