# Habilitation universitaire

20 janvier 2010

# Table des matières


**1 A perturbation expansion method** **7**
- 1.0.1 Introduction . . . 7
- 1.0.2 Perturbation Expansion : Outlines of the Method . . . 8
- 1.0.3 Some definitions . . . 9
- 1.0.4 Main results . . . 10
- 1.0.5 A worked-out example . . . 11

**2 Some limit results on random trees** **13**
- 2.1 Introduction . . . 13
- 2.2 Distances in Trees . . . 13
  - 2.2.1 Introduction . . . 13
  - 2.2.2 Digital trees . . . 13
  - 2.2.3 Notation and methodology . . . 15
  - 2.2.4 Distances in DST . . . 15
  - 2.2.5 Tries . . . 16
- 2.3 Binary tree . . . 17
  - 2.3.1 Introduction . . . 17
  - 2.3.2 Weighted path to the minimal label . . . 18
  - 2.3.3 Weighted path to the maximal label . . . 19
  - 2.3.4 Some general questions . . . 21

**3 Some Limit Theorems** **23**
- 3.1 Asymptotic behavior for sums of some weakly dependant random variables 23
  - 3.1.1 introduction . . . 23
  - 3.1.2 Main results . . . 24
  - 3.1.3 Discussions . . . 25
- 3.2 Polya urns . . . 26
  - 3.2.1 Introduction . . . 26
  - 3.2.2 Embedding method . . . 27
  - 3.2.3 Asymptotic composition of the discrete urn . . . 28
  - 3.2.4 A project work . . . 30
- 3.3 Random fragmentation . . . 30
  - 3.3.1 Introduction . . . 30

3.3.2 Homogeneous random fragmentation process . . . . . 31
3.3.3 Exponential fragmentation probability . . . . . 32
3.3.4 A project work . . . . . 36
3.4 A pattern matching problem . . . . . 36
3.4.1 Introduction . . . . . 36
3.4.2 Some examples . . . . . 37
3.4.3 Generating function of the generalized Moran model . . . . . 37
3.4.4 Some future works . . . . . 39

# Report on Research Activities

Our research work is concerned with various questions related to probability theory and its applications to adaptive control algorithms and random trees. Namely, we have studied the following :

- stochastic tracking algorithms,
- a version of the Central Limit Theorem via weak dependence,
- random trees,
- pòlya urns,
- fragmentation processes,
- Moran model.

In our Phd thesis, we proved an explicit formula for the asymptotic error of certain classes of stochastic tracking algorithms. We have developed a method of analysis of the performance called "perturbation expansion". More precisely, given a recursive equation in the form

$$\eta_{n+1} = (I - \mu F_n)\eta_n + \xi_{n+1},$$

where

$$\begin{aligned} \xi_n : &\quad \text{is a perturbation,} \\ \mu : &\quad \text{the adaptation step-size,} \\ F_n : &\quad \text{a given square random matrix.} \end{aligned}$$

The "perturbation expansion" method is based on replacing the random matrix $(I - \mu F_n)$ by an appropriate deterministic matrix to obtain an explicit development of the moments of $\eta_n$ in the form of a finite sequence of size $r$ ($r$ is a given integer which depends on processes $F_n$, $\mu$).

In collaboration with Louhichi Sana ( [5]), we have considered a triangular array of row-wise stationary, centered and square integrable real valued random variables $(X_{i,n})_{0 \leq i \leq n}$, satisfying the following condition of weak dependence

$$\lim_{N \to N_0} \lim \sup_{n \to +\infty} n \sum_{r=N}^{n} cov(X_{0,n}, X_{r,n}) = 0;$$

where $N_0$ is the smallest integer $N$ such that $\lim_{n\to+\infty} n\sum_{r=N}^{n} cov(X_{0,n}, X_{r,n}) = 0$. We have constructed a sequence $(\tilde{X}_{i,n})_{0\leq i\leq n}$ of independent random variables such that $S_n := \sum_{i=1}^{n} X_{i,n}$ and $\tilde{S}_n := \sum_{i=1}^{n} \tilde{X}_{i,n}$ have, asymptotically, the same distribution.

In collaboration with Hosam Mahmoud and Nabil Lasmar ([1, 2, 4]), we have studied distribution of inter-node distances in digital trees. We have also studied the extremal weighted path lengths in random binary search trees constructed from a random permutation of 1, 2, $\cdots n$.

We have also been interested in some "Polya urns" ([3]) (A model which consists of drawing one ball, then adding $n$ black balls and $m$ red balls, where the values of $n$ and $m$ depend on the color of the drawn ball ). We studied the asymptotic composition, almost surely and in distribution, for a family of random distributions of $n$ and $m$.

Fragmentation process represents another focus of our research. We have been interested in the following problem. Starting with an object of size $x$ large enough and let $(U, 1-U)$ be a random vector, where $U$ is a uniform distributed random variable on $[0, 1]$. At time $n = 0$,with probability $1 - e^{-x}$, we break $x$ into two pieces with sizes $Ux$ and $(1-U)x$ and with probability $e^{-x}$ we let it indefinitely stable. At every time $n$ every piece with size $l$ , with probability $1 - e^{-l}$, is broken into two pieces (independently from the others) using a new (independent) copies of the random vector $(U, 1-U)$, and with the complimentary probability is indefinitely stable. Using a technique introduced by Svante Janson and Ralph Neininger [52], we have proved a distribution limit theorem for the random variable $N(x)$ which represents the number of stable pieces at the end of the process.

Together with Cyril Banderier we have started studying the height of a random walk which at each step either advances by a step or get back to the origin. We also considered the generalization of this model to dimension $d$. This is constructed from Moran Model which is related to population biology and failure theory. Our approach is based on generating functions and Complex Analysis.

# Chapitre 1

# A perturbation expansion method

### 1.0.1 Introduction

An important issue in system identification, signal processing, automatic control is that of tracking the parameter variations in a linear regression model

$$y_t = \phi_t^T \theta_t + v_t \ ; \ t \geq 0 \tag{1.1}$$

where $\{y_t\}_{t\geq 0}$ and $\{v_t\}_{t\geq 0}$ are respectively the scalar observation and noise, $\{\phi_t\}_{t\geq 0}$ and $\{\theta_t\}_{t\geq 0}$ are the $d$-dimensional stochastic regressor and the unknown time-varying parameter. This model encompasses many different applications, including channel equalization, time delay estimation and echo cancellation [87]. In the sequel it is assumed that the parameter variation obeys

$$\theta_{t+1} = \theta_t + w_{t+1} \tag{1.2}$$

where $w_{t+1}$ is referred to the lag-noise. To track the variations of the parameter, it is customary to use a recursive algorithm for updating an estimate $\hat{\theta}_t$ of the parameter (see [31, 87] and the references therein). Most of theses algorithms can be put in the form

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \mu L_t (y_t - \phi_t^T \hat{\theta}_t). \tag{1.3}$$

where $\mu$ is referred to as the adaptation step-size and $L_t$ is a random vector, which can be chosen in a number of different ways. There is a vast literature on the analysis of algorithms of type (1.3). In most contributions, the main goal is to obtain bounds n the tracking errors. Results in that directions have been obtained in [31, 43, 42]. In this paper a different approach is pursued. Our goal is to obtain explicit expression and not only bounds for the tracking error. To that purpose, we will use a technique, referred to as "perturbation expansion", consisting in getting approximations (1.3) by nested processes, with much simpler structure than the original error process. This particular decomposition enables the computation of explicit expressions for the moments and other related quantities.

### 1.0.2 Perturbation Expansion : Outlines of the Method

From (1.1) and (1.3), we can write

$$\tilde{\theta}_{t+1} = (I - \mu L_t \phi_t^T)\tilde{\theta}_t + \mu L_t v_t - w_{t+1}, \tag{1.4}$$

where $\tilde{\theta}_t = \hat{\theta}_t - \theta_t$ is the weight-error vector. Since this equation is linear, $\tilde{\theta}_{t+1}$ can be decomposed as

$$\begin{aligned}
\tilde{\theta}_t &= \tilde{\theta}_t^u + \mu\tilde{\theta}_t^v + \tilde{\theta}_t^w, \\
\tilde{\theta}_{t+1}^u &= (I - \mu L_t \phi_t^T)\tilde{\theta}_t^u,\ \tilde{\theta}_0^u = \tilde{\theta}_0 = -\theta_0, && (1.5)\\
\tilde{\theta}_{t+1}^v &= (I - \mu L_t \phi_t^T)\tilde{\theta}_t^v + L_t v_t,\ \tilde{\theta}_0^v = 0, && (1.6)\\
\tilde{\theta}_{t+1}^w &= (I - \mu L_t \phi_t^T)\tilde{\theta}_t^w - w_{t+1},\ \tilde{\theta}_0^w = 0. && (1.7)
\end{aligned}$$

$\{\tilde{\theta}_t^u\}$ is a transient term, reflecting the way the successive estimates of the regression coefficients forget the initial conditions. $\{\tilde{\theta}_t^v\}$ accounts for the errors introduced by the measurement noise, $\{v_t\}$ ; similarly, $\{\tilde{\theta}_{t+1}^w\}$ accounts for the errors associated lag-noise $\{w_t\}$. According to these definitions, $\tilde{\theta}_t^v$ and $\tilde{\theta}^w$ obey an inhomogeneous stochastic recurrence equation

$$\begin{aligned}
\delta_{t+1} &= (I - \mu F_t)\delta_t + \xi_t,\ \delta_0 = 0 && (1.8)\\
&= \sum_{s=0}^{t} \Phi(t,s)\xi_s && (1.9)
\end{aligned}$$

where $\{F_t\}_{t\geq 0}$ is a matrix-valued random process, $\{\xi_t\}_{t\geq 0}$ is a $(d\times 1)$ vector-valued random process, and $\Phi(t,s)$ is defined as

$$\begin{cases}
(I - \mu F_t)(I - \mu F_{t-1})\cdots(I - \mu F_{s+1}), & \text{if } t > s \\
I, & \text{if } t = s \\
0, & \text{otherwise.}
\end{cases}$$

Here, the dependance of $\delta_t$ upon the step-size $\mu$ is implicit. Eqs (1.6) and (1.7) may be rewritten as (1.8) with $F_t = L_t\phi_t^T$ and, recursively,

$$\xi_t = L_t v_t \text{ measurement noise, } \xi_t = -w_{t+1} \text{ lag noise.} \tag{1.10}$$

Applied to the recurrence equation (1.8), the whole procedure goes as follows. Denote $\bar{F}_t = \mathbf{E}(F_t)$ and $Z_t = \bar{F}_t - F_t$. We may decompose $(I - \mu F_t)$ according to

$$I - \mu F_t = (I - \mu\bar{F}_t) + \mu Z_t. \tag{1.11}$$

Now, decompose the recurrence equations (1.8) into two separate recursions :

$$\begin{aligned}
J_{t+1}^{(0)} &= (I - \mu\bar{F}_t)J_t^{(0)} + \xi_t,\ J_0^{(0)} = 0 && (1.12)\\
H_{t+1}^{(0)} &= (I - \mu\bar{F}_t)H_t^{(0)} + \mu Z_t J_t^{(0)},\ H_0^{(0)} = 0 && (1.13)\\
\delta_t &= J_t^{(0)} + H_t^{(0)}. && (1.14)
\end{aligned}$$

According to (1.12), $J_t^{(0)}$ satisfy a deterministic inhomogeneous first-order linear equation

$$J_{t+1}^{(0)} = \sum_{s=0}^{t} \psi(t,s)\xi_s$$

where, as above,

$$\psi(t,s) = \begin{cases} (I-\mu\bar{F}_t)(I-\mu\bar{F}_t)\cdots(I-\mu\bar{F}_{s+1}), & t>s \\ I, & t=s \\ 0, & \text{otherwise.} \end{cases}$$

Using this technique, $\delta_t$ may be written as

$$\delta_t = J_t^{(0)} + J_t^{(1)} + \cdots + J_t^{(n)} + H_t^{(n)},$$

where the processes $J_t^{(r)}$, $0 \le r \le n$ and $H_t^{(n)}$ are respectively defined as

$$\begin{aligned} J_{t+1}^{(0)} &= (I-\mu\bar{F}_t)J_t^{(0)} + \xi_t \ ; \ J_0^{(0)} = 0 \\ J_{t+1}^{(r)} &= (I-\mu\bar{F}_t)J_t^{(r)} + \mu Z_t J_t^{(r-1)} \ ; \ J_t^{(r)} = 0, \ 0 \le t < r \\ H_{t+1}^{(n)} &= (I-\mu F_t)H_t^{(n)} + \mu Z_t J_t^{(n)} \ ; \ J_t^{(n)} = 0, \ 0 \le t < n \end{aligned}$$

### 1.0.3 Some definitions

**Definition 1 (Weak dependence)** *Let $q \ge 1$ and let $X = \{X_n\}_{n\ge 0}$ be a $(l \times 1)$ vector-valued process. Let $\delta = (\delta(r))_{r\in\mathbb{N}}$ a sequence decreasing to zero at infinity. $X$ is called $(\delta, q)$-weak dependent, if there exists finite constants $C = \{C_1, \cdots, C_q\}$, such that for any $1 \le m < s \le q$, any $m$-tuple $t_1, \cdots, t_m$ and any $(s-m)$-tuple $t_{m+1}, \cdots, t_s$, with $t_1 \le \cdots \le t_m < t_m + r \le t_{m+1} \le \cdots \le t_s$, it holds*

$$\sup_{i_1,\cdots,i_s} |\mathrm{Cov}(X_{t_1,i_1}\cdots X_{t_m,i_m}, X_{t_{m+1},i_{m+1}}\cdots X_{t_s,i_s})| \le C_s\delta(r)$$

*where $X_{n,i}$ denotes the $i$-th component of $X_n$.*

*The notion of weak-dependence, introduced by Doukhan and Louhichi[28]. Weak-mixing processes encompass a large class of models and in particular strongly processes. Weak-dependence is essential in the developments that follow, and appear at several places. The following result plays a key role in the sequel*

**Proposition 1** *Let $p \ge 1$ and $n \in \mathbb{N}$. Let $G = \{G_t\}_{t\ge 0}$ be a $(d \times d)$ zero-mean matrix-valued process. Assume that $(\delta, p(n+2))$ weak-dependent and that*

$$\sum_r (r+1)^{p(n+2)/2-1}\delta(r) < \infty.$$

*Then, there exists a finite constant $D_{p,n}(G)$, such that for all $j \in \{1, \cdots, n\}$ and all $0 \le s \le t < \infty$, we have*

$$\Big\| \sum_{s\le i_1<\cdots<i_j\le t} G_{i_1}\cdots G_{i_j} \Big\|_{pn/j} \le D_{p,n}(G)(t-s)^{j/2}.$$

Let $\mathcal{B}$ be a subfield of the basic probability space $(\Omega, \mathcal{A})$.

**Definition 2 (Rosental's class)** *Let $q \geq p > 0$ be two real numbers. For any $\eta = \{\eta_k\}_{k\geq 0}$ $(d \times 1)$ vector-valued sequence, define $\mathcal{N}(p, q, \mathcal{B})$ as the set*

$$\Big\{\eta : \Big\| \sum_{k=s}^{t} G_k \eta_k \Big\|_p \leq \rho_{p,q}(\eta) \Big( \sum_{k=s}^{t} \|G_k\|_q^2 \Big)^{1/2} \forall 0 \leq s \leq t \, \forall G = \{G_k\}_{k\geq 0} \in L_q(\Omega, \mathcal{B}) \Big\}.$$

The Rosental's class gather all the process that verify a Rosental's type inequality. By direct applications of the Rosental's inequality [28], $\mathcal{N}(p)$ include sequences $\varepsilon = \{\varepsilon_k\}_{k\geq 0}$ of independent random variables with zero-mean and uniformly bounded $p$-th order moment. By application of the Burkholder's inequality for martingales, martingales with $L_p$ stable increments also belong to this class.

**Definition 3 (Exponential stability)** *Define, for $p \geq 1$, $\mu^* > 0$ and $0 < \beta < 1/\mu^*$, $\mathcal{S}(p, \beta, \mu^*)$ the $L_p$ exponentially stable family as the set of random matrix-valued process $\{A_k(\mu)\}_{k\in\mathbb{N}}$ satisfying*

$$\mathcal{S}(p, \beta, \mu^*) = \Big\{ A(\mu) = \{A_k(\mu)\}_{k\in\mathbb{N}} : \Big\| \prod_{j=i+1}^{k} (I - \mu A_j(\mu)) \Big\|_p \quad \leq K_{\beta,\mu^*}(A)(1-\beta\mu)^{k-i} \quad \forall \, \mu \in (0, \mu^*], \; \forall k \geq i \geq 0 \Big\}.$$

*Where the constant $K_{\beta,\mu^*}(A)$ does not depend on $\mu$. Likewise, define the averaged exponentially stable family as the set of deterministic matrix-valued process $\{\bar{A}_k(\mu)\}_{k\in\mathbb{N}}$ :*

$$\mathcal{S}(\beta, \mu^*) = \Big\{ \bar{A}(\mu) = \{\bar{A}_k(\mu)\}_{k\in\mathbb{N}} : \Big| \prod_{j=i+1}^{k} (I - \mu \bar{A}_j(\mu)) \Big| \quad \leq K_{\beta,\mu^*}(\bar{A})(1-\beta\mu)^{k-i} \quad \forall \, \mu \in (0, \mu^*], \; \forall k \geq i \geq 0 \Big\}.$$

$L_p$ exponential stability forms the basis of the stability analysis state space system (see, e.g., [41, 42, 43, 87]. It is a natural extension of the notion of exponential stability for deterministic linear systems.

### 1.0.4 Main results

The first result gives condition upon which $J_s^{(r)}$ is uniformly bounded in $L_p$ and provides an expression for that bound

**Theorem 1** *Assume that, for some integer $n$ and real numbers $q \geq p \geq 2$ :*
*(i) $F = \{F_t\}_{t\geq 0}$ is averaged exponentially stable,*
*(ii) $F$ is $(\delta, q(n+2))$ weakly-dependent, and*

$$\sum_r (r+1)^{p(n+2)/2-1} \delta(r) < \infty.$$

*(iii)* $\xi = \{\xi_t\}_{t\geq 0} \in \mathcal{N}(p, q, \mathcal{B})$

*Then, there exists a constant* $K < \infty$ *(depending on* $F$ *and on the numerical constants* $p, q, n, \mu_0, \beta$ *but not on* $\{\xi_t\}$ *or on the stepsize parameter* $\mu$*), such that for all* $0 < \mu \leq \mu_0$*, for all* $0 \leq r \leq n$

$$\sup_{s\geq 1} \|J_s^{(r)}\|_p \leq K \rho_{p,q}(\xi) \mu^{(r-1)/2}.$$

The second result deals with the $L_p$ bound of the rest $(H_s^{(n)})_{s\geq 0}$

**Theorem 2** *Let* $p \geq 2$ *and let* $a, b, c > 0$ *such that* $1/a + 1/b + 1/c = 1/p$*. Let* $n \in \mathbb{N}$*. Assume that*

- $\{F_t\}$ *is* $L_a$*-exponentially stable,*
- $\sup_{s\geq 0} \|Z_s\|_b < \infty$ *and,*
- $\sup_{s\geq 0} \|J_s^{(n+1)}\|_c < \infty$*.*

*Then, there exists a constant* $K' < \infty$ *(depending on the numerical constants* $a, b, c, \beta, \mu_0, n$ *and on the process* $\{F_t\}$ *but not on the process* $\{\xi_t\}$ *or on the step-size parameter* $\mu$*), such that for all* $0 < \mu \leq \mu_0$*,*

$$\sup_{s\geq 0} \|H_s^{(n)}\|_p \leq K' \sup_{s\geq 0} \|J_s^{(n+1)}\|_c.$$

### 1.0.5 A worked-out example

To illustrate our findings, it is shown in this section that first order approximation of the tracking error covariance may fail, in certain situation, to capture the behavior of the algorithm.

(M1) $\{\phi_t\}_{t\geq 0}$ is VAR process

$$\phi_{t+1} = \kappa \phi_t + u_{t+1}$$

where $\kappa \in ]-1, 1[$ is a scalar, $\{u_t\}_{t\geq 0}$ is i.i.d Gaussian with zero-mean and covariance matrix $\sigma_u^2 I$.

(M2) $\{v_t\}_{t\geq 0}$ and $\{w_t\}_{t\geq 0}$ are i.i.d, with bound moments of order $r$, where $r > 2$. Moreover : $\mathbf{E}(w_0 w_0^T) = \gamma^2 I$.

(M3) $\mathcal{M}_0^\infty(v)$, $\mathcal{M}_0^\infty(\phi)$ and $\mathcal{M}(\theta)$ are independent.

Under assumptions $(M1 - M2 - M3)$, Theorems 1, 2 show that

**Proposition 2** *Assume* $(M1 - M2 - M3)$

$$\Gamma := \lim_{t\to+\infty} \mathbf{E}(\tilde{\theta}_t \tilde{\theta}_t^T) = \mu^2 \Gamma^v + \Gamma^w$$

*where*

$$\begin{aligned}
\Gamma^v &= \frac{\sigma_v^2}{2\mu} I + \alpha \sigma_v^2 \frac{d+2}{4} I + O(\sqrt{\mu}) \\
\Gamma^w &= \frac{\gamma^2}{2\mu\alpha} I + \frac{\gamma^2}{4}\Big(1 + (d+1)\frac{1+2\kappa^2}{1-\kappa^2}\Big) I + O(\gamma^2\sqrt{\mu})
\end{aligned}$$

# Chapitre 2

# Some limit results on random trees

## 2.1 Introduction

Trees are a fundamental object in graph theory and combinatorics as well as a basic object for data structures and algorithms in computer science. During the last years research related to (random) trees has been constantly increasing and several asymptotic and probabilistic techniques have been developed in order to describe characteristics of interest of large trees in different settings.

In this chapter we study distances and weighted path lengths in some kinds of random trees.

## 2.2 Distances in Trees

### 2.2.1 Introduction

Various types of distances in random trees have lately become a topic of interest, as can be seen in half a dozen or so of recent papers. The distance between pairs of nodes in a recursive tree was investigated in [67]. The distance between pairs of nodes in a random binary search tree was investigated in [64, 25] ; Panholzer and Prodinger [71] give a generalization.

The standard data model for digital search trees is the Bernoulli probability distribution (infinitely long independent keys of independent bits). The probability model should ideally be unbiased. In practice this unbias is not guaranteed. So, our study is not limited to the unbiased Bernoulli case, and puts in good perspective the contrast between biased and unbiased data models.

### 2.2.2 Digital trees

There are two flavors of naturally grown digital trees, the digital search tree and the trie. We consider both. Digital trees are suited for digital data, which abound in science,

engineering and technology. For instance, they are the building blocks of computer files. For ease of exposition, we shall deal with the binary case. Generalization to larger alphabets should not be hard. For example, for DNA strands one uses a 4-letter alphabet of protein nucleotides.

Let $\delta_n$ be the depth of a randomly selected node in a random digital tree of size $n$, with *random* meaning that all nodes are equally likely choices. Let $\Delta_n$ be the distance (i.e. the number of tree edges) between two randomly selected keys in a random digital tree of size $n$, where all $\binom{n}{2}$ pairs of keys are equally likely. The recurrence equations for $\Delta_n$ will use $\delta_n$.

### Digital search trees

The digital search tree was invented in Coffman and Eve [19]. In addition to the uses already mentioned as a data structure, the digital search tree provides a model for the analysis of several important algorithms, such as the Lempel-Ziv parsing algorithm (see Louchard and Szpankowski [59]), and Conflict Resolution (see Mathys and Flajolet [65]).

The binary DST grows according to an algorithm. The keys $K_1, K_2 \ldots, K_n$ come in serially. Initially we have an empty tree. For the first key, a root is allocated. The key $K_2$ is guided to the left subtree, where it becomes a left child of the root, if its first bit is 0, otherwise it goes to the right subtree, where it is linked as a right child of the root. Subsequent keys are treated similarly, they are taken into the left or right subtree according as whether the first bit is 0 or 1, and in the subtree the algorithm is applied recursively, but at level $\ell$ of the recursion the $(\ell+1)$st bit is used for guiding the search.

### Tries

The trie was invented independently by De La Briandais[20] and Fredkin [38] for information re*trie*val. A binary trie is a digital tree consisting of internal nodes that each has one or two children, and leaves that hold data. The trie grows from $n$ keys according to a construction algorithm. If $n = 0$, the insertion algorithm terminates. If $n = 1$, a leaf is allocated for the key given. If $n \geq 2$, an internal node is allocated as a root of the tree ; keys starting with 0 go to the left subtree, and keys starting with 1 go to the right. The construction proceeds recursively in the subtrees, but at level $\ell$ the $(\ell+1)$st bit of the key is used for branching. When the algorithm terminates, each key is in a leaf by itself, and the root-to-leaf paths correspond to minimal prefixes sufficient to distinguish the keys.

In addition to the uses already mentioned as a data structure, the trie provides a model for the analysis of several important algorithms, such as Radix Exchange Sort (see [55]), and Extendible Hashing (see [32]).

A main distinction between the algorithm for digital search trees and that for tries is that all the nodes of the digital tree hold keys, whereas in tries the keys reside only in leaves.

### 2.2.3 Notation and methodology

The Mellin transform of a function $f(x)$ is

$$\int_0^\infty f(x)x^{s-1}\,ds,$$

and will be denoted by $f^*(s)$. For a survey of the Mellin transform in the context of the analysis of algorithms we refer the reader to the comprehensive survey in Flajolet, Gourdon and Dumas [33].

Another tool we rely on in the analysis is depoissonization. This method is now standard and we shall not produce the details in any great length. We refer the reader to an original source such as Jacquet and Szpankowski [48], or a textbook such as Szpankowski [88].

Instrumental to our presentation is the functions

$$Q_k(s) = \prod_{j=0}^{k}(1 - p^{j-s} - q^{j-s}),$$

with $q = 1 - p$, and the data entropy

$$h_p = -p\ln p - q\ln q.$$

We shall also need the two functions

$$\tilde{h}_p = p\ln^2 p + q\ln^2 q \qquad \text{and} \qquad \hat{h}_p = p^2\ln p + q^2\ln q.$$

In the sequel the symbol $\gamma$ is Euler's constant.

### 2.2.4 Distances in DST

Our first result concern the moments of $\Delta_n$ in a digital search tree

**Proposition 3** *In a random digital search tree of $n$ random keys, the average distance between two randomly selected keys is*

$$\begin{aligned}\mathbf{E}[\Delta_n] &= \frac{2}{h_p}\ln n + \frac{\hat{h}_p}{pqh_p} + \frac{\tilde{h}_p}{h_p^2} - \frac{2(1-\gamma)}{h_p} + \frac{\ln(pq)}{h_p} \\ &\qquad -\frac{2\alpha_\infty}{h_p} + 2 - 2\xi_p(\ln n) + O\Big(\frac{1}{n}\Big),\end{aligned}$$

*where $\xi_p(.)$ is a small oscillating function. The variance is*

$$\mathbf{V}[\Delta_n] = 2\sigma_p^2\ln n + O(1),$$

*where $\sigma_p^2 = (\tilde{h}_p - h_p^2)h_p^{-3}$.*

**Remark** Except for the symmetric case, the variance grows logarithmically with the number of keys inserted in the tree. In the symmetric case the variance is $O(1)$ (oscillating but uniformly bounded), showing the stiff resistance of inter-node distances in digital search trees to change with the number of keys. In either case we have a concentration law as an immediate corollary (by Chebyshev's inequality).

**Corollary 1** *As* $n \to \infty$,

$$\frac{\Delta_n}{\ln n} \xrightarrow{\mathcal{P}} \frac{2}{h_p}.$$

**Limit laws**

In principle, one can continue pumping higher moments by the methods utilized for the mean and variance, and aspire to determine limit distributions by a method of recursive moments (see [17], for example). However, as already mentioned, the explosive complexity is forbidding.

The contraction method offers a shortcut. Let

$$\Delta_n^* := \frac{\Delta_n - \mathbf{E}[\Delta_n]}{\sqrt{\ln n}}.$$

Based on some heuristics in the structure of the problem, a solution is guessed for the limit distribution of $\Delta_n^*$. The guess is then verified by showing convergence of the distribution function to that of the guessed limit in some metric space. The contraction method was introduced by Rösler [81]. Rachev and Rüschendorf [75] added several useful extensions. Recently general contraction theorems and multivariate extensions were added by Rösler [82], and Neininger [82]. Rösler and Rüschendorf [83] provide a valuable survey. Using this method we prove our principal result

**Theorem 3** *In a digital search tree of* $n$ *random keys following the biased Bernoulli model, the distance* $\Delta_n$ *between two randomly selected keys satisfies*

$$\frac{\Delta_n - \frac{2}{h_p}\ln n}{\sqrt{\ln n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\sigma_p^2).$$

### 2.2.5 Tries

In a random trie, we prove that

**Proposition 4** *In a trie of* $n$ *random keys following the Bernoulli model, the average distance between two randomly selected keys is*

$$\begin{aligned}\mathbf{E}[\Delta_n] &= \frac{2}{h_p}\ln n + \frac{2\gamma}{h_p} + 2 - \frac{1}{pqh_p^2}(p^3\ln^2 p + 2pq\ln p\ln q + q^3\ln^2 q) \\ &\quad +4\beta(n) + o(1),\end{aligned}$$

*where $\beta(n)$ is a small oscillating function. The variance is*

$$\mathbf{V}[\Delta_n] = 2\frac{pq}{h_p^3}(\ln p - \ln q)^2 \ln n + O(1) := 2\tilde{\sigma}_p^2 \ln n + O(1).$$

The case $p = q$ presents a degeneracy, which was handled in [83], where the details of the $O(1)$ term are specified, and where it is proved that no limit exists.

**Corollary 2**

$$\frac{\Delta_n}{\ln n} \xrightarrow{\mathcal{P}} \frac{2}{h_p}.$$

By an argument similar in its general gist to the one we used for the DST, but differing in many of its details we arrive at the main result for inter-distance in random tries.

**Theorem 4** *In a trie of $n$ random keys following the biased Bernoulli model, the distance $\Delta_n$ between two randomly selected keys satisfies*

$$\frac{\Delta_n - \frac{2}{h_p}\ln n}{\sqrt{\ln n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\tilde{\sigma}_p^2).$$

## 2.3 Binary tree

### 2.3.1 Introduction

A binary tree is a hierarchical structure of nodes each having no children, one left child, one right child, or two children (one left and one right). The nodes of such a tree can be labeled from some ordered set, say the natural numbers. The tree can further be endowed with a *search property* (to support fast searching of the items (also called *keys*) stored in it), which imposes the restriction on the labeling scheme that the label of any node is larger than the labels in its left subtree and no greater than any label in its right subtree. For definitions and combinatorial properties see [60], and for applications in sorting see [55, 63].

Suppose we want to store in a computer a set of data, whose elements can be compared by an ordering relation like $\leq$ . One common choice for the data structure is the binary search tree constructed by the following algorithm. The data are fed serially to the computer as an input stream. A root node is created for the first element ; then subsequent elements are guided to the left or right subtree, according to whether they are less than the root label or not, where they are subjected recursively to the same treatment, until a unique insertion position is found. As an example, suppose we want to grow a binary search tree from the input sequence $(5, 8, 7, 3, 9, 1, 6, 2, 4)$. The first item 5 is placed in the root node.

The next item is 8 and it goes to the right of the root because $8 \geq 5$. When 7 comes along it is first compared with the label of the root, and as $7 \geq 5$, it goes to the right subtree and is compared with 8, and because $7 \leq 8$ it is guided to the left of 8, and that is where 7 is inserted. The process continues with the rest of the numbers until we end up with the last tree of Figure 2.1.

Several models of randomness are in common use on binary trees. The uniform model in which all trees are equally likely has been proposed for applications in formal languages, compilers, computer algebra, etc. (see [54]). However, for the searching and sorting algorithms alluded to the *random permutation model* is considered to be more appropriate. In this model of randomness we assume that the tree is built from permutations of $\{1, \ldots, n\}$, where a uniform probability model is imposed on the *permutations* instead of the trees. When all $n!$ permutations are equally likely or *random*, binary search trees are not equally likely. Several permutations give rise to the same tree, favoring shorter and well balanced trees rather than scrawny and tall shapes, which is a desirable property in searching and sorting algorithms (see [60]). The term *random tree* (and occasionally just the tree) will refer to a binary search tree built from a random permutation. The random permutation model is not really restrictive, as it covers a rather wide variety of instances, such as when the input is a sample drawn from *any* continuous probability distribution, and the construction algorithm is concerned only with the ranks of the keys, not their actual values.

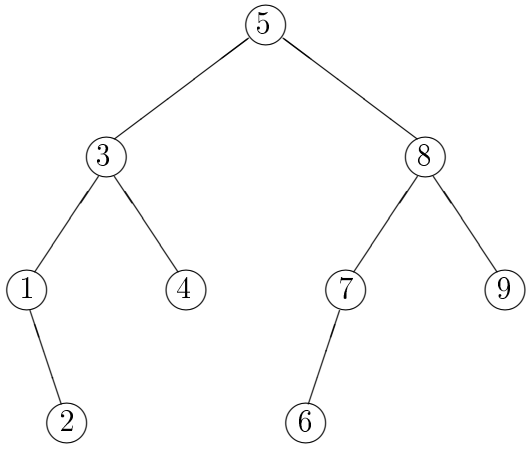


FIG. 2.1 – A binary search tree.

### 2.3.2 Weighted path to the minimal label

We denote by $L_n$ the number of items that appear in the left subtree, and thus $L_n + 1$ is the value of the root. For $n \geq 1$, the weighted path length $W_1(n)$ from the root to the

node labeled 1 (that is, the sum of the collection of values on the leftmost path in the tree) satisfies the stochastic recurrence

$$W_1(n) = L_n + 1 + W_1(L_n).$$

**Proposition 5** *Let $W_1(n)$ be the weighted path length from the root to the least ranked label in a binary search tree built from a random permutation. We then have*

$$\begin{aligned} \mathbf{E}[W_1(n)] &= n; \\ \mathbf{V}[W_1(n)] &= \frac{n(n+1)}{2} \sim \frac{1}{2}n^2. \end{aligned}$$

Guided by the rate of growth of the variance, we next proceed to argue the infinite divisibility of $n^{-1}W_1(n)$. We obtain a convergence law theorem

**Theorem 5** *Let $W_1(n)$ be the weighted path length from the root to the least ranked label in a binary search tree built from a random permutation. Then,*

$$\frac{W_1(n)}{n} \xrightarrow{\mathcal{D}} 1 + X,$$

*where $X$ is Dickman's random variable.*

**Remark :** The limiting random variable for $n^{-1}W_1(n)$ bears some similarity to the limiting random variable for $n^{-1}C_n^{[1]}$, the normalized number of comparisons required by Quickselect to find the least item in a random input (of size $n$) with ranks following the random permutation model. It is shown in [62] that $n^{-1}C_n^{[1]}$ converges in distribution to $2 + X$. Thus, asymptotically, the distribution of $n^{-1}C_n^{[1]}$ behaves like that of $1 + n^{-1}W_1(n)$.

We study the weighted path length leading to the rightmost and leftmost nodes. For instance, in the tree of Figure 2.1, $W_1(9) = 5 + 3 + 1 = 9$, $W_2(9) = 5 + 3 + 1 + 2 = 11$, ..., $W_9(9) = 5 + 8 + 9 = 22$.

### 2.3.3 Weighted path to the maximal label

In the set $\mathcal{T}_n$ of binary search trees of size $n$, we introduce a reflection operator $\mathcal{R}e$ defined as follows :

$$\begin{aligned} \mathcal{R}e : \mathcal{T}_n &\longrightarrow \mathcal{T}_n \\ \text{a node with key} : K &\longmapsto \text{a node with key} : n + 1 - K. \end{aligned}$$

That is, to obtain the reflected tree $T'_n := \mathcal{R}e(T_n)$ of a binary search tree of size $n$, exchange the right and left children of every node, starting at the root and progressing recursively

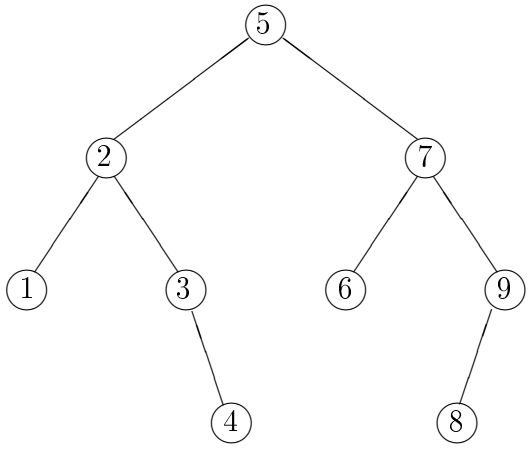


FIG. 2.2 – The reflection of the tree of Figure 2.1

toward the leaves. This reflection concerns only the shape of the tree, and not the labels. To maintain the binary search property in $T'_n$, we reinsert the numbers $1, \ldots, n$ in a manner consistent with the search property. For example, the reflected tree of that in Figure 2.1, is shown in Figure 2.2.

Because there are the same number of permutations of $\{1, 2, \ldots, n\}$ producing $T_n$ as those producing $\mathcal{R}e(T_n)$, we have the following lemma :

**Lemma 1**

$$\mathbf{P}\Big(T_n\Big) = \mathbf{P}\Big(\mathcal{R}e(T_n)\Big).$$

Let the length of the rightmost path in $T_n$ be $Q_n$, and suppose the chain of values appearing on it from the root to the rightmost node (containing $n$) is $Y_1, Y_2, \ldots, Y_{Q_n+1}$. Observe that the rightmost path in $T_n$ becomes a leftmost one (of the same length) in $T'_n$, and suppose that the corresponding labels in the reflection are $Y'_1, Y'_2, \ldots, Y'_{Q_n+1}$. This connection suggests that we can use the distribution of the path to the minimal value, which was established in Section 2.3.2, for the rightmost path as follows. We have

$$W_n(n) = \sum_{j=1}^{Q_n+1} Y_j \stackrel{\mathcal{L}}{=} \sum_{j=1}^{Q_n+1} (n+1-Y'_j) = (Q_n+1)(n+1) - W_1(n). \tag{2.1}$$

In one hand we have the following result do to Devroye [24]

**Theorem 6 (Devroye (1988))**

$$\frac{Q_n - \ln n}{\sqrt{\ln n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

On the other hand, centering and scaling the relation (2.1) with asymptotic mean and standard deviation of $nQ_n$ will yield a limit distribution. Let

$$W_n^* := \frac{W_n(n) - n\ln n}{n\sqrt{\ln n}} \stackrel{\mathcal{L}}{=} \frac{nQ_n - n\ln n}{n\sqrt{\ln n}} + \frac{Q_n}{n\sqrt{\ln n}} + \frac{n+1}{n\sqrt{\ln n}} - \frac{W_1(n)}{n\sqrt{\ln n}}.$$

According to Theorem 5, we have

$$\frac{W_1(n)}{n\sqrt{\ln n}} \stackrel{a.s.}{\longrightarrow} 0,$$

indicating that the main contribution in $W_n(n)$ comes from the length of the rightmost path. Also, according to the limit law of $Q_n$, $Q_n/(n\sqrt{\ln n}) \stackrel{a.s.}{\longrightarrow} 0$, and of course $(n+1)/(n\sqrt{\ln n}) \stackrel{a.s.}{\longrightarrow} 0$. Hence,

**Theorem 7**

$$W_n^* \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}(0,1).$$

### 2.3.4 Some general questions

Numerate the external nodes from left to Wright from 1 to $n+1$. Select an external node at random in the tree, its index $K_n$ will uniformly distributed on the set $\{1, \cdots, n\}$. Define its path length : $W_{K_n}(n) = \sum_{j=1}^{n} j\mathbf{1}_{\{j \triangleright K_n\}}$, where the event $\{j \triangleright k\}$ denote that the label $j$ is encountered on the path to $k$. We prove that $\mathbf{E}[W_{K_n}(n)] \sim (n+1)\ln(n+1)$. Define $W(n) := \sum_{k=1}^{n+1} j\mathbf{1}_{\{j \triangleright K\}}$. What can be said a bout the limit law of

(1) $W_{K_n}(n)$ ?

(2) $W(n)$ ?

# Chapitre 3

# Some Limit Theorems

In this chapter we present some limit theorems corresponding to sums of some weakly dependant random variables, Polya urns, random fragmentation and pattern matching problem.

## 3.1 Asymptotic behavior for sums of some weakly dependant random variables

### 3.1.1 introduction

Let $(X_{i,n})_{0\leq i\leq n}$ be a triangular array of stationary row-wise, centered and square integrable real valued random variables. We suppose that this sequence satisfies some weak dependance condition. Our aim is to find a sequence of independent random variables $(\tilde{X}_{i,n})$ such that the two sums $S_n = \sum_{i=1}^n X_{i,n}$ and $\tilde{S}_n = \sum_{i=1}^n \tilde{X}_{i,n}$ have the same asymptotic limiting behavior in distribution. Such a sequence $(\tilde{X}_{i,n})$ will be said as Jakubowski's terminology [cf. [49]] *an asymptotic independent representation* (a.i.r.) for sums of $(X_{i,n})$. For this purpose, we suppose that the sequence $(X_{i,n})_{0\leq i\leq n}$ is *associated*

**Definition 4** *A vector $X = (X_1, X_2, ..., X_n)$ is associated if for all non-decreasing functions $f, g$ defined on $\mathbb{R}^n$*

$$\mathrm{Cov}(f(X), g(X)) \geq 0. \tag{3.1}$$

For this definition and for its main properties, we refer the reader to Esary *et al.*[30]. For such sequences, we have the well known Newman's inequality [cf. [69]] :

**Theorem 8 (Newman inequality)**

$$\left| \mathbb{E}\left(\exp\left(izS_n\right)\right) - \prod_{j=1}^{n} \left(\mathbb{E}\exp(izX_{j,n})\right) \right| \leq \frac{z^2}{2}\left(\mathrm{Var}S_n - n\mathrm{Var}X_{0,n}\right). \tag{3.2}$$

As a consequence of the last Theorem, if we consider a sequence of i.i.d random variables $\tilde{X}_{i,n}$ distributed like $X_{0,n}$, then we can prove easily that

**Corollary 3** *Under the Condition*

$$\lim_{n\to+\infty} n\sum_{r=1}^{n}\mathrm{Cov}(X_{0,n},X_{r,n})=0, \tag{3.3}$$

$(\tilde{X}_{i,n})_{0\le i\le n}$ *is an asymptotic independent representation for sums of* $(X_{i,n})_{0\le i\le n}$.

We assume, instead of (3.3), that

$$\lim_{N\to N_0}\limsup_{n\to+\infty} n\sum_{r=N}^{n}\mathrm{Cov}(X_{0,n},X_{r,n})=0,\quad \lim_{n\to+\infty}\mathrm{Var}S_n=:\sigma^2, \tag{3.4}$$

where $\sigma^2$ is a finite positive real number.
Under this condition (3.4), our task is to describe the asymptotic independent representations of $S_n$. For this, we need to define the following class of functions :

**Notations :** Let $\eta>0$ be fixed and consider the two continuous and 1-Lipschitz decoupling functions $f_\eta$ and $f^{(\eta)}$ defined respectively by
$f_\eta(x)=(\eta\wedge x)\vee(-\eta)$ and $f^{(\eta)}(x)=x-f_\eta(x)$.
Set

$$X_{i,n}(\eta)=f_\eta(X_{i,n})-\mathbb{E}(f_\eta(X_{i,n})),\quad X_{i,n}^{(\eta)}=X_{i,n}-X_{i,n}(\eta),$$

and let, $S_n^{(\eta)}$, and $S_{n,\eta}$ be respectively the sums

$$S_n^{(\eta)}=\sum_{i=1}^{n}X_{i,n}^{(\eta)},\quad S_{n,\eta}=\sum_{i=1}^{n}X_{i,n}(\eta).$$

Clearly $S_n=S_n^{(\eta)}+S_{n,\eta}$. Finally, for any $t\in[0,1]$, we denote by $S_n(t)=\sum_{i=1}^{[nt]}X_{i,n}$, where [.] denotes the integer part as usual.
We consider the set

$$\widetilde{\mathcal{F}}_{0,C}=\left\{x\to\cos(zx)-1-\frac{z^2x^2}{2},\quad x\to\sin(zx)-zx,\quad \text{for}\quad z\in\mathbb{R}\right\} \tag{3.5}$$

### 3.1.2 Main results

Our main result is the following.

**Theorem 9** *Let* $(X_{i,n})$ *be a triangular array of stationary, centered and associated square integrable real valued random variables, such that*

$$\sup_{n\in\mathbb{N}} n\mathbb{E}(X_{0,n}^2)<\infty.$$

*Suppose that (3.4) holds. Let $(\widetilde{X}_{i,n})$ be a sequence of centered and i.i.d. random variables, for which*

$$\lim_{N\to N_0} \limsup_{\eta\to 0} \limsup_{n\to+\infty} \sup_{h\in\ \widetilde{\mathcal{F}}_{0,C}} n\left|\mathbb{E}h(\widetilde{X}_{0,n}^{(\eta)}) - \mathbb{E}\left[h(S_{N,n}^{(\eta)}) - h(S_{N-1,n}^{(\eta)})\right]\right| = 0, \tag{3.6}$$

*and*

$$\lim_{n\to+\infty} n\mathrm{Var}\widetilde{X}_{0,n} = \sigma^2. \tag{3.7}$$

*Then $(\widetilde{X}_{i,n})$ is an a.i.r. for sums of $(X_{i,n})$.*

### 3.1.3 Discussions

Let us now explain how we can deduce a central limit theorem for $S_n$. Let $(\widetilde{X}_{i,n})_{i\in\mathbb{N},n\in\mathbb{N}}$ be a stationary array of i.i.d. square-integrable and centered random variables. Suppose that $n\mathbb{E}(\widetilde{X}_{0,n}^2)$ converges to some positive real number $\sigma^2$ as $n$ tends to infinity. If moreover

$$\lim_{n\to+\infty} n\mathbb{E}\left(\widetilde{X}_{0,n}^2 \mathbb{1}_{|\widetilde{X}_{0,n}|\geq\epsilon}\right) = 0 \quad \text{for all} \quad \epsilon > 0, \tag{3.8}$$

then we know that $\widetilde{S}_n$ converges in distribution to a centered normal law with variance $\sigma^2$ (cf. for instance [15]). If this sequence $(\widetilde{X}_{i,n})_{(i,n)\in\mathbb{N}^2}$ fulfills moreover Conditions (3.6) and (3.7) of Theorem 9, then we deduce from the conclusion of Theorem 9 that $S_n$ converges also in distribution to a centered normal law with variance $\sigma^2$.

So our task is to give sufficient conditions on the sequence $(X_{i,n})_{(i,n)\in\mathbb{N}^2}$ under which Conditions (3.6) and (3.7) of Theorem 9 are satisfied. Clearly Condition (3.7) is satisfied as soon as $\mathrm{Var}S_n$ converges to $\sigma^2$ as $n$ goes to infinity. In the other hand, we prove that Condition (3.6) is satisfied as soon as

$$\lim_{n\to+\infty} n\mathbb{E}\left(X_{0,n}^2 \mathbb{1}_{|X_{0,n}|\geq\epsilon}\right) = 0 \quad \text{for all} \quad \epsilon > 0. \tag{3.9}$$

This is can be summarized by the following corollary

**Corollary 4** *Let $(X_{i,n})_{(i,n)\in\mathbb{N}^2}$ be a stationary array of associated square-integrable and centered random variables. Suppose that $\mathbb{E}(X_{0,n}^2)$ tends to zero as $n$ tends to infinity and that Conditions (3.4) and (3.9) are satisfied with $\sigma^2 = 1$. Then $S_n$ converges in distribution to the standard normal law.*

**Example :** We suppose now that $X_{i,n} = \frac{X_i}{B_n}$ where $(X_i)$ is a centered sequence of stationary and associated random variables having finite second moment and $B_n = \sqrt{\mathrm{Var}(X_1 + \ldots + X_n)}$. We suppose moreover that this sequence is of $m$-dependent random variables, fulfilling (3.4) and

$$\lim_{\eta\to 0} \limsup_{n\to+\infty} n\mathbb{E}(X_{0,n}(\eta))^2 = 0. \tag{3.10}$$

The following corollary gives sufficient conditions for the existence of an i.i.d. sequence $\widetilde{X}_{i,n}$ as described in Theorem 9.

**Corollary 5** *Let $(X_i)$ be a sequence of m-dependent associated random variables as defined in the last example, fulfilling (3.10) and (3.4). Let $(\widetilde{X}_{i,n} = \frac{\widetilde{X}_i}{B_n})$, where $\widetilde{X}_i$ is a sequence of centered and i.i.d. random variables, for which*

$$\lim_{\eta \to 0} \limsup_{n \to +\infty} n\mathbb{E}(\widetilde{X}_{0,n}(\eta))^2 = 0. \tag{3.11}$$

*If the characteristic function $\mathbb{E}(e^{iz\widetilde{X_1}})$ of $\widetilde{X_1}$ fulfills*

$$\mathbb{E}(e^{iz\widetilde{X_1}}) = \frac{\mathbb{E}(e^{iz(X_1+\ldots+X_{m+1})})}{\mathbb{E}(e^{iz(X_1+\ldots+X_m)})}, \tag{3.12}$$

*then Conditions (3.6) and (3.7) are satisfied and the conclusion of Theorem 9 applies.*

## 3.2 Polya urns

### 3.2.1 Introduction

A Polya urn is an urn containing balls of up $k$ different colors. The urn evolves in discrete time steps. At each step, a ball is sampled uniformly at random (all balls being equally likely).The color of the ball withdrawn is observed, and the ball is returned to the urn. If at any step the color of the ball withdrawn is $i = 1, \cdots, k$, then $A_{i,j}$ balls of color $j$ are placed in the urn, $j = 1, \cdots, k$, where $A_{i,j}$ follows a discrete distribution on a set of integers. Generally speaking, the entries $A_{i,j}$ can be deterministic or random, positive or negative. It is customary to represent the urn scheme by a square ball addition matrix, or schemata :

$$B = [A_{i,j}], \quad i, j = 1, \cdots, k,$$

the rows of which are indexed by the color of the ball picked, and the columns are indexed by the color of the balls added. The primary interest lies in the long-term composition of the urn and in the stochastic path leading to it. So, the number of balls of each color and the number of splits of a particular color ( the number of times a ball of that color is drawn) are examples of important parameters.

Our interest is the case of $2 \times 2$ schemata. We suppose that we have two colors white and black with the top row corresponding to additions upon drawing a white ball, and the first column corresponding to the number of white balls added. We shall use $W_n$ (resp. $N_n^W$) to denote the number of white balls (resp the number of white splits ) after $n$ draws, and $B_n$ (resp. $N_n^B$) to denote the number of black balls (resp. the number of black splits ) after $n$ draws, with $W_0$ and $B_0$ being the initial conditions. The total number of balls after $n$ draws is $S_n = W_n + B_n$ and it is clear that $N_n^W + N_n^B = n$.

Several Polya urn models with various settings for the ball addition matrix have been studied by many authors. In particular, when the mean of the replacement matrix is irreducible, Janson [50] carries out a study in which he characterizes the number of balls of each colors. When the replacement matrix is triangular, not irreducible and non random, Janson [51] characterizes the limit law and the almost sure limits of both $W_n$ and $B_n$.

To extend this results we use the so called "Embedding method" do to Athreya and Karlin [9, 10].

### 3.2.2 Embedding method

The idea of embedding discrete urn models in continuous time branching processes goes back at least to Athreya and Karlin [9]. A description is given in Athreya and Ney ([9], section 9). The method has been recently revisited and developed by Janson [50].

We define the continuous time Markov branching process $\mathcal{X}(t) := (W(t),\, B(t),\, t \geq 0)$ as being the embedded process of $(W_n, B_n,\, n \geq 0)$. One gets this way a branching process whose dynamical description in terms of white and black balls in an urn is the following. In the urn, at any moment, each ball is equipped with an $\mathcal{E}xp(1)$-distributed random clock, all the clocks being independent. When the clock of a white ball rings, a random number, with law as $A_{1,1}$, white balls and a random number, with the same law that $A_{1,2}$, black balls are added in the urn ; when the ringing clock belongs to a black ball, one adds a number, with the same law that $A_{2,1}$, white balls and a random number, with the same law as $A_{2,2}$, black balls, so that the replacement rules are the same as in the discrete time urn process.

The successive jumping times of $\mathcal{X}(t) := (W(t),\, B(t),\, t \geq 0)$, will be denoted by

$$0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots$$

The $n$th jumping time is the time of the $n$th dislocation of the branching process. The process is thus constant on any interval $[\tau_n, \tau_{n+1}[$.

**Lemma 2 (Athreya and Karlin [9])** *The process* $(\mathcal{X}(\tau_n))_{n\geq 0}$ *has the same law as* $((W_n,\, B_n))_{n\geq 0}$ *and* $\tau_n \overset{a.s}{\to}_n +\infty$.

**Remarks :**

(a) As a consequence of the last Lemma, limit theorems for $(W_n,\, B_n)$ can be derived from limit theorems for $\mathcal{X}(t)$.

(b) From now on, thanks to last Lemma, we will classically consider that the discrete time process and the continuous time process are built on a same probability space on which

$$\mathcal{X}(\tau_n) := (W(\tau_n),\, B(\tau_n),\, n \geq 0) = (W_n, B_n,\, n \geq 0)\ a.s..$$

In order to study the process $\mathcal{X}(t) = (W(t), B(t))$, we define a suitable martingale $\mathcal{Y}(t)$. The martingale that we use, is a standard one in branching process theory. Let

$$A^T = \Big(\mathbf{E}[A_{i,j}],\ 1 \leq i, j \leq 2\Big) \text{ and } \mathcal{Y}(t) = e^{-tA}\mathcal{X}(t)^T.$$

Our study concerns the triangular urn, which means, almost surely, that $A_{1,2} = 0$. For all $i, j = 1, 2$ let $\mu_{i,j} = \mathbf{E}(A_{i,j})$ and $\sigma_{i,j} = \mathbf{V}(A_{i,j})$.

Using the Markov property, in the following theorem we extend a fundamental well-known result described in [Lemma 9.8 [50], and in Theorem V.7.2 [10]].

**Theorem 10** *If $\mu_{1,1} \geq \mu_{2,2}$, the Martingale $\{\mathcal{Y}(t), \mathcal{F}_t, t \geq 0\}$ is an $L^2$ bounded martingale, and hence converges almost surely and in $L^2$. Moreover, $\mathcal{Y}(t) = \mathbf{E}(\widetilde{\mathcal{Y}}/\mathcal{F}_t)$, where $\widetilde{\mathcal{Y}} := \begin{pmatrix} W \\ B \end{pmatrix}$ is the almost sure and $L^2$ limit of $\mathcal{Y}(t)$.*

### 3.2.3 Asymptotic composition of the discrete urn

Our basic assumptions for this model are the following

**Assumptions (H) :**

Our basic assumptions are as follows

**(H1)** $(A_{i,j})_{1\leq i,j\leq 2}$ are nonnegative independent integer-valued random variables satisfying

$(*)$ $\sigma_{1,1} + \sigma_{2,1} + \sigma_{2,2} < \infty$,

$(**)$ $\mu_{1,1} \geq \mu_{2,2}$,

$(***)$ $\mu_{1,1}(\mu_{2,1} + \mu_{2,2}) > 0$.

**(H2)** The initial composition of the urn is $(W_0, B_0)$, with $B_0 > 0$.

**Remarks**

(a) The case when $\mu_{1,1} < \mu_{2,2}$ has been studied by Janson [51].

(b) The assumption that $(A_{i,j})_{1\leq i,j\leq 2}$ are nonnegative integer-valued random variables, guarantee the non-extinction of the urn.

(c) The assumption (**H1**)$(*)$ is to avoid any explosion of the urn.

Kotz, Mahmoud and Robert [56] give exact formulas and some asymptotics for 2-type urns. They comment that the case $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ gives asymptotics "of an essentially different character". They prove, "heuristically", that $\mathbf{E}(B_n)$ is of order $\dfrac{n}{\ln n}$. This is expressed in the following theorem.

**Theorem 11** *Consider a generalized Pólya urn with two colors having a triangular random replacement matrix*

$$\begin{pmatrix} A_{1,1} & 0 \\ A_{2,1} & A_{2,2} \end{pmatrix}.$$

*Let* $\rho = \frac{\mu_{2,2}}{\mu_{1,1}}$ *and* $K = (\mu_{1,1})^\rho B \left(W + \frac{\mu_{2,1}}{\mu_{1,1}-\mu_{2,2}} B\right)^{-\rho}$,
*under hypothesis **H** and the fact that $\mu_{2,1} > 0$, we obtain*

1. *in the case $\mu_{1,1} > \mu_{2,2} > 0$,*
   (a) *almost surely*

$$\begin{aligned} W_n &= \mu_{1,1}\, n + o(n), & B_n &= Kn^\rho + o(n^\rho) \\ N_n^W &= n - \frac{K}{\mu_{2,2}} n^\rho + o(n^\rho), & N_n^B &= \frac{K}{\mu_{2,2}} n^\rho + o(n^\rho). \end{aligned}$$

*(b) if $\rho \leq \frac{1}{2}$, then,*

$$\frac{W_n - n\mu_{1,1}}{\sqrt{n}} - \frac{K}{\mu_{2,2}}(\mu_{2,1} - \mu_{1,1})\mathbb{1}_{\{\rho=\frac{1}{2}\}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_{1,1}^2),$$

*(c) if $\frac{1}{2} < \rho < 1$,*

$$\frac{W_n - n\mu_{1,1}}{n^\rho} \xrightarrow{\mathbf{P}} \frac{K}{\mu_{2,2}}(\mu_Y - \mu_{1,1}).$$

*2. if $\mu_{1,1} = \mu_{2,2} > 0$, we have almost surely*

$$\begin{aligned} W_n &= \mu_{1,1} n + o(n), \\ B_n &= \frac{\mu_{1,1}^2}{\mu_{2,1}} \frac{n}{\ln n} + o\Big(\frac{n}{\ln n}\Big). \end{aligned}$$

The next theorem deals with the case $A_{2,2} \overset{a.s.}{=} 0$. In this case we have $B_n \equiv B_0$. For the asymptotic of $W_n$, we have the following result

**Theorem 12** *Consider a generalized Pólya urn with two colors having a triangular random replacement matrix*

$$\begin{pmatrix} A_{1,1} & 0 \\ A_{2,1} & 0 \end{pmatrix}.$$

*Under (H) we have,*

*1. almost surely : $W_n = \mu_{1,1} n + o(n)$,*

*2. if $\sigma_{1,1} \neq 0$, then we have the central limit theorem,*

$$\frac{W_n - \mu_{1,1} n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_{1,1}^2),$$

*3. if $A_{1,1} = \alpha \neq 0$ is non random, then*

$$\frac{1}{\sqrt{\ln n}}\Big(W_n - n\alpha - \frac{\mu_{2,1} - \alpha}{\alpha} B_0 \ln n\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \Big(\sigma_{2,1}^2 + (\mu_{2,1} - \alpha)^2\Big)\frac{B_0}{\alpha}\Big).$$

The next result deals with the diagonal case : $A_{2,1} \overset{a.s}{=} 0$.

**Theorem 13** *Consider a generalized Pólya urn with diagonal replacement matrix*

$$\begin{pmatrix} A_{1,1} & 0 \\ 0 & A_{2,2} \end{pmatrix}.$$

*Let $\rho = \frac{\mu_{2,2}}{\mu_{1,1}}$ and $D = \mu_{1,1}^\rho \, B \, W^{-\rho}$, then*

*1. when $\mu_{1,1} > \mu_{2,2} > 0$,*

*(a) almost surely as* $n \to +\infty$

$$W_n = \mu_{1,1}\, n + o(n), \quad B_n = D\, n^{\rho} + o(n^{\rho}),$$

*(b) if* $0 < \rho \leq \frac{1}{2}$*, we have as* $n \to +\infty$

$$\frac{W_n - n\mu_{1,1}}{\sqrt{n}} + \sqrt{\mu_{1,1}}\frac{\mu_{1,1}}{\mu_{2,2}} B W^{-1/2} \mathbb{1}_{\{\rho=\frac{1}{2}\}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_{1,1}^2),$$

*(c) if* $\frac{1}{2} < \rho < 1$*, we have as* $n \to +\infty$

$$\frac{W_n - n\mu_{1,1}}{n^{\rho}} \xrightarrow{\mathbf{P}} -\mu_{1,1}^{\rho}\frac{\mu_{1,1}}{\mu_{2,2}} B W^{-\rho}.$$

2. *when* $\mu_{1,1} = \mu_{2,2}$ *almost surely as* $n \to +\infty$

$$W_n = \mu_{1,1}\frac{W}{W+B} n + o(n), \qquad B_n = \mu_{1,1}\frac{B}{W+B} n + o(n).$$

### 3.2.4 A project work

1. It would be fine to relax the $L^2$ assumption on $A_{1,1}$, $A_{1,2}$ and $A_{2,1}$ and to replace it by some "$(A_{i,j} \ln A_{i,j},\ i, j)$" hypothesis. Do one still get an almost sure and $L^1$ convergence of the martingale $\{\mathcal{Y}(t),\, \mathcal{F}_t,\, t \geq 0\}$ in such a case ?
2. What can be said on the variable $B$ ?
3. What about the joint law $(W_n,\ B_n)$ ?
4. It would be interesting to extend the results to cases with more two colors and where conditions $(A4)$ or $(A6)$ in [50] does not hold.

## 3.3 Random fragmentation

### 3.3.1 Introduction

Fragmentation is a widely studied phenomena [78] with applications ranging from conventional fracture of solids [58] and collision induced fragmentation in atomic nuclei/ aggregates [16] to seemingly unrelated fields such as disordered systems [21] and geology. Fragmentation processes are also relevant to spin glasses, Boolean networks and genetic population.
We are concerned with two classes of stochastic fragmentation processes : homogeneous and exponential fragmentation.

### 3.3.2 Homogeneous random fragmentation process

Let $(\xi_1, \xi_2, \cdots, \xi_m)$ be a given random vector. We assume that each $\xi_j$ has a distribution that is absolutely continuous on $(0,1)$ with density $f_j$ and

$$\sum_{j=1}^{m} \xi_j \stackrel{a.s}{=} 1.$$

In certain collision processes (see Feller, 1971, page 25), an unstable particle, with size 1, splits into $m$ offspring fragments with some probability $p$. If splitting occurs, the parental mass is shared between them at random, according to different partition laws $(\xi_1, \xi_2, \cdots, \xi_m)$. With complementary probability $1-p$, splitting does not take place and, as a result, stable parental fragments are left unchanged for ever. This splitting process is then iterated independently on first generation sub-fragments and so on until exhaustion of the fragmentation process, as in the context of Galton-Watson trees with positive probability of becoming extinct. Similar processes have been investigated by Krapivsky, Ben-Naim and Grosse [57]. The term homogeneous refers to the fact that in such models, the splitting probability is independent of fragment sizes at each step.

For $t \in [0,1]$, let $K(t)$ be the number of stable fragments of length less than $t$ at the end of the fragmentation process, and $M(t) = \mathbf{E}(K(t))$.
For $y \geq 0$ let $K(y,t)$ be the total number of stable fragments of length less than $t$ given by a fragment of length $y$ and let $M(y,t) = \mathbf{E}(K(y,t))$. Using the recursive nature of the process we obtain, almost surely,

$$\begin{aligned} K(t) &= \sum_{j=1}^{m} \mathbb{1}_{\{j^{th}\ stable\}} \mathbb{1}_{\{\xi_j \leq t\}} + \sum_{j=1}^{m} \mathbb{1}_{\{j^{th}\ unstable\}} K(\xi_j, t) \\ &= \sum_{j=1}^{m} \mathbb{1}_{\{j^{th}\ stable\}} \mathbb{1}_{\{\xi_j \leq t\}} + \sum_{j=1}^{m} \mathbb{1}_{\{j^{th}\ unstable\}} \mathbb{1}_{\{\xi_j \leq t\}} K(\xi_j, t) \\ &\quad + \sum_{j=1}^{m} \mathbb{1}_{\{j^{th}\ unstable\}} \mathbb{1}_{\{\xi_j > t\}} K(\xi_j, t). \end{aligned}$$

By Mellin transformation, we prove that, the Mellin transform of $M$ exists in the half of complex plane

$$\mathcal{S} = \{s \in \mathbb{C},\ \mathcal{R}e(s) > 0\}.$$

Denoting, for each $j \leq m$, by ${f^*}_j$ the Mellin transform of $f_j$. The Mellin transform of $M$ reads

$$M^*(s) = \frac{\mathbf{E}(N)\Big(1 - \frac{1}{m}\sum_{j=1}^{m} {f^*}_j(s+1)\Big)}{s(1 - p\sum_{j=1}^{m} {f^*}_j(s+1))},\ \forall s \in \mathcal{S}. \tag{3.13}$$

Using the Mellin inverse, we deduce that, for all $c > 0$ and $t \in [0,1]$

$$M(t) = \frac{1}{2i\pi} \int_{c-\infty}^{c+i\infty} M^*(s) t^{-s} ds.$$

**Application**

Each unstable fragment with size $l$ splits into 3 fragments, using 2 independent, uniformly distributed cut points in the interval $[0, l]$. We get an explicit expression for $M(t)$. In fact

$$\begin{aligned} M(t) &= 3q\mathbf{P}(\xi_1 \le t) + 6p\int_0^t M(s,t)(1-s)ds + 6pt\int_t^1 M(s,t)(1-s)ds \\ &= 3q\mathbf{P}(\xi_1 \le t) + 6p\mathbf{E}(N)\int_0^t (1-s)ds + 6p\int_t^1 M(t/s)(1-s)ds \\ &= 3qt(2-t) + 6pt(1-\frac{t}{2})\mathbf{E}(N) + 6pt\int_t^1 \frac{M(y)}{y^2}dy - 6pt^2\int_t^1 \frac{M(y)}{y^3}dy. \end{aligned}$$

Using the Mellin inverse and (3.13), it leads to the following Theorem

**Theorem 14** *For all $t \in [0,1]$,*

$$M(t) = \begin{cases} \frac{3q}{(1-3p)\sqrt{1+24p}} t^{\frac{3+\sqrt{1+24p}}{2}} \left[\frac{3+\sqrt{1+24p}}{2} t^{-\sqrt{1+24p}} + \frac{\sqrt{1+24p}-3}{2}\right], & \text{if } p < \frac{1}{3} \\ +\infty, & \text{if } p \ge \frac{1}{3}. \end{cases}$$

### 3.3.3 Exponential fragmentation probability

We assume that the first segment is of length $x$ (large enough) and the probability $p(s)$ that a new fragment remains unstable depends on the fragment size $s$. This is relevant for impact fragmentation and DNA segmentation where fragments have an intrinsic size scale below which the fragmentation probability becomes negligible.
If $p(s) = \mathbb{1}_{s\ge 1}$, $m = 2$ and $\xi$ is uniform on $[0,1]$, this model has been studied by Sibuya and Itoh [86]. Recently Janson and Neininger [52] studied the general case where $p(s) = \mathbb{1}_{s\ge 1}$, $m \ge 2$ and the support of the distribution of $(\xi_1, \xi_2, \cdots, \xi_m)$ on the standard simplex has an interior point.

We assume, in our model, that $m = 2$, $\xi_1$ is uniform on $[0,1]$ and

$$p(s) = 1 - e^{-s}.$$

This model can help to approximate the binary search tree by a fragmentation tree (see [52]). In fact, for binary search trees, we have $n$ random (uniformly distributed) points in an interval, split the interval by the first of these points, and continue recursively splitting each subinterval that contains at least one of the points. If we scale the initial interval to have length $n$, then the probability that a subinterval of length $x$ contains at least one point is $\approx 1 - e^{-x}$.

Let $N(x)$ be the total number of stable fragments when we start with an interval of length $x$ and

$$m(x) = \mathbf{E}(N(x)),\ V(x) = \mathrm{Var}(N(x)).$$

**Moments of $N(x)$**

By the recursive construction of the fragmentation process, we have, almost surely

$$N(x) = \mathbb{1}_{\{x \text{ stable}\}} + \mathbb{1}_{\{x \text{ unstable}\}} \sum_{j=1}^{2} N^{(j)}(\xi_j x), \tag{3.14}$$

where $(N^{(j)}(.))_{j\geq 1}$ are copies of $N(.)$, independent of each other and of $(\xi_1, \xi_2)$. From equation(3.14), we prove that

**Theorem 15** *As $x \to +\infty$ :*

$$m(x) = 2\gamma x + O(1), \tag{3.15}$$

*and*

$$V(x) = 2\lambda x - \frac{2}{x} + O(e^{-x/2}), \tag{3.16}$$

*where*

$$\begin{aligned} \gamma &= \exp\Big[\int_0^1 \frac{2(1-e^{-t})}{t}dt\Big] \exp\Big[-2\int_1^{+\infty} \frac{e^{-t}}{t}dt\Big] \\ &\quad + \int_0^{+\infty} \Big\{\frac{1}{t^2}\Big[e^{-t} - \frac{2(1-e^{-t})}{t}\Big] \exp\Big[-2\int_t^{+\infty} \frac{e^{-s}}{s}ds\Big]\Big\}dt. \end{aligned}$$

*and*

$$\lambda = \exp\Big(\int_0^1 (\frac{1-e^{-t}}{t})dt - \int_1^{+\infty} \frac{e^{-t}}{t}dt\Big) + \int_0^{+\infty} \frac{H(t)}{t^2} \exp\Big(-2\int_t^{+\infty} \frac{e^{-s}}{s}ds\Big)dt.$$

*with*

$$\begin{aligned} H(x) &= 2(1-e^{-x})\mathbf{E}\Big[m(\xi x)m((1-\xi)x)\Big] - \frac{2(1-e^{-x})}{x} \\ &+ (e^{-x} - e^{-2x}) + \frac{(2e^{-x}-1)e^{-x}(m(x)-e^{-x})^2}{2(1-e^{-x})} - 2e^{-x}(m(x)-e^{-x}). \end{aligned}$$

**Limit law of $N_*(x) := \frac{N(x)-m(x)}{\sqrt{V(x)}}$**

For the limit law, we use a general contraction Theorem due to Janson and Neininger [52]. In the sequel we present this method

**General contraction Theorem in continuous time**

Assume that we have

$$Y_t \stackrel{\mathcal{L}}{=} \sum_{r=1}^{K} A_r(t) Y^{(r)}_{T_r^{(t)}} + b_t, \ t \geq 0, \tag{3.17}$$

where $K$ is a positive integer, $(Y_t^{(1)})_t, \cdots, (Y_t^{(K)})_t, (A_1(t), \cdots, A_K(t), b_t, T^{(t)})_t$ are independent, where $T^{(t)} = (T_1^{(t)}, \cdots, T_K^{(t)})$ is a vector of random indices $T_r^{(t)} \in [0, t]$, the $A_r(t)$ are random $d \times d$ matrices for $r = 1, \cdots, K$ and $b_t$ is a random $d$ dimensional vector. Finally in (3.17), we have that for each $t \geq 0$, $Y_t$ and $Y_t^{(r)}$ are identically distributed for all $r = 1, \cdots, K$.

We assume that all $Y_t$ as well as $A_r(t)$, $b_t$ and $T^{(t)}$ are defined on some probability space $(\Omega, \mathcal{F}, \mu)$ and that they are measurable functions of $(t, w)$.

We introduce the normalized random vectors

$$X_t := C_t^{-1/2}(Y_t - M_t),\ t \geq 0, \tag{3.18}$$

where $M(t) = \mathbf{E}(Y_t)$ and $C_t = \mathrm{Cov}(Y_t)$. The recurrence (3.17) implies a recurrence for $X_t$,

$$X_t \stackrel{\mathcal{L}}{=} \sum_{r=1}^{K} A_r^{(t)} X^{(r)}_{T_r^{(t)}} + b^{(t)},\ \ t \geq 0, \tag{3.19}$$

with independence relations as in (3.17) and

$$A_r^{(t)} = C_t^{-1/2} A_r(t) C^{1/2}_{T_r^{(t)}},\ b^{(t)} = C_t^{-1/2}\Big(b_t - M_t + \sum_{r=1}^{K} A_r(t) M_{T_r^{(t)}}\Big).$$

Introduce the map $T$ on the space $\mathcal{M}^d$ of probability measures on $\mathbb{R}^d$ by

$$T:\ \mathcal{M}^d \longrightarrow \mathcal{M}^d, \quad \eta \longrightarrow \mathcal{L}\Big(\sum_{r=1}^{K} A_r^* Z^{(r)} + b^*\Big), \tag{3.20}$$

where $(A_1^*, \cdots, A_K^*, b^*)$, $Z^{(1)}, \cdots, Z^{(K)}$ are some independent random variables and $\mathcal{L}(Z^{(r)}) = \eta$ for $r = 1, \cdots, K$

**Theorem 16 (Janson Neininger (2008))** *Let $0 < s \leq 3$ and let $(Y_t)_{t \geq 0}$ be a process of random vectors satisfying (3.17) such that $\mathbf{E}((Y_t)^s) < \infty$ fo every $t$. Denote by $X_t$ the rescaled quantities in (3.18). Assume that*

$$\mathbf{E}((A_r^{(t)})^s) + \mathbf{E}((b^{(t)})^s) + \sup_{0 \leq u \leq t} \mathbf{E}((X_u)^s) < \infty,\ \forall\, t \geq 0,$$

*and*

$$\begin{aligned} (A_1^{(t)}, \cdots, A_k^{(t)}, b^{(t)}) \quad &\xrightarrow{L^s} (A_1^*, \cdots, A_k^*, b^*), \\ \mathbf{E}\sum_{r=1}^{K} \|A_r^*\|_{op}^s \quad &< 1 \end{aligned}$$

*Then $X_t$ converges in distribution to a limit $X$, where $\mathcal{L}(X)$ is the unique fixed point of $T$ given in (3.20) subject to $\mathbf{E}(X^s) < \infty$.*

## Limit law of $N_*$

Start from the recursive decomposition (3.14), adapted in the form

$$\begin{aligned}\frac{N(x)-m(x)}{\sqrt{V(x)}} &= \frac{1\!\!1_{\{x\ stable\}}-e^{-x}}{\sqrt{V(x)}} \\ &\quad +1\!\!1_{\{x\ instable\}}\frac{N^{(1)}(\xi_1 x)-m(\xi_1 x)}{\sqrt{V(\xi_1 x)}}\sqrt{\frac{V(\xi_1 x)}{V(x)}} \\ &\quad +1\!\!1_{\{x\ instable\}}\frac{N^{(2)}(\xi_2 x)-m(\xi_2 x)}{\sqrt{V(\xi_2 x)}}\sqrt{\frac{V(\xi_2 x)}{V(x)}} \\ &\quad +\frac{\mathbf{E}\Big[N^{(1)}(\xi_1 x)+N^{(2)}(\xi_2 x)\Big]}{\sqrt{V(x)}}\Big[1\!\!1_{\{x\ instable\}}-(1-e^{-x})\Big],\end{aligned}$$

where $\xi$ is a random variable with law $\mathcal{U}_{[0,1]}$. The last equation can be written as

$$N_*(x)=\sum_{j=1}^{2}N_*^{(j)}(\xi_j x)\sqrt{\frac{V(\xi_j x)}{V(x)}}+D(x), \tag{3.21}$$

with

$$\begin{aligned}D(x) &= 1\!\!1_{\{x\ stable\}}\sum_{j=1}^{2}N_*^{(j)}(\xi_j x)\sqrt{\frac{V(\xi_j x)}{V(x)}}+\frac{1\!\!1_{\{x\ stable\}}-e^{-x}}{\sqrt{V(x)}} \\ &\quad +\frac{\mathbf{E}\Big[N^{(1)}(\xi x)+N^{(2)}((1-\xi)x)\Big]}{\sqrt{V(x)}}\Big[1\!\!1_{\{x\ unstable\}}-(1-e^{-x})\Big].\end{aligned}$$

In order to apply Theorem 16, we introduce the map $T$ on the space $\mathcal{M}$ of probability mesures on $\mathbb{R}$ by

$$\begin{aligned}T &: \mathcal{M}\longrightarrow\mathcal{M} \\ T(\mu) &\stackrel{\mathcal{L}}{=} \textstyle\sum_{j=1}^{2}\sqrt{\xi_j}X^{(j)},\end{aligned} \tag{3.22}$$

where $X^{(j)}\stackrel{\mathcal{L}}{=}\mu$ for $j=1,2$. We obtain the following result

**Proposition 6** *As $x\longrightarrow+\infty$, we have*

*(i)* $D(x)\stackrel{L^3}{\longrightarrow}0$.

*(ii)* $\sqrt{\frac{V(\xi_j x)}{V(x)}}\stackrel{L^3}{\longrightarrow}\sqrt{\xi_j}$, $j=1,2$.

*(iii)* $\sum_{j=1}^{2}\mathbf{E}\Big(\sqrt{\xi_j}^{3}\Big)=\frac{4}{5}$.

Since the conditions of Theorem 16 are satisfied one can derive the following result.

**Proposition 7** *$N_*(x)$ converges in distribution to a limit $N_*$, where $\mathcal{L}(N_*)$ is the unique fixed point of $T$ given in (3.22) subject to*

$$\mathbf{E}\Big(N_*^3\Big) < \infty,\ \mathbf{E}(N_*) = 0 \text{ and } \mathrm{Var}(N_*) = 1.$$

It's not difficult to prove that the law $\mathcal{N}(0,1)$ is a fixed point of the map $T$, leading to our last Theorem.

**Theorem 17** *As $x \longrightarrow +\infty$, we have a version of the central limit Theorem*

$$N_*(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

### 3.3.4 A project work

1. We intend to extend this work in the case when $m \geq 2$ and $p(s) = 1 - e^{-s}$.
2. By our assumptions, the label of a node equals the sum of the labels of its children. Another version would be to allow a (possibly random) loss at each node. One important case is Rényi' s parking problem [80], where a node with label $x$ is interpreted as an interval of length $x$ on a street, where cars of length 1 park at random. Each car splits an interval of length $x \geq 1$ into two free intervals with the lengths $\xi(x-1)$ and $(1-\xi)(x-1)$, where $\xi \sim U(0,1)$. An obvious generalization is to split $(x-1)$ using an arbitrary random vector $(\xi_1, ..., \xi_m)$.

## 3.4 A pattern matching problem

### 3.4.1 Introduction

Let $\omega$ and $\omega'$ be two words of length $n$ (the letters $\omega_i$ and $\omega'_i$ of these words are i.i.d., i.e., chosen according to any previously fixed distribution). One has then a probability $p$ that $\omega_i = \omega'_i$ . What is the length of the longest common sequence ? I.e., what is the largest $l$ such that $\omega_{i+1} = \omega'_{i+1}; \cdots; \omega_{i+l} = \omega'_{i+l}$ ? (Note that we considerer here subsequences beginning at the same position ; in biology, this model makes sense e.g. when we have a synchronisation forced by a "start" codon). Another natural question is what is the length of the last common suffix ? I.e., what is the largest $s$ such that $\omega_{n+1-s} = \omega'_{n+1-s}, \cdots, \omega_n = \omega'_n$ ? More generally, given a reference pattern $\omega$ and $m$ samplings $\omega'; \omega"; \cdots$, what is the length of the longest common sequence between $\omega$ and any one of the $\omega', \omega", \cdots$ ?

For example, consider $m = 2$ ; $n = 24$, and

$$\begin{aligned} \omega &= abceacbcddacbccacabbaaaa, \\ \omega' &= abcebcbcddacbccbcabbacaa, \\ \omega" &= abceccbcddacaccccacbaaaa. \end{aligned}$$

The length of largest last common sequence is $H_n = 9$ (because of the common pattern *bcddacbcc* between $\omega$ and $\omega'$). The length of largest last common sequence is $L_n = 5$ (because of the common end *baaaa* between $\omega$ and $\omega$").

As illustrated by D. Knuth [55], knowing the typical behavior of these quantities allows to tune faster algorithms. It is also possible to use such knowledges to design more precise approximation algorithms, indeed it could be the case that solving a given problem exactly is exponential, and therefore it is wise to use early abort strategies, relying on the knowledge of the typical behavior. This idea is used in computational biology for RNA alignment questions (e.g. in softwares like BLAST).

### 3.4.2 Some examples

**Example 1 :** Air France has $m$ airplanes. Assume that the mean time between failures follows an exponential law. Each week, *one* of the airplanes is immobilized and the speed detector follows a complete checkup. With a probability $p$, it is replaced by a new speed detector.
NB1 : Each day, one has then 0 or 1 replaced speed detector.
NB2 : When $p = 1$, this model is the classical "Moran Model".

This example can be generalized as follows :
We suppose that each day, *each* component has a probability $p$ to have a failure. All the broken components are then replaced by new ones.
NB : Each day, one has from 0 to $m$ broken components. The independence property allows to reexpress everything in terms of $m = 1$, and, in this case, the models with or without independence are equivalent. Therefore, we concentrate hereafter on the case $m = 1$ of the Moran Model.
**Example 2 (Formulation as a random walk problem) :** Let $Y_n^{(1)}, Y_n^{(2)}, \cdots, Y_n^{(m)}$ be the age of the components 1 to $m$ at day $n$. One has then $Y_n = \max(Y_n^{(1)}, Y_n^{(2)}, \cdots, Y_n^{(m)})$ and $H_n = \max(Y_0, Y_1, Y_2, \cdots, Y_n)$. For such a walk of length $n$, $Y_n$ is the final altitude (at time $n$) and $H_n$ is the height (the age of the oldest component) between time 0 and $n$.

### 3.4.3 Generating function of the generalized Moran model

We consider here the case of $m = 1$ component. This case directly translates into the following recurrence :

$$\begin{aligned} Y_0 &= 0 \\ Y_{n+1} &= \begin{cases} 1 + Y_n, & \text{with probability } q = 1 - p\,; \\ 0, & \text{with probability } p. \end{cases} \end{aligned}$$

We introduce the bivariate generating function

$$F(z,u) = \sum_{n\geq 0} f_n(u) z^n = \sum_{n\geq 0} \mathbf{E}[u^{Y_n}] z^n = \sum_{n\geq 0} \Big( \sum_{k=0}^{n} \mathbf{P}(Y_n = k) u^k \Big) z^n \sum_{k\geq 0} F_k(z) u^k$$

After some manipulation, one gets

$$F(z,u) = \frac{1+\frac{pz}{1-z}}{1-quz}.$$

From this closed form, it is standard to get the mean and the standard deviation :

**Theorem 18 (Length of the last common suffix)** .

$$\begin{aligned} \mathbf{E}(Y_n) &= \frac{q^{n+1}-q}{q-1} \\ \mathrm{Var}(Y_n) &= q^n(q-3q^2-2qn+2q^2n-q^{2+n})+q-2q^2/p. \end{aligned}$$

In order to study $H_n$, we first introduce walks of height bounded by $h$ (from a probabilistic point of view, one conditions the walk not to go above the height $h$, please note that this does not modify the transition probability : one does not have a reflecting border, but just a conditioning). The following equalities implicitly defines all our objects :

$$F^{(\leq h)}(z,u) = \sum_{n\geq 0} f_n^{(\leq h)}(u) z^n = \sum_{k\geq 0} F_k^{(\leq h)}(z) u^k.$$

After tedious algebra, we get

$$F^{(\leq h)}(z,u) = \frac{(1-qz)(1-(quz)^{h+1})}{(1-quz)(1-z+(qz)^{h+1}zp)}.$$

In fact, if we care only for $H_n$, $u$ is just a catalytic variable : we can remove it by setting $u=1$, though it would have not been possible to get the following generating function without introducing $u$. The variable $u$ is however useful if we want to study the random variable $Y_n$ (length of the last common suffix).

**Theorem 19 (closed form solution for $H_n$)** *The probability that the longest common pattern (between two texts of length $n$) of size $\leq h$ is :*

$$\begin{aligned} \mathbf{P}(H_n \leq h) &= [z^n]F^{(\leq h)}(z,1) \\ \textit{where} & \\ F^{(\leq h)}(z,1) &= \frac{1-(qz)^{h+1}}{1-z+p/q(zq)^{h+2}} \end{aligned}$$

*For any fixed $p$, $h$, and $n$, it is possible to get the exact value of $\mathbf{P}[H_n \leq h]$ in time $O(ln(n))$ (via binary exponentiation).*

Let $F^{(=h)}(z)$ be the generating function of walks reaching at least one height $h$, but never exceeding it and let $F^{(>h)}(z)$ the generating function of walks reaching at least one height exceeds $h$. The average height is then given by

$$\mathbf{E}[H_n] = [z^n] \sum_{k\geq 0} hF^{(=h)}(z)$$

or, equivalently, by

$$\mathbf{E}[H_n] = [z^n] \sum_{h \geq 0} F^{(>h)}(z).$$

It is then possible to give a closed form formula involving a double sum and binomial coefficients (this is related to integer compositions into l parts), but we do not go to this direction here, as such a double sum will be of no use for getting the asymptotics of $H_n$, on which we concentrate now.

**Theorem 20 (Asymptotics of $H_n$)** *The average length of the longest common pattern is*

$$\mathbf{E}(H_n) \sim \frac{\ln n}{\ln(1/q)}.$$

### 3.4.4 Some future works

It is possible to say much more on these models.

- What can be said about the deviation of $H_n$ ?, about the asymptotic distribution of $H_n^* := \frac{H_n - \mathbf{E}[H_n]}{\sqrt{\mathrm{Var}[H_n]}}$ ?.
- It could also be possible to extend these results to text according to a distribution following a higher order Markovian dependency, or to patterns which are a regular expression.
- We hope to find many other new problems/algorithms for which these maximal statistics (and the fine asymptotics on them that we can derive with our approach) are relevant.

# Bibliographie


[1] Aguech R., Lasmar N., and Mahmoud H. (2006). *Distances in random digital search trees.* Acta Informatica, **43**, 243-264.

[2] Aguech R., Lasmar, N., and Mahmoud H. (2006). *Limit distribution of distances in biased random tries.* Journal of Applied Probability, **43** , 1-14.

[3] Aguech R. (2009) *Limit Theorems for random triangular urn schemes.* Journal of Applied Probability, **46** , 827-843.

[4] Aguech R., Lasmar N. and Mahmoud H.(2006) *Extremal Weighted Path Lengths in Random Binary Search Trees.* Probability in the Engineering and Informational Sciences, **21**, 133-141.

[5] Aguech R., Louhichi S. and Louhichi S. (2005) *On the asymptotic independent representation for sums of some weakly dependent random variables.* Studia Scientiarum Mathematicarum Hungaruca, **42 (4)**, 371-383.

[6] Aguech R. (2008) *The size of random fragmentation intervals. DMTCS proc AI*, 523-534

[7] Aldous D., Flannery B. and Palacios J. (1988). *Two applications of urn processes : The fringe analysis of search trees and the simulation of quesi-stationary distributions of Markov chains.* Probability in Engineering and information sciences. **2**, 293-307.

[8] Athreya K.B., and Karlin S. (1967). *Limit Theorems for the split times of branching processes.* J.Math.Mech.,**17**, 257-277.

[9] Athreya K.B., and Karlin S. (1968). *Embedding of urn schemes into continous time Markov branching processes and related limit Theorems. Ann. Math. Statist.*, **39**, 1801-1817.

[10] Athreya K.B., and Ney P.E . (1972). *Branching Processes.* Springer, Berlin.

[11] Bagchi A. and Pal A.K. (1985). *Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures.* SIAM Journal on Algebraic and Discrete Methods **6**, 394-405.

[12] Banderier C. (2001). *Combinatoire analytique des chemins et des cartes.* Ph.D. thesis, Université Paris VI.

[13] Banderier C. and Flajolet P. (2002). *Basic analytic combinatorics of directed lattice paths.* Theor. Comput. Sci. **281** 37-80

[14] Bernstein S. (1940). *Sur un problème du shèma des urnes à composition variable.* C.R. (Dokl.), Acad. Sci. URSS **28**, 5-7.

[15] Billingsley P. (1995). *Probability and Measure*. Wiley, New York.

[16] Campi X. , Krivine H. , Sator N. and Plagnol E. (2000). *Analyzing fragmentation of simple fluids with percolation theory.* Eur. Phys. J. D., **11, 233**.

[17] Chern H., Hwang H. and Tsai T. (2002). *An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms.* Journal of Algorithms, **44**, 177–225.

[18] Christophi C. and Mahmoud H. (2005). *The oscillatory distribution of distances in random tries.* The Annals of Applied Probability, **15**, 1536-1564.

[19] Coffman E. and Eve J. (1970). *File structures using hashing functions.* Communications of the ACM, **13**, 427–432, and 436.

[20] De La Briandais R. (1959). *File searching using variable length keys.* Proceedings of the Western Joint Computer Conference, 295–298, AFIPS, San Francisco, California.

[21] Derrida B. and Flyvbjerg H. (1987). *Distribution of local magnetisations in random networks of automata.* J. Phys. A , **20**, 5273 .

[22] Devroye, L. (1986). *A note on the height of binary search trees.* journal of the ACM, **33**, 489–498.

[23] Devroye L. (1987). *Branching processes in the analysis of the height of trees.* Acta Informatica, **24**, 277–298.

[24] Devroye L. (1988). *Applications of the theory of records in the study of random trees.* Acta Informatica, **26**, 123–130.

[25] Devroye L. and Neininger, R. (2004). *Distances and finger search in random binary search trees.* SIAM Journal on Computing, **33**, 647–658.

[26] Drmota. M. (2001). *An analytic approach to the height of binary search trees.* Algorithmica, **29**, 89–119.

[27] Drmota. M. (2002). *The variance of the height of binary search trees.* Theoretical Computer Science, **270**, 913–919.

[28] Doukhan P. and Sana L. (1997). *Weak dependence and moment inequalities.* Pre-print Université de Paris-Sud,

[29] Eggenberger F. and Pòlya G. (1923). *U¨ber die statistik verketetter vorgäge. Z,.* schrift für Angewandte Mathematik and Mechanik, **1**, 279-289.

[30] Esary, J., Proschan, F. and Walkup, D. (1967) *Association of random variables with applications.* Ann. Math. Statist, **38**, 1466-1474. MR 36, 915

[31] Eweda E. and Macchi O. (1985). *Tracking error bounds of adaptive nonstationary filtering.* Automatica, **2(3)**, 293-302.

[32] Fagin R., Nievergelt J., PippengerN., and Strong H.(1979). *Extendible hashing—a fast access method for dynamic files.* ACM Transactions on Database Systems, **4**, 315–344.

[33] Flajolet P., Gourdon X. and Dumas P. (1995). *Mellin transform and asymptotic harmonic sums.* Theoretical Computer Science, **144**, 3–58.

[34] Flajolet P., Joaquim K. and Pekari H. (2005) *Analytic Urns.* The Annals of Probability , **33** no. 3, 1200-1233.

[35] Flajolet P. and Vincent P. (2005) *Triangular Urns with dimension 2 and 3.* Phd Thesis from the Ecole Polytechnique.

[36] Flajolet P., Nicodème P. and Salvy B. (2002) *Motif statistics.* Theore. Comput Sci., **287 (2)** 593-618.

[37] Flajolet P. and Sedgewick R. (2008) *Analytic Combinatorics.* Cambridge university Press 824p, **287 (2)** 593-618.

[38] Fredkin E. (1960). *Trie memory.* Communications of the ACM, **3**, 490–499.

[39] Friedman B. (1949). *A simple urn model.* Communications of Pure and Applied Mathematics, **2**, 59-70.

[40] Gouet R. (1993). *Martingale functional central limit theorems for a generalized Pólya-urn.* The Annals of Probability, **21**, 1624-1639.

[41] Guo L. and Ljung L. (1995). *Exponential stability of general tracking algorithms.* IEEE Trans. on automatic Contro, **40 no 8** 1376-1387

[42] Guo L., Ljung L. and Wang G.-J.(1997). *Necessary and sufficient cnditions for stability of LMS.* IEEE Trans. on Automatic Control, **40**, 1376-1387.

[43] Guo L.(1994). *Stability of recursive stochastic tracking algorithms.* SIAM journal of control and Optimisation, **23**, 1195-1225.

[44] Harris T.E (1951). *Some mathematical models for branching Processes.* 2nd Berk. Symp. 305-328.

[45] Hwang, H. and Tsai, T. (2002). *Quickselect and Dickman function.* Combinatorics, Probability and Computing, **11**, 353–371.

[46] Itoh Y. and Mahmoud H. (2005). *Age statistics in the Moran population model.* Stat. and Prob. Lett. 305-328. **74(1)** 21-30

[47] Itoh Y., Mahmoud H. and Daisuke Takahashi. (2004). *A stochastic model for solitons.* Random Structures and Algorithms, **24(1)** 51-64

[48] Jacquet P. and Szpankowski W. (1998). *Analytical depoissonization and its applications.* Theoretical of Computer Science, **201**, 1–62.

[49] Jakubowski A.(1993). *An asymptotic independent representation in the limit theorems for maxima of nonstationary rndom sequences.* Ann. Probab, **21** No 2, 819-830

[50] Janson S. (2004). *Functional limit theorems for multitype branching processes and generalized Pólya urns.* Stochastic Process. Appl., **110** no 2, 909–930.

[51] Janson S. (2005). *Limit theorems for triangular urn schemes.* Probability Theory and Related Fields, **134**, 417–452.

[52] Janson S. and Neininger R.(2008). *The size of random fragmentation trees.* Probab. Theory Relat. Fields, **142**, no. 3-4, 399-442.

[53] Karlin S. and McGregor J. (1968). *Embeddability of discrete time simple branching processes into contunious time branching processes.* TAMS, **132** 115-136.

[54] Kemp R. (1984). *Fundamentals of the Average Case Analysis of Particular Algorithms.* Wiley-Teubner Series in Computer Science, John Wiley & Sons, New York.

[55] Knuth D. (1998). *The Art of Computer Programming*, Vol. 3 : *Sorting and Searching*, 2nd ed. Addison-Wesley, Reading, Massachusetts.

[56] Kotz S., Mahmoud H. and Robert P. (2000). *On generalized Pólya urn models.* Statistic probability letters, **49** no. 2, 163-173.

[57] Krapivsky, P.L., Grosse, I. and Ben-Naim, E. (2004). *Stable distributions in stochastic fragmentation.* J.Phys. A : Math. Gen., **37**, 2863-2880.

[58] Lawn B.R. and Wilshaw T.R. (1975). *Fracture of Brittle Solids.* Cambridge University Press, Cambridge.

[59] Louchard G. and Szpankowski W. (1995). *Average profile and limiting distribution for a phrase size in the Lempel-Ziv parsing algorithm.* IEEE Transactions on Information Theory, **41**, 478–488.

[60] Mahmoud H. (1992). *Evolution of Random Search Trees*. Wiley, New York.

[61] Mahmoud, H. and Pittel, B. (1984). *On the most probable shape of a search tree grown from a random permutation.* SIAM Journal on Algebraic and Discrete Methods, **5**, 69–81.

[62] Mahmoud H., Modarres R. and Smythe R. (1995). *Analysis of quickselect : An algorithm for order statistics.* RAIRO : Theoretical Informatics and Its Applications, **29**, 255–276.

[63] Mahmoud H. (2000). *Sorting : A Distribution Theory*. Wiley, New York.

[64] Mahmoud M. and Neininger R. (2003). *Distribution of distances in random binary search trees.* The Annals of Applied Probability, **13**, 253–276.

[65] Mathys P. and Flajolet P. (1985). *Q-ary collision resolution algorithms in random-access systems with free and blocked channel access.* IEEE Transactions on Information Theory, **31**, 217–243.

[66] Neininger R. (2001). *On a multivariate contraction method for random recursive structures with applications to Quicksort.* Random Structures Algorithms, **19**, 498–524.

[67] Neininger, R. (2002). *The Wiener index of random trees.*Combinatorics, Probability and Computing, **11**, 587–597.

[68] Neininger, R. and Rüschendorf, L. (2004). *A general limit theorem for recursive algorithms and combinatorial structures.* Ann. Appl. Probab, **14**, 378–418 .

[69] Newman C.M. (1984). *Asymptotic independence and limit theorems for positively nd negatively dependent random variables.* Y.L. Tong, editor, inequalities in statistics and Probability, IMA Lecture Notes-Monograph Series, **5**, 127-140.

[70] Nicodème P. (2001) *Fast approximate motif statistics.* Journal of Computational Biology, **8(3)** 235-248

[71] Panholzer A. and Prodinger H. (2004). *Spanning tree size in random binary search trees.* The Annals of Applied Probability, **14**, 718–733.

[72] Pittel, B. (1984). *On growing random binary trees.* Journal of Mathematical Analysis and its Applications, **103**, 461–480.

[73] Poblete, P., Papadakis, T. and Munro, I. (1995). *The binomial transform and its applications to the analysis of skip lists.* Proc. ESA 95, Lecture Notes in Computer Science, **979**, 554-569. Springer-Verlag, New York.

[74] Pólya G. (1931). *Sur quelques points de la théorie des probabilités.* Annales de l'institut Henri Poincaré, **1**, 117-161.

[75] Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models.* Wiely, New York.

[76] Rachev, S. and Rüschendorf, L. (1995). *Probability metrics and recursive algorithms.* Advances in Applied Probability, **27**, 770–799.

[77] Reed, B. (2003). *The height of a random binary search tree.* Journal of the Association for Computing Machinery, **50**, 306–332.

[78] RednerFor S. (1990). *Statistical Models for the Fracture of Disordered Media.* ed. H.J. Herrmann and S. Roux (Elsevier Science, New York).

[79] Robson, J. (1979). *The height of binary search trees.* The Australian Computer Journal, **11**, 151–153.

[80] Rényi A. (1958). *On a one-dimensional random space-filling problem.*(ungarian.) Magyar Tud. Akad. Mat. Kutató Int. Kozl. English transl, **3**, 109-127. in Selected papers of Alfréd Rényi, Vol. II : 1956-1961. Ed. Pál Turán. Akad´emiai Kiadó, Budapest, 1976, pp. 173-188.

[81] Rösler, U. (1991). *A limit theorem for "Quicksort".* RAIRO Inform. Théor. Appl.,**25**, 85–100.

[82] Rösler, U. (2001). *On the analysis of stochastic divide and conquer algorithms.* Algorithmica, **29**, 238–261.

[83] Rösler, U. and Rüschendorf, L. (2001). *The contraction method for recursive algorithms.* Algorithmica, **29**, 3–33.

[84] Salvy B. and Zimmermann P. (1994) *Gfun : a maple package for manipulation of generating and holonomic functions in one variable.* Transactions of Mathematical Software, **20(2)** 163-177.

[85] Savkevich V. (1940). *Sur le schéma des urnes à composition variable.* C.R (Dokl.) Acad.Sci. URSS, **28**, 8-12.

[86] Sibuya M. and Itoh Y. (1987). *Random sequential bisection and its associated binary tree.* Ann. Inst. Satist. Math., **39**, 69–84.

[87] Solo V. and Kong X. (1995). *Adaptive signal processing algorithms : stability and performance.* Prentice Hall,

[88] Szpankowski W. (2001). *Average Case Analysis of Algorithms on Sequences*. Wiley, New York.

[89] Zolotarev, V.M. (1977). *Ideal metrics in the problem of approximating the distributions of sums of independent random variables (Russian).* Teor. Verojatnost. i Primenen, **22**, 449–465.