m-ary Search Trees when m ≥ 27: A Strong Asymptotics for the Space Requirements

Brigitte Chauvin, Nicolas Pouyanne
Département de Mathématiques, Université de Versailles-Saint-Quentin, 45, avenue des Etats-Unis, 78035 Versailles cedex, France; e-mail: chauvin@math.uvsq.fr, pouyanne@math.uvsq.fr

Received 4 July 2002; revised 28 May 2003; accepted 5 September 2003

DOI 10.1002/rsa.10108

ABSTRACT: It is known that the joint distribution of the number of nodes of each type of an m-ary search tree is asymptotically multivariate normal when $m \leq 26$. When $m \geq 27$, we show the following strong asymptotics of the random vector $X_n = {}^t(X_n^{(1)}, \ldots, X_n^{(m-1)})$, where $X_n^{(i)}$ denotes the number of nodes containing $i - 1$ keys after having introduced $n - 1$ keys in the tree: There exist (nonrandom) vectors X, C, and S and random variables ρ and φ such that $(X_n - nX)/n^{\sigma_2} - \rho(C \cos(\tau_2 \log n + \varphi) + S \sin(\tau_2 \log n + \varphi)) \to_{n\to\infty} 0$ almost surely and in L^2; σ_2 and τ_2 denote the real and imaginary parts of one of the eigenvalues of the transition matrix, having the second greatest real part. © 2004 Wiley Periodicals, Inc. Random Struct. Alg., 24: 133–154, 2004

1. INTRODUCTION

An m-ary search tree is a data structure that grows by the progressive insertion of keys into a tree with branch factor m (first sentence in Lew and Mahmoud's paper [8]). Each node of such a tree contains 0, 1, $\cdots$ or $m - 1$ keys and gives rise to m branches (see Section 2 for the detailed definition of an m-ary search tree).

Our purpose is to make precise the asymptotic behavior of the vector X_n whose

---

Correspondence to: B. Chauvin

coordinates are (with our notations) the number of nodes containing $0, \ldots, m-2$ keys in a random m-ary search tree holding $n-1$ keys, as n tends to infinity.

Several cost measures on random m-ary search trees have been studied in the literature, one of the most frequent being the total number of nodes S_n also called space requirement; in the vectorial frame, we note that S_n is an affine function of X_n. This cost is classically studied using generating functions and the method of moments. Mahmoud and Pittel [11] describe the asymptotics of the mean and the variance of S_n and derive a normal limit distribution for $m \leq 15$. Lew and Mahmoud [8] extend this range to $m \leq 26$. Smythe [16] and Mahmoud and Smythe [12] conjecture that the limit distribution is not normal for $m > 26$. In the related frame of branching processes, the change of normal limit laws to nonnormal ones depends on the second eigenvalue of the transition matrix (which corresponds for m-ary search trees to the transition at $m = 26$) and already appears for instance in Athreya and Ney's book [1]. This has been often noted by the previous authors dealing with m-ary search trees (see, for instance, example 3.1 in Smythe [16]).

The state-of-the-art can be found in Chern and Hwang's paper [3]: A phase transition occurs between $m = 26$ and $m = 27$. It is easy to see the following from the variance of the space requirement, where the asymptotics have two types of behavior, depending on the values of m: For small m ($m \leq 26$), the variance is of order n and the rescaled space requirement is asymptotically normal, but, for $m \geq 27$, the variance is of order $n^{2\sigma}$ for some (known) real number σ, $\sigma > 1/2$, and a periodic phenomenon appears.

In the range $m \geq 27$, the challenge comes from the questions asked by Chern and Hwang, who prove (in [3], Corollary 2) that the distribution of S_n, even conveniently renormalized, does not approach any fixed distribution function but fluctuates via some periodic function. They seek more intuitive explanations of the phase transition than pure analytic reasons.

The asymptotic normality for $m \leq 26$ can also be found by contraction method (see, for instance, Neininger and Rueschendorf [14]). Interestingly, the same phase transition for the variance is noted by physicists in the close context of random fragmentation problem (for instance, in Dean and Majumdar [4]).

The literature on the subject, including limit distribution results by the contraction method, mostly takes advantage of the "divide-and-conquer" recursivity (sometimes called the "backward" method). Another point of view on these processes is based on the dynamical recursivity (sometimes called the "forward" method), already used in Smythe's [16] and Mahmoud and Smythe's [12] papers.

We consider $(X_n)_{n\geq 1}$ a Markov process, and we notice that X_n is a kind of Pólya urn model, random walk or a multitype branching process, depending on one's point of view.

In Sections 2 and 3, we see how $(X_n)_{n\geq 1}$ can be viewed as a Markovian process with values in $\mathbb{R}^{m-1}$ and that its evolution is driven by a transition-type matrix A in the following remarkable (*linear*) way:

$$E^{\mathcal{F}_n}(X_{n+1}) = \left(\mathrm{Id} + \frac{A}{n}\right)X_n, \tag{1}$$

where $\mathcal{F}_n$ is the past before time n and Id is the identity matrix. Our method is based on exploiting the linearity of this evolution.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Thus $\mathbb{C}^{m-1}$ is decomposed along the eigenspaces of A, and, if $\mathrm{Sp}(A)$ denotes the set of eigenvalues of A (all its eigenvalues are simple), we have

$$\mathbb{C}^{m-1} = \bigoplus_{\lambda \in \mathrm{Sp}(A)} ker(A - \lambda\,\mathrm{Id}),$$

$$\mathrm{Id} = \sum_{\lambda \in \mathrm{Sp}(A)} \pi_\lambda$$

$$A = \sum_{\lambda \in \mathrm{Sp}(A)} \lambda \pi_\lambda \tag{2}$$

where π_λ denotes the projection on the eigenspace $ker(A - \lambda\,\mathrm{Id})$ relative to the decomposition (2). Moreover, 1 is an eigenvalue, the other ones having a real part strictly less than 1. If λ_2 and $\overline{\lambda_2}$ are the eigenvalues having the greatest real part, say σ_2, $\sigma_2 < 1$, we write the following fundamental decomposition of vector X_n ($\bar{x}$ denotes the conjugate of a complex number x):

$$X_n = \pi_1 X_n + \pi_{\lambda_2} X_n + \pi_{\overline{\lambda_2}} X_n + \sum_{\lambda \neq 1, \lambda_2, \overline{\lambda_2}} \pi_\lambda X_n. \tag{3}$$

This spectral decomposition of X_n coincides, as $m \geq 27$, with the almost sure asymptotic expansion of X_n for the first three terms; this is a key phenomenon. For this purpose, the analysis of each projection $\pi_\lambda X_n$ is performed by rescaling it in order to get a martingale. Notice that the appearance of martingale methods is not surprising, considering the evolution given by formula (1). The result then comes from the spectral decomposition (3) and from the lemmas in Section 4 explaining successively that the first projection is of order n, the projections $\pi_\lambda X_n$ for $\Re(\lambda) > 1/2$ are of order n^λ by an L^2-convergence theorem of martingales, and the remaining projections $\pi_\lambda X_n$ for $\Re(\lambda) \leq 1/2$ are asymptotically almost surely negligible. One can find the complete theorem with its proof in Section 5. Simulations in Section 6 help to visualize the phenomena.

Notice that our approach is somehow complementary to Mahmoud's one in a recent paper [10], where the frame (Pólya schemes) is quite large, including m-ary search trees, and focuses on the leading term of X_n; our study goes further in the expansion of X_n but is restricted here to m-ary search trees.

Using similar arguments, we hope that the asymptotics of the "profile" (meaning the number of nodes level by level in the tree) of an m-ary search tree is tractable: A natural generalization of the binary search tree case [6] to higher dimensions would consist of considering the number of nodes of each type level by level, and introducing some "level polynomial" vectors. This will be the subject of a forthcoming paper.

2. DEFINITION AND MARKOVIANITY OF THE PROCESS

One throws a sequence of numbers in [0, 1], named the *keys,* uniformly in $[0, 1]^{\mathbb{N}^*}$. The keys are placed one after another in an m-ary tree (one node-root, from each node grow m branches). The following recursive rule describes the way a key named k is inserted in the tree.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

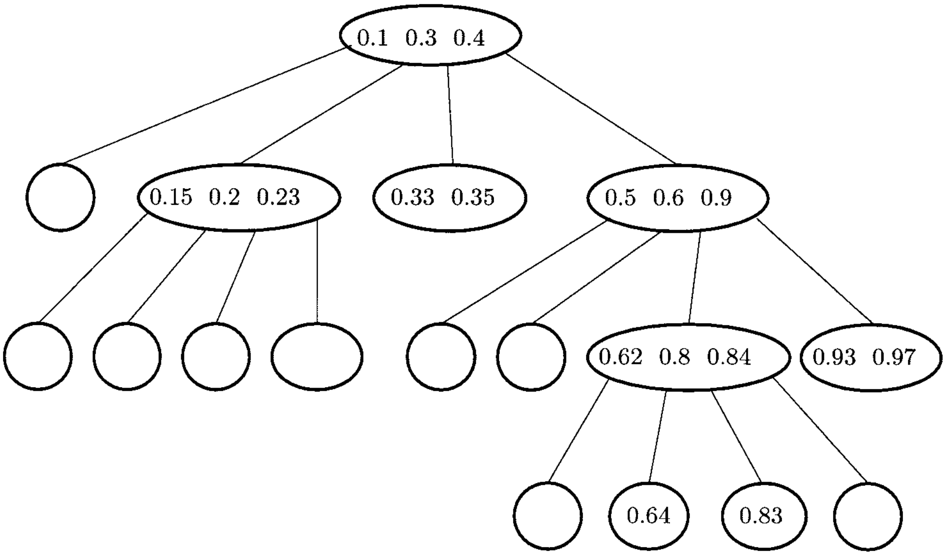

Fig. 1. Insertion of the keys 0.3, 0.1, 0.4, 0.15, 0.9, 0.2, 0.6, 0.5, 0.35, 0.8, 0.97, 0.93, 0.23, 0.84, 0.62, 0.64, 0.33, 0.83 in a 4-ary tree.

i. If the root contains strictly less than $m - 1$ keys, then k is inserted in the root. One draws usually keys in a root from left to right in increasing order.

ii. If the root is already saturated, i.e., if it contains $m - 1$ keys named $k_1, \ldots, k_{m-1}$, ordered such that $k_i < k_{i+1}$, then corresponding to each interval $I_1 =]-\infty, k_1[$, $I_{j+1} =]k_j, k_{j+1}[$ $(1 \leq j \leq m-2)$, $I_m =]k_{m-1}, +\infty[$ a subtree, itself an m-ary search tree. One draws usually the branches corresponding to $I_1, \ldots, I_m$ from left to right. In this situation, k in inserted in the subtree that corresponds to the interval I_j such that $k \in I_j$.[1]

Figure 1 is an example of 4-ary tree obtained by insertion of the numbers 0.3, 0.1, 0.4, 0.15, 0.9, 0.2, 0.6, 0.5, 0.35, 0.8, 0.97, 0.93, 0.23, 0.84, 0.62, 0.64, 0.33, 0.83 in this order.

Although it is not explicitly used later on, let us mention (see Mahmoud's book [9] for details) that such a sequence $(T_n)_{n\in\mathbb{N}}$ of trees has the same distribution as the one obtained by construction of T_n from a random permutation of n integers, with a uniform distribution on the set of permutations. This is the so-called random permutation model.

In the sequel, $(T_n)_{n\in\mathbb{N}}$ and other parameters of interest are random variables on the space Ω of infinite m-ary trees.[2] The space is endowed with the natural filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, where $\mathcal{F}_n$ is the σ-field generated up to time n.

For each $i = \{1, \ldots, m\}$ and $n \geq 1$, we define the number $X_n^{(i)}$ as the number of nodes

[1]In this paper, our convention is that empty nodes (corresponding to the m above intervals) appear when the concerned internal node has just been saturated by the insertion of an $(m-1)$st key. Other conventions are possible; for instance, empty nodes could appear once the first key is stored in the concerned internal node. Anyway, the choice of any convention has no impact on the results.

[2]It is necessary to define random variables on this (big) probability space in order to give a meaning to almost sure and L^2-convergences later on. For tree probability spaces, see, for instance, Neveu [15].

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that contain $i - 1$ keys after insertion of the $(n - 1)$st key; such nodes are named nodes of type i. The question consists of describing the asymptotic behavior of the $X_n^{(i)}$ as n tends to infinity.

Counting the total number of keys in nodes of each type in the tree holding $(n - 1)$ keys leads to the formula

$$n - 1 = \sum_{i=1}^{m} (i - 1)X_n^{(i)}. \tag{4}$$

This formula which binds the $X_n^{(i)}$ allows to limit the study to the first $m - 1$ indices i. It should not be confused with the relation (8) later on which counts the number of free places (or gaps) in the tree.

Throughout, we call

$$V = \mathbb{R}^{m-1}$$

(or more exactly the real vector space of matrices having one column and $m - 1$ rows). The random vector $X_n \in V$ is defined for all $n \geq 1$ as

$$X_n = \begin{pmatrix} X_n^{(1)} \\ \vdots \\ X_n^{(m-1)} \end{pmatrix},$$

and evolves as follows. The first $m + 1$ vectors are nonrandom:

$$\begin{cases} X_1 = {}^t(1, 0, \ldots, 0), \\ X_2 = {}^t(0, 1, \ldots, 0), \\ \vdots \\ X_{m-1} = {}^t(0, \ldots, 0, 1), \\ X_m = {}^t(m, 0, \ldots, 0), \\ X_{m+1} = {}^t(m - 1, 1, 0, \cdots). \end{cases}$$

The following ones are random. For instance, $X_{m+2} = {}^t(m - 2, 2, 0, \cdots)$ with probability $(m - 1)/(m + 1)$, and $X_{m+2} = {}^t(m - 1, 0, 1, 0, \cdots)$ with probability $2/(m + 1)$. These probabilities are computed with the rules of the random permutation model: When $n - 1$ keys are inserted, the probability that the nth one falls between two of them is $1/n$ (the probability that it falls on the left-hand side of the smallest one or on the right-hand side of the greatest one is $1/n$, too). Consequently, only the relative order of the keys is taken into account (not their values).

More generally, the transition rules between the states at time n and $n + 1$ are the following: For each i between 1 and $m - 1$, if the nth key falls on a node of type i, then

$$X_{n+1} = X_n + \Delta_i,$$

where

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\begin{cases} \Delta_1 = {}^t(-1, 1, 0, 0, \cdots), \\ \Delta_2 = {}^t(0, -1, 1, 0, \cdots), \\ \vdots \\ \Delta_{m-2} = {}^t(0, \ldots, 0, -1, 1), \\ \Delta_{m-1} = {}^t(m, 0, \ldots, 0, -1), \end{cases}$$

and this event takes place with probability $(i/n)X_n^{(i)}$ because each node of type i contains i free places.

Let us emphasize here that this last probability, containing the randomness of the evolution of the process, is *linear* in X_n. For this reason, for each $i \in \{1, \ldots, m-1\}$, let l_i be the *linear* form of V defined as

$$l_i = idx_i,$$

where dx_i is the ith coordinate form of $V = \mathbb{R}^{m-1}$. The process $(X_n)_n$ in V is now defined by the first vector X_1 and the transition condition for each $n \geq 1$:

$$\begin{cases} X_{n+1} = X_n + \Delta_1, & \text{with probability } (1/n)l_1(X_n), \\ \vdots \\ X_{n+1} = X_n + \Delta_{m-1}, & \text{with probability } (1/n)l_{m-1}(X_n). \end{cases} \tag{5}$$

In other words, the process is a random walk in V defined by X_1 and a random increment $\Delta(n+1)$ between times n and $n+1$:

$$X_{n+1} = X_n + \Delta(n+1), \tag{6}$$

with the transition probabilities

$$\mathrm{P}(\Delta(n+1) = \Delta_i | X_n) = \frac{1}{n} l_i(X_n), \qquad 1 \leq i \leq m-1. \tag{7}$$

Note that the process (X_n) satisfies the relation

$$\sum_{i=1}^{m-1} iX_n^{(i)} = \sum_{i=1}^{m-1} l_i(X_n) = n, \tag{8}$$

available for each $n \geq 1$, meaning that the numbers $l_i(X_n)/n$ are probabilities of disjoint events whose union is the total probability space. The interpretation of this relation in terms of m-ary trees is just the distribution of the n free places where the nth key may be inserted into nodes of different types (each node of type i contains i free places). This relation plays a crucial role in the theorem.

Other relations are satisfied by the l_i and Δ_i, namely,

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\forall\, i \in \{1, \ldots, m-1\}, \qquad \sum_{j=1}^{m-1} l_j(\Delta_i) = 1. \tag{9}$$

Remarks. (i) As noticed in [10], X_n also describes the composition of a Pólya urn model, where the $m-1$ colors are the $m-1$ types of the nodes and where the balls are the free places. The addition matrix of this Pólya urn, say A_{dd}, is thus

$$A_{dd} = \begin{pmatrix} -1 & 2 & & & & \\ & -2 & 3 & & & \\ & & -3 & 4 & & \\ & & & \ddots & \ddots & \\ & & & & -(m-2) & m-1 \\ m & & & & & -(m-1) \end{pmatrix}$$

as given in [10, Section 8.2]. This addition matrix A_{dd} is similar to our transition matrix A given later on in (11). A straightforward computation gives

$$A = P^{-1}\, {}^{t}(A_{dd}) P,$$

where the change of basis from counting nodes to counting free places is given by matrix P

$$P = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & m-1 \end{pmatrix}.$$

(ii) System (5) is also the description of a discrete multitype branching process (X_n), where transitions from state $i = {}^t(i_1, \ldots, i_{m-1})$ to state $j = {}^t(j_1, \ldots, j_{m-1})$ in $\mathbb{R}^{m-1}$ are given by $\mathrm{P}(i, j) = \mathrm{P}(X_{n+1} = j | X_n = i)$: All the $\mathrm{P}(i, j)$ equal 0 except if $j = i + \Delta_k$, $1 \leq k \leq m-1$. In that case,

$$\mathrm{P}(i, i + \Delta_k) = \frac{1}{n}\, l_k(i).$$

The set of types is $S = \{1, 2, \ldots, m-1\}$ and the offspring distribution satisfies any moment condition, since the number of descendants is bounded above by m. These processes are well known in the homogeneous case where the transition *does not depend* on the current state [1, 13]. Thus we are in the so-called finite-type varying environment case, studied, for instance, in [7] and [2], mainly by martingale methods, in the same way as our Lemma 3 later on.

(iii) *Note on this kind of process.* The above random walk of an m-ary search tree belongs to a larger family of vector processes $(Z_n)_n$ in $\mathbb{R}^s$ (for any integer $s \geq 1$). Such a process can be defined as a random walk starting from some $Z_1 \in \mathbb{R}^s$, with random increments which take their values in a finite set of vectors $\{\Delta_1, \ldots, \Delta_s\}$:

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\forall\, n \geq 1, \qquad Z_{n+1} = Z_n + \Delta(n+1),$$

with the transition probabilities

$$\forall\, n \geq 1, \qquad \mathrm{P}(\Delta(n+1) = \Delta_i | Z_n) = \frac{1}{n}\, l_i(Z_n), \qquad 1 \leq i \leq s$$

where the l_i's are linear forms on $\mathbb{R}^s$. The process is Markovian and the transition probabilities between time n and time $n + 1$ depend linearly on the state at time n.

In order to guarantee that such a process is well defined, that is to say that the numbers $l_i(Z_n)/n$ are almost surely nonnegative and that their sum equals 1 for all n, one needs further assumptions on the parameters, namely, on Z_1, the l_i, and the Δ_i (all these assumptions are satisfied by m-ary search trees). First, we have hypotheses that allow Z_2 to be well defined:

$$\sum_{i=1}^{s} l_i(Z_1) = 1 \qquad \text{and} \qquad \forall\, j \in \{1, \ldots, s\}, \quad l_j(Z_1) \geq 0.$$

Then we have the hypotheses on the increments (an elementary induction shows that they are enough to make sure that the process is well defined): For all $j, k \in \{1, \ldots, s\}$,

$$\begin{cases} \sum_{i=1}^{s} l_i(\Delta_j) = 1, \\ j \neq k \Rightarrow l_j(\Delta_k) \geq 0, \\ l_j(\Delta_j) = 0 \qquad \text{or} \qquad l_j(Z_1)\mathbb{Z} + \sum_{i=1}^{s} l_j(\Delta_i)\mathbb{Z} = l_j(\Delta_j)\mathbb{Z}. \end{cases}$$

Only the diagonal terms $l_j(\Delta_j)$ are allowed to be negative. The last arithmetical condition just indicates that if $l_j(\Delta_j)$ is nonzero for some j, it divides (as a real number) $l_j(Z_1)$ and all the $l_j(\Delta_i)$.

The conditions defining such a model remain stable after an invertible linear change of coordinates. Keeping in mind remark (i), it means that these conditions are sufficient to guarantee that the corresponding generalized Pólya urn is tenable. The choice of a good basis of V is the key point in what follows.

3. EVOLUTION OF THE PROCESS AND AVERAGE-CASE ANALYSIS

Both are based on the computation of the conditional expectation:

$$E^{\mathcal{F}_n}(X_{n+1}) = \sum_{i=1}^{m-1} \frac{1}{n}\, l_i(X_n)(X_n + \Delta_i).$$

If one denotes by A the endomorphism of V defined by

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\forall\, Z \in V, \qquad AZ = \sum_{i=1}^{m-1} l_i(Z)\Delta_i,$$

one gets the following formula, which specifies that the above conditional expectation is a linear function of the state at time n:

$$E^{\mathscr{F}_n}(X_{n+1}) = \left(\mathrm{Id}_V + \frac{A}{n}\right)X_n, \tag{10}$$

where Id_V is the identity map of V. An immediate consequence of this fact is the computation of the expectation of the random vector X_n: Define $\Gamma_1 = \mathrm{Id}_V$ and

$$\Gamma_n = \prod_{k=1}^{n-1} \left(\mathrm{Id}_V + \frac{A}{k}\right)$$

for all $n \geq 2$, so that one gets the expression

$$E(X_n) = \Gamma_n X_1.$$

In the canonical basis of V, the matrix of A is

$$A = \begin{pmatrix} -1 & & & & & m(m-1) \\ 1 & -2 & & & & \\ & 2 & -3 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & -(m-2) & \\ & & & & m-2 & -(m-1) \end{pmatrix}, \tag{11}$$

where an empty entry means a zero entry. This matrix A is the *transition matrix* (*or endomorphism*) of the process. The characteristic polynomial of A is

$$\chi_A(z) = \prod_{k=1}^{m-1} (z+k) - m!. \tag{12}$$

The matrix A has only simple (complex) eigenvalues and 1 is the eigenvalue having the greatest real part. Furthermore, when m is even, 1 is the only real eigenvalue; when m is odd, the only other real eigenvalue is $-m-1$. Figure 2, made with the help of Maple, shows the complex eigenvalues of A when m equals 50. The plot of all roots of A in the complex plane seems to have always the same shape: regularly spaced points on the algebraic curve defined by equation $\prod_{1\leq k\leq m-1} |z+k|^2 = (m!)^2$. An important fact for the sequel is that **all the eigenvalues different from 1 have a real part less than 1/2 if and only if** $m \leq 26$.

We denote by $\mathrm{Sp}(A)$ the set of (complex) eigenvalues of A, and

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

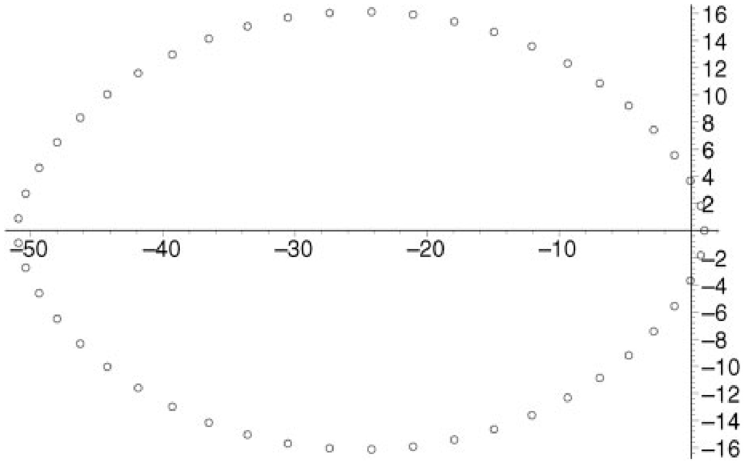

Fig. 2. Roots of χ_A when $m = 50$.

$$V_{\mathbb{C}} = \mathbb{C}^{m-1},$$

or, more precisely, $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$. For every $\lambda \in \mathrm{Sp}(A)$, we denote by π_λ the projection of $V_{\mathbb{C}}$ on the eigenspace $ker(A - \lambda\,\mathrm{Id}_{V_{\mathbb{C}}})$ relative to the decomposition

$$V_{\mathbb{C}} = \bigoplus_{\mu \in \mathrm{Sp}(A)} ker(A - \mu\,\mathrm{Id}_{V_{\mathbb{C}}}).$$

Then, define $\gamma_1(\lambda) = 1$ and

$$\gamma_n(\lambda) = \prod_{k=1}^{n-1} \left(1 + \frac{\lambda}{k}\right)$$

for all $n \geq 2$, so that the endomorphism Γ_n splits into the sum $\Gamma_n = \Sigma_\lambda\, \gamma_n(\lambda)\pi_\lambda$, and the expectation of X_n equals $E(X_n) = \Sigma_\lambda\, \gamma_n(\lambda)\pi_\lambda X_1$. Since 1 is the eigenvalue of A having the greatest real part, and since, by the Stirling formula,

$$\gamma_n(\lambda) = \frac{\Gamma(n + \lambda)}{\Gamma(\lambda + 1)\Gamma(n)} = \frac{n^\lambda}{\Gamma(\lambda + 1)} + O(n^{\lambda - 1}), \tag{13}$$

as n tends to infinity, the first term in the above expansion of $E(X_n)$ is $\gamma_n(1)\pi_1 X_1$, and one gets

$$\lim_{n\to\infty} \frac{E(X_n)}{n} = \pi_1 X_1. \tag{14}$$

Note that this limit is nonzero, since otherwise $E(X_n) = o(n)$ and taking the expectation

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

value in formula (8) provides a contradiction. The coordinates of the vector $\pi_1 X_1$ are explicitly given in Section 5 [see (17)], proving once again that $\pi_1 X_1$ is nonzero.

4. LOCAL STUDY ALONG PRINCIPAL DIRECTIONS

Keeping in mind the spectral decomposition of the process

$$X_n = \pi_1 X_n + \pi_{\lambda_2} X_n + \pi_{\overline{\lambda_2}} X_n + \sum_{\lambda \neq 1, \lambda_2, \overline{\lambda_2}} \pi_\lambda X_n,$$

we study locally in this section the projections $\pi_\lambda X_n$ for every eigenvalue λ of the transition matrix A, via three lemmas.

The first lemma describes explicitly the projection π_1 on the fixed points of the matrix A. It is available for every process at the one defined at the end of Section 2 (see Remark (iii)) as long as 1 is a *simple* eigenvalue of A. It is applied for the final result to $Y = X_n$ and $\pi_1 X_1$ is explicitly computed in Section 5, thus giving the first term of the expansion of X_n.

Lemma 1 (First Projection Lemma).

$$\forall\, Y \in V, \qquad \pi_1 Y = \left(\sum_{j=1}^{m-1} l_j(Y) \right) \pi_1 X_1.$$

Proof of Lemma 1. Let L be the endomorphism of V defined for all Y in V by

$$LY = \left(\sum_{j=1}^{m-1} l_j(Y) \right) \pi_1 X_1.$$

Note that L is nonzero because $\pi_1 X_1 \neq 0$ (recall the end of the previous section). Because of the relation $\sum_j l_j(\Delta_i) = 1$ for all i [see (9)], the value of L at each Δ_i is $\pi_1 X_1$. Thus, for all Y in V,

$$LA(Y) = L\left(\sum_{i=1}^{m-1} l_i(Y)\Delta_i \right) = \sum_{i=1}^{m-1} l_i(Y)\pi_1 X_1 = L(Y);$$

hence $LA = L$. But since $\pi_1 X_1$ is a fixed point of A, one has $AL = L$ too. Then A and L commute, and this product is L.

Since π_1 is a polynomial in A, the endomorphisms π_1 and L commute. Because $\sum_i l_i(X_1) = 1$ [relation (8) for $n = 1$], the endomorphisms π_1 and $L\pi_1$ (and $\pi_1 L$ too) have the same value at X_1. Since they are zero on the hyperplane spanned by the eigenvectors associated with eigenvalues different from 1, they are equal. Then, $\pi_1 = L\pi_1 = \pi_1 L$. But $\pi_1 L = L$, obviously. Thus $\pi_1 = L$. ∎

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lemma 2 gives the asymptotics of the moments of all projections $\pi_\lambda X_n$. The different behaviors of these moments appear, depending on the position of the real part of λ with respect to 1/2. This result is to be compared with similar ones for the second moments in the literature and contains the technical reason of the *phase transition* mentioned by these authors. It is available for every process defined at the end of Section 2. Notice that the L^2-convergence in the final result only requires the second moment asymptotics, but the almost sure convergence comes from the higher moments asymptotics.

Lemma 2 (Moments Lemma). *Let λ be an eigenvalue of A, σ its real part, $(\cdot|\cdot)$ any positive definite Hermitian form on $V_{\mathbb{C}}$, and $Z \in V_{\mathbb{C}}$. Then, for every nonnegative integer p, if $\sigma \neq 1/2$,*

$$\begin{cases} E(|(Z|\pi_\lambda X_n)|^{2p}) = O(n^p + n^{2p\sigma}), \\ E(|(Z|\pi_\lambda X_n)|^{2p+1}) = O(n^{p+1} + n^{2p\sigma+1}) \end{cases}$$

as n tends to infinity. If $\sigma = 1/2$,

$$\begin{cases} E(|(Z|\pi_\lambda X_n)|^{2p}) = O(n^p \log n) \\ E(|(Z|\pi_\lambda X_n)|^{2p+1}) = O(n^{p+1} \log n) \end{cases}$$

as n tends to infinity.

In other words, as n tends to infinity,

$$\text{if } \Re(\lambda) < 1/2, \qquad \text{then} \begin{cases} E(|(Z|\pi_\lambda X_n)|^{2p}) = O(n^p), \\ E(|(Z|\pi_\lambda X_n)|^{2p+1}) = O(n^{p+1}), \end{cases}$$

$$\text{if } \Re(\lambda) = 1/2, \qquad \text{then} \begin{cases} E(|(Z|\pi_\lambda X_n)|^{2p}) = O(n^p \log n), \\ E(|(Z|\pi_\lambda X_n)|^{2p+1}) = O(n^{p+1} \log n), \end{cases}$$

$$\text{if } \Re(\lambda) > 1/2, \qquad \text{then} \begin{cases} E(|(Z|\pi_\lambda X_n)|^{2p}) = O(n^{2p\sigma}), \\ E(|(Z|\pi_\lambda X_n)|^{2p+1}) = O(n^{2p\sigma+1}). \end{cases}$$

Remark. Note that we do not know if some value of m leads to $\Re(\lambda) = 1/2$ for some eigenvalue λ. It does not affect the final result.

Proof of Lemma 2. By induction of the integer p. If $p = 0$, only the assertion on the moment of order $2p + 1$ is nontrivial. If $\|\cdot\|$ denotes the norm associated with the Hermitian form, it follows directly from the definition of the process $(X_n)_n$ that almost surely

$$\|X_{n+1}\| \leq \|X_n\| + \max_{1 \leq i \leq m-1} \|\Delta_i\|$$

for every $n \geq 1$. Therefore, there is some positive constant c depending only on m such that, almost surely, for every $n \geq 1$,

$$\|X_n\| \leq cn. \tag{15}$$

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The result for $p = 0$ follows from this inequality.

Although it is not needed to make the proof complete, we proved the second moments inequality before presenting the induction; it helps the understanding of the general case, and it is used several times later on. An elementary computation of the conditional expectation, based on the dynamics of the process (X_n) leads to:

$$E^{\mathcal{F}_n}(|(Z|X_{n+1})|^2) = \sum_{i=1}^{m-1} \frac{1}{n} l_i(X_n)(Z|X_n + \Delta_i)\overline{(Z|X_n + \Delta_i)}$$

$$= |(Z|X_n)|^2 + 2\Re\left[\sum_{i=1}^{m-1} \frac{1}{n} l_i(X_n)(Z|X_n)\overline{(Z|\Delta_i)}\right] + \sum_{i=1}^{m-1} \frac{1}{n} l_i(X_n)|(Z|\Delta_i)|^2$$

$$= \Re\left[(Z|X_n)\overline{\left(Z\middle|\left(I + \frac{2A}{n}\right)X_n\right)}\right] + \sum_{i=1}^{m-1} \frac{1}{n} l_i(X_n)|(Z|\Delta_i)|^2.$$

Take now the expectation and apply this formula to the vector $\pi_\lambda^* Z$, where u^* denotes the adjoint endomorphism of u relative to the positive definite Hermitian form $(\cdot|\cdot)$. If σ is the real part of λ, one gets

$$E(|(Z|\pi_\lambda X_{n+1})|^2) = \left(1 + \frac{2\sigma}{n}\right)E(|(Z|\pi_\lambda X_n)|^2) + b_n,$$

where $b_n = \sum_i l_i(EX_n/n)|(Z|\pi_\lambda \Delta_i)|^2$, where the sum is extended to all i between 1 and $m - 1$. Since b_n has a limit as n tends to infinity [see (14)], $b_n = O(1)$. We get the explicit form

$$E(|(Z|\pi_\lambda X_n)|^2) = \gamma_n(2\sigma)\left(|(Z|\pi_\lambda X_1)|^2 + \sum_{k=1}^{n-1} \frac{b_k}{\gamma_{k+1}(2\sigma)}\right),$$

and, since by (13),

$$\gamma_n(2\sigma) = \frac{n^{2\sigma}}{\Gamma(1 + 2\sigma)} + O(n^{2\sigma-1}),$$

the above series has not the same behavior depending on the position of 2σ with respect to 1. This shows the following second moment asymptotics:

$$E(|(Z|\pi_\lambda X_n)|^2) = \begin{cases} O(n) & \text{if } \sigma < 1/2, \\ O(n \log n) & \text{if } \sigma = 1/2, \\ O(n^{2\sigma}) & \text{if } \sigma > 1/2. \end{cases} \tag{16}$$

Suppose now $p \ge 1$. On one hand, if x and y are complex numbers, the binomial

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

formula implies that $|x + y|^{2p} = |x^p + px^{p-1}y + z|^2$, where z is a polynomial in x and y whose degree in x equals $p - 2$. Thus

$$|x + y|^{2p} \leq |x|^{2p-2}\Re[x\overline{(x + 2py)}] + P(|x|, |y|)$$

where $P(X, Y)$ is a polynomial whose degree in X does not exceed $2p - 2$. On the other hand, the inequality (15) provides a positive constant (depending only on m) which bounds from above the number $|l_i(X_n)/n|$ for all i and for all n. The use of the last two facts to bound from above the conditional expectation

$$E^{\mathcal{F}_n}(|(Z|X_{n+1})|^{2p}) = \sum_{i=1}^{m-1} \frac{1}{n} l_i(X_n)|(Z|X_n) + (Z|\Delta_i)|^{2p}$$

leads to the existence of a polynomial Q of degree $\leq 2p - 2$ such that, for every $n \geq 1$,

$$E^{\mathcal{F}_n}(|(Z|X_{n+1})|^{2p}) \leq |(Z|X_n)|^{2p-2}\Re\left[(Z|X_n)\overline{\left(Z\left|\left(I + \frac{2pA}{n}\right)X_n\right.\right)}\right] + Q(|(Z|X_n)|).$$

Now, the same arguments as in the preceding proof for the second moments allow us to show the inequality

$$E(|(Z|\pi_\lambda X_{n+1})|^{2p}) \leq \left(1 + \frac{2p\sigma}{n}\right)E(|(Z|\pi_\lambda X_n)|^{2p}) + EQ(|(Z|\pi_\lambda X_n)|),$$

which gives the result by induction, assuming the result for all integers $< 2p$. Using (15), the result for the moments of order $2p + 1$ is a straightforward consequence of the inequality

$$E|(Z|X_n)|^{2p+1} \leq \max_\Omega |(Z|X_n)| \times E|(Z|X_n)|^{2p},$$

where Ω is the underlying probability space (see the beginning of Section 2). ∎

Lemma 3 is a direct consequence of Lemma 2. It makes precise the convergence of the martingale associated with X_n after rescaling with relation (1), establishing that some projections $\pi_\lambda X_n$ have an L^2 and a.s. limit. This lemma is applied as the final result to the first terms of the spectral decomposition of X_n.

Lemma 3 (L^2-Convergence Lemma). *Let λ be an eigenvalue of A. If $\Re(\lambda) > 1/2$, then the martingale $\gamma_n^{-1}(\lambda)\pi_\lambda X_n$ converges in L^2 (thus almost surely).*

Proof of Lemma 3. The random vector $\gamma_n^{-1}(\lambda)\pi_\lambda X_n$ is a $\mathcal{F}_n$ martingale from Eq. (1) because the restriction of Γ_n to the image of π_λ is the multiplication by $\gamma_n(\lambda)$. Moreover, under the hypothesis on $\Re(\lambda)$, estimation (16) on the second moments implies that

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$E(\|\pi_\lambda X_n\|^2) = O(n^{2\sigma})$; indeed, it is enough to choose a suitable (orthonormal) basis $(Z_i)_{1\le i\le m-1}$ of $V_{\mathbb{C}}$ such that

$$\|\pi_\lambda X_n\|^2 = \sum_{i=1}^{m-1} |(Z_i|\pi_\lambda X_n)|^2$$

and apply Lemma 2 to each vector Z_i. Combining this with (13) gives that

$$E(\|\gamma_n^{-1}(\lambda)\pi_\lambda X_n\|^2) = |\gamma_n^{-2}(\lambda)|E(\|\pi_\lambda X_n\|^2)$$

is a bounded sequence indexed by n so that the martingale $\gamma_n^{-1}(\lambda)\pi_\lambda X_n$ converges in L^2 and thus almost surely by standard theorems on martingales. ∎

5. THEOREM

Theorem. *Assume $m \geq 27$. Let $\lambda_2 = \sigma_2 + i\tau_2$ be the eigenvalue of the transition matrix A, having the second greatest real part σ_2 ($\sigma_2 > 1/2$) and a positive imaginary part $\tau_2 > 0$. For every eigenvalue λ of the transition matrix A, let π_λ be the projection on the eigenspace ker$(A - \lambda$ Id$)$ associated to λ, relative to the decomposition of $V_{\mathbb{C}}$ in eigenspaces of A. Let $X := \pi_1 X_1$.*

1.

$$X = \lim_{n\to\infty} \frac{EX_n}{n} = \frac{1}{H_m - 1}\begin{pmatrix} \frac{1}{1\times 2} \\ \frac{1}{2\times 3} \\ \vdots \\ \frac{1}{(m-1)\times m} \end{pmatrix}, \tag{17}$$

where H_m is the harmonic sum $H_m = \sum_{1\le k\le m} 1/k$.

2. *If Λ denotes the limit of the L^2-convergent martingale $\gamma_n^{-1}(\lambda_2)\pi_{\lambda_2}X_n$, then*

$$X_n = nX + 2\Re\left[\frac{n^{\lambda_2}\Lambda}{\Gamma(1+\lambda_2)}\right] + n^{\sigma_2}\varepsilon_n, \tag{18}$$

where the random vector ε_n converges to zero almost surely and in L^2 as n tends to infinity.

Corollary 1. *With the same notations as in the theorem, let C and S be the real (and nonrandom) vectors of $V_{\mathbb{C}}$ defined by the relation*

$$\pi_{\lambda_2}X_1 = C - iS. \tag{19}$$

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Let ρ *and* φ *be, respectively, the modulus and the argument of the random vector* $2\Lambda/(\Gamma(1 + \lambda_2))$ *along the line generated by* $\pi_{\lambda_2}X_1$:

$$\rho \exp(i\varphi)\pi_{\lambda_2}X_1 = \frac{2\Lambda}{\Gamma(1 + \lambda_2)}, \qquad \rho \geq 0, \quad \varphi \in [0, 2\pi]. \tag{20}$$

Then

$$X_n = nX + n^{\sigma_2}\rho(C \cos(\tau_2 \log n + \varphi) + S \sin(\tau_2 \log n + \varphi)) + n^{\sigma_2}\varepsilon_n,$$

where the random vector ε_n *converges to zero almost surely and in* L^2 *as* n *tends to infinity.*

In other words, the random vector

$$\frac{X_n - nX}{n^{\sigma_2}} - \rho(C \cos(\tau_2 \log n + \varphi) + S \sin(\tau_2 \log n + \varphi))$$

converges to zero almost surely and in L^2.

The corollary is a straightforward consequence of the theorem. Just write Λ in (20) as the product of a complex random variable and of the nonrandom complex vector $\pi_{\lambda_2}X_1$, and separate the real and imaginary parts of $\pi_{\lambda_2}X_1$ (19). Also note that $n^{\lambda_2} = n^{\sigma_2}e^{i\tau_2 \log n}$. Notice that X, C, and S are linearly independent vectors of $V_{\mathbb{C}}$ (because $\pi_{\lambda_1}X_1$, $\pi_{\lambda_2}X_1$, and $\pi_{\overline{\lambda_2}}X_1$ are).

Computation of X, C, S. X is the projection of the first vector X_1 on the vector line of the fixed vectors of A. The first equality of (17) has already been given [see (14)]. An easy computation [compute a fixed vector, and add the condition $\Sigma_i\ l_i(\pi_1X_1) = \lim_n \Sigma_i\ l_i(E(X_n/n)) = 1$)] gives (17).

To express the vectors C and S, we sum up how one can compute the projection $\pi_\lambda X_1$ of X_1 on the eigenspace $ker(A - \lambda\ \mathrm{Id})$ for every eigenvalue λ, and give the result: For each λ, compute first the eigenvector of A associated with λ having 1 as the $(m - 1)$st coordinate. Call it F_λ. Decompose $X_1 = \Sigma_{\lambda \in \mathrm{Sp}(A)}\ a_\lambda F_\lambda$, where a_λ is the complex number such that $\pi_\lambda X_1 = a_\lambda F_\lambda$. Then, for all $p \geq 0$, one has $A^pX_1 = \Sigma_\lambda\ a_\lambda \lambda^p F_\lambda$. With the explicit form of A, one can easily compute the $(m - 1)$st coordinate of the vectors A^pX_1 for $0 \leq p \leq m - 2$ (induction shows that its pth coordinate is $p!$ and its jth ones are zero for all $j \geq p + 1$) and solve the system

$$dx_{m-1}A^pX_1 = \sum_{\lambda \in \mathrm{Sp}(A)} a_\lambda \lambda^p, \qquad 0 \leq p \leq m - 2$$

with Cramer's formula. Likewise, one writes the number a_λ as the product of $(m - 2)!$ by the quotient of two Vandermonde determinants. After simplification, one gets $a_\lambda = (m - 2)!/\chi_A'(\lambda)$, where χ_A denotes the characteristic polynomial of A [see (12)]. The computation of the logarithmic derivative of $\chi_A + m!$ gives the expression

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\chi'_A(\lambda) = \prod_{\mu \in \mathrm{Sp}(A)\setminus\{\lambda\}} (\lambda - \mu) = m! \sum_{j=1}^{m-1} \frac{1}{\lambda + j}.$$

The result is now the following: For every eigenvalue $\lambda \in \mathrm{Sp}(A)$,

$$\pi_1 X_1 = \frac{1}{\chi'_A(\lambda)} \begin{pmatrix} \varpi_1(\lambda) \\ \vdots \\ \varpi_{m-1}(\lambda) \end{pmatrix},$$

where, for every $j \in \{1, \ldots, m-1\}$,

$$\varpi_j(\lambda) = (j-1)! \prod_{k=j+1}^{m-1} (k + \lambda) = (j-1)! \frac{\Gamma(m+\lambda)}{\Gamma(j+1+\lambda)} = \frac{m!}{j\gamma_{j+1}(\lambda)}.$$

Proof of the Theorem. The proof consists of examining the decomposition

$$X_n = \pi_1 X_n + \pi_{\lambda_2} X_n + \pi_{\overline{\lambda_2}} X_n + \sum_{\Re(\lambda) < \sigma_2} \pi_\lambda X_n, \tag{21}$$

in order to get the expected asymptotic order of magnitude of each term.

The first projection lemma describes the first term, because relation (8) between the number of nodes of each type gives that, for every n,

$$\sum_{i=1}^{m-1} l_i(X_n) = n,$$

so that

$$\pi_1 X_n = n \pi_1 X_1. \tag{22}$$

For the following two terms in (21), recall that the assumption $m \geq 27$ implies that $\sigma_2 > 1/2$. Let

$$\Lambda = \lim_{n \to +\infty} \gamma_n^{-1}(\lambda_2) \pi_{\lambda_2} X_n,$$

and notice that the random vector Λ is both the L^2 and the almost sure limit of this martingale as guaranteed by the L^2-convergence lemma (Lemma 3) applied to λ_2. In other words,

$$\gamma_n^{-1}(\lambda_2) \pi_{\lambda_2} X_n = \Lambda + \varepsilon_n,$$

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where ε_n tends to zero a.s. and in L^2. In the following, ε_n always denotes a generic random variable which tends to zero a.s. and in L^2 when n tends to infinity, even if it changes from place to place.

Multiply by $\gamma_n(\lambda_2)$ the previous equality and recall the asymptotics of the γ_n given in formula (13) to get

$$\pi_{\lambda_2} X_n = \gamma_n(\lambda_2)\Lambda + n^{\sigma_2}\varepsilon_n = \frac{n^{\lambda_2}}{\Gamma(1+\lambda_2)}\Lambda + n^{\sigma_2}\varepsilon_n.$$

Summing with $\pi_{\overline{\lambda_2}} X_n$ and noticing that $\pi_{\bar{\lambda}} X_n = \overline{\pi_\lambda X_n}$ gives

$$\pi_{\lambda_2} X_n + \pi_{\overline{\lambda_2}} X_n = 2\Re\left[\frac{n^{\lambda_2}\Lambda}{\Gamma(1+\lambda_2)}\right] + n^{\sigma_2}\varepsilon_n, \tag{23}$$

which provides the second term in (18).

It remains to show that if λ is an eigenvalue of A different from 1, λ_2, and $\overline{\lambda_2}$, then $\pi_\lambda X_n = n^{\sigma_2}\varepsilon_n$, where ε_n tends to zero a.s. and in L^2 as n tends to infinity. Let σ be the real part of such an eigenvalue; we know that $\sigma < \sigma_2$. The case $\sigma > 1/2$ is easy: Lemma 3 of martingale convergence still holds; hence L^2 and almost sure convergence are shown together for the martingale $\gamma_n^{-1}(\lambda)\pi_\lambda X_n$. Thus $\pi_\lambda X_n$ is of order n^σ, which is negligible to n^{σ_2}.

In case $\sigma \le 1/2$, let us first prove L^2-convergence: As in the proof of Lemma 3, we have as a corollary of the moments lemma

$$E(\|\pi_\lambda X_n\|^2) = O(n) \qquad \text{or} \qquad O(n \log n);$$

hence (recall that $\sigma_2 > 1/2$),

$$\frac{\pi_\lambda X_n}{n^{\sigma_2}} \underset{L_2}{\longrightarrow} 0.$$

For the almost sure convergence to zero of $\pi_\lambda X_n/n^{\sigma_2}$, we use the Borel-Cantelli lemma: It is sufficient to show that, for any $\varepsilon > 0$, the series $\sum_n P(\|\pi_\lambda X_n/n^{\sigma_2}\| > \varepsilon)$ is convergent. By the Markov inequality, it is sufficient to show that, for some integer p, the moment $E\|\pi_\lambda X_n/n^{\sigma_2}\|^{2p}$ is the general term of a convergent numerical series. It is true, for p large enough, because of the moments lemma: For every positive definite Hermitian form and complex vector Z, for every nonnegative integer p,

$$E\left(\left|\left(Z\middle|\frac{\pi_\lambda X_n}{n^{\sigma_2}}\right)\right|^{2p}\right) = O\left(\frac{1}{n^{p(2\sigma_2-1)}}\right).$$

Summarizing, for every eigenvalue λ of A different from 1, λ_2, and $\overline{\lambda_2}$,

$$\frac{\pi_\lambda X_n}{n^{\sigma_2}} \underset{n\to\infty}{\longrightarrow} 0 \qquad \text{a.s. and in } L^2. \tag{24}$$

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

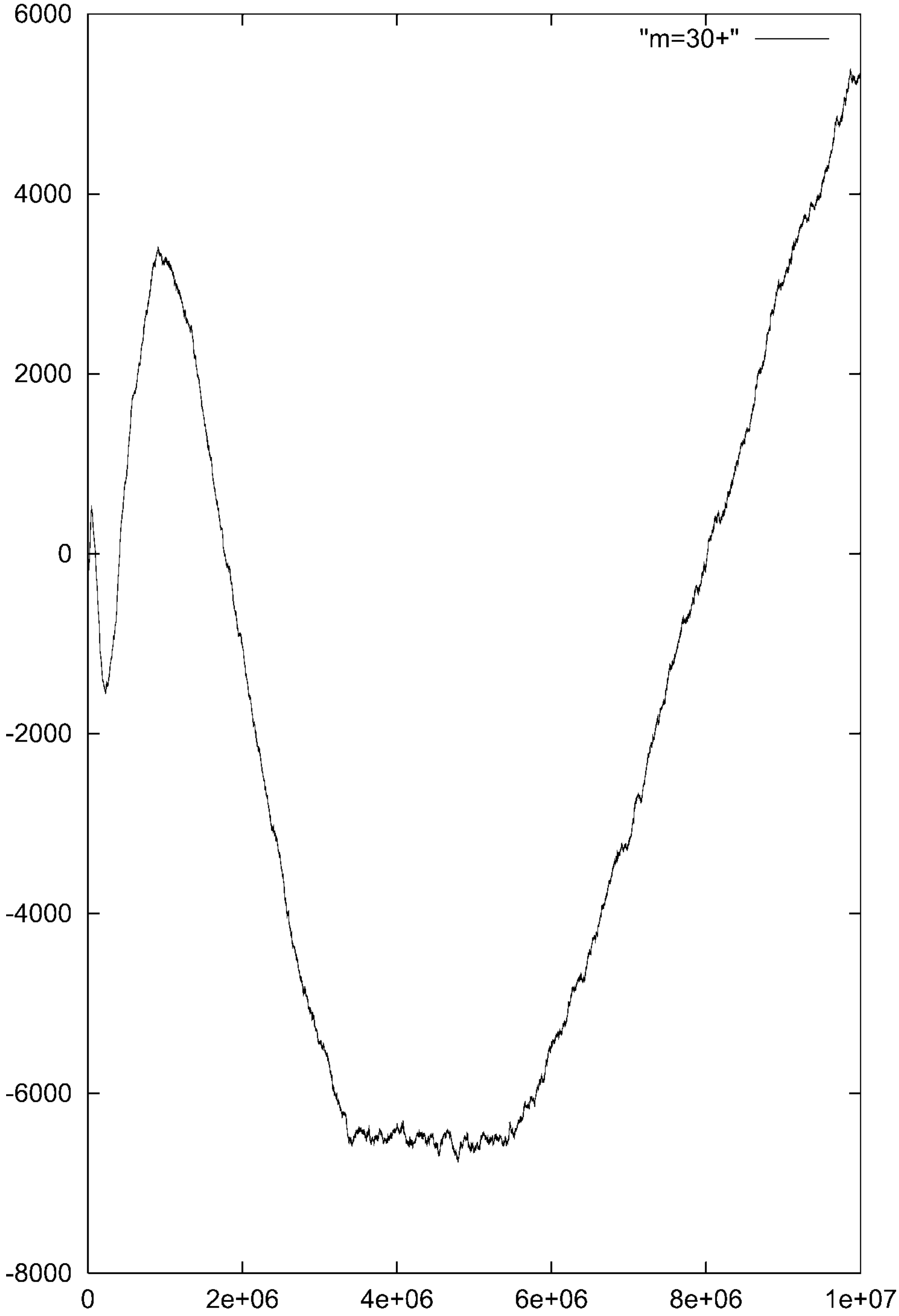

Fig. 3. $m = 30$.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

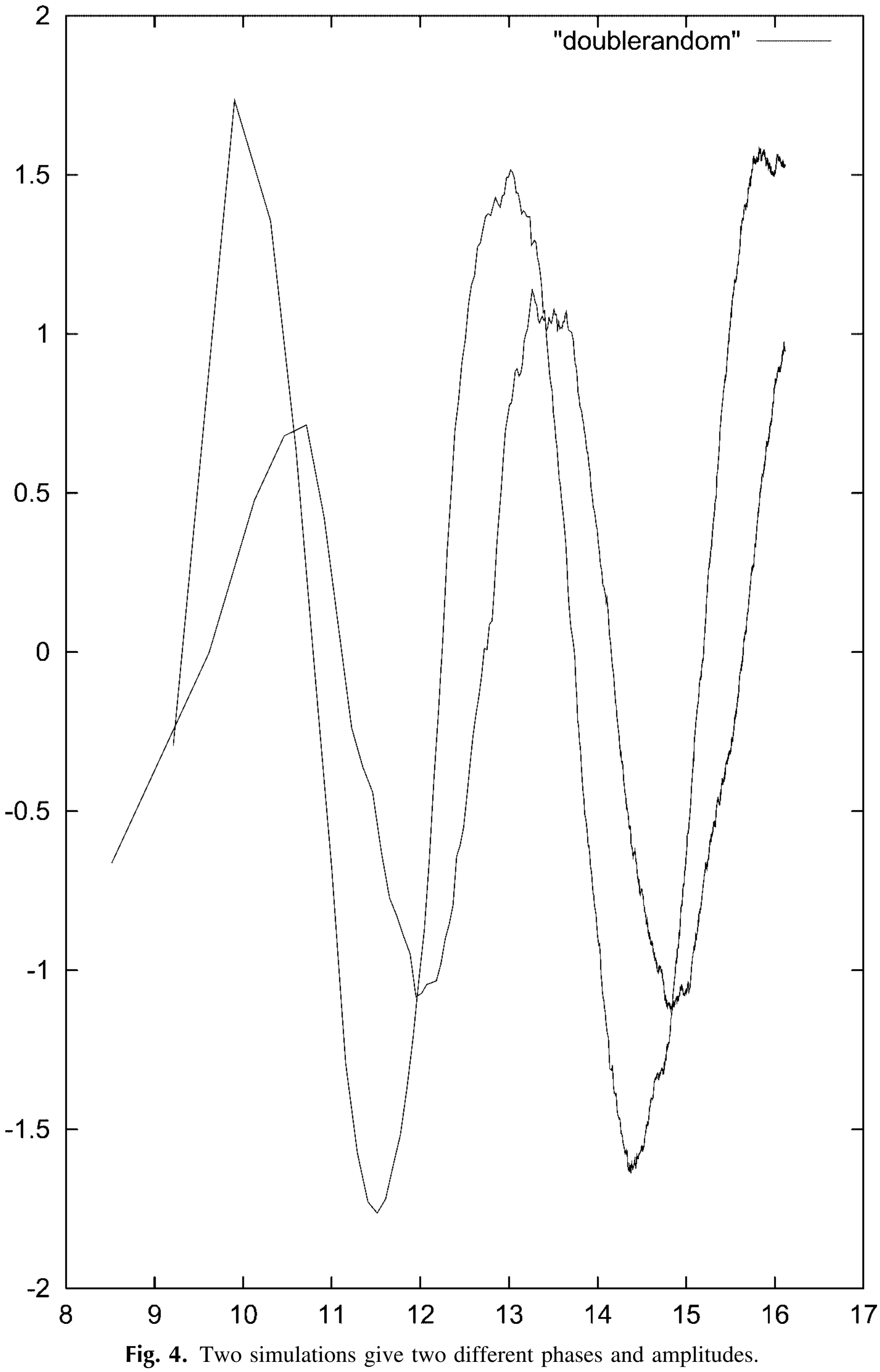

Fig. 4. Two simulations give two different phases and amplitudes.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To get the final result, it is now enough use (21), (22), (23), and (24) together. ■

Corollary 2. *Suppose $m \geq 27$. If χ is any linear form on V, then there exist a real number x_χ and real random variables ρ_χ and φ_χ such that*

$$\chi(X_n) = nx_\chi + n^{\sigma_2}\rho_\chi \cos(\tau_2 \log n + \varphi_\chi) + n^{\sigma_2}\varepsilon_n,$$

where ε_n tends to zero almost surely and in L^2 as n tends to infinity.

To prove this, see what happens to $\chi(X_n)$ with Corollary 1, and put the sine and cosine terms together to get a new random phase and a new random amplitude.

This corollary describes for example the asymptotic behavior of the number of nodes of a given type (take $\chi = dx_i$, the ith coordinate of $\mathbb{R}^{m-1}$), or of the total (except saturated nodes) number of nodes (take $\chi = \Sigma\, dx_i$, where i ranges over all i between 1 and $m - 1$).

The following question naturally arises: What are the laws of the random variables ρ and φ of the theorem?

6. SIMULATIONS

Figure 3 represents simulations for the total number of nodes for $m = 30$. We put the number n of keys inserted in the tree on the x-axis, and $x_n - nx_\chi$ on the y-axis, where x_n is the total number of nodes (except saturated nodes, those with $m - 1$ keys) at time n and x_χ the coefficient $\lim_{+\infty} E(x_n)/n$ of its drift. The graph remains fairly smooth around an "$n^{\sigma_2}\cos \log n$" curve. Note that we only drew one point over one thousand.

Figure 4 illustrates the random amplitude ρ_χ and the random phase φ_χ for the asymptotics of the total number of nodes x_n: On the x-axis, $\log n$; on the y-axis, $(x_n - nx_\chi)/n^{\sigma_2}$ for two simulations. Note the difference between the amplitudes and the phases of both simulations.

ACKNOWLEDGMENTS

We are indebted to the referees for careful reading and accurate suggestions. We warmly thank Jean-François Marckert for helpful discussions.

REFERENCES

[1] K. B. Athreya and P. Ney, Branching processes, Springer-Verlag, New York, 1972.

[2] J. D. Biggins, H. Cohn, and O. Nerman, Multi-type branching in varying environment, Stochastic Process Appl 83(2) (1999), 357–400.

[3] H. H. Chern and H. K. Hwang, Phase changes in random m-ary trees and generalized quicksort, Random Structures Algorithms 19(3–4) (2001), 316–358.

[4] D. S. Dean and S. N. Majumdar, Phase transition in a random fragmentation problem with applications to computer science, J Phys A 35 (2002), 501–507.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[5] H. K. Hwang, Second phase changes in random m-ary trees and generalized quicksort: Convergence rates, Ann Probab 31(2) (2003), 609–629.

[6] J. Jabbour-Hattab, Martingales and large deviations for binary search trees, Random Structures Algorithms 19(2) (2001), 112–127.

[7] O. D. Jones, On the convergence of multitype branching processes with varying environments, Ann Appl Probab 7(3) (1997), 772–801.

[8] W. Lew and H. Mahmoud, The joint distribution of elastic buckets in multiway search trees, SIAM J Comput 23(5) (1994), 1050–1074.

[9] H. M. Mahmoud, Evolution of random search trees, Wiley, New York, 1992.

[10] H. M. Mahmoud, The size of random bucket trees via urn models, Acta Inform 38 (2002), 813–838.

[11] H. M. Mahmoud and B. Pittel, Analysis of the space of search trees under the random insertion algorithm, J Algorithms 10 (1989), 52–75.

[12] H. M. Mahmoud and R. T. Smythe, Probabilistic analysis of bucket recursive trees, Theoret Comput Sci 144 (1995), 180–205.

[13] C. J. Mode, "Multitype branching processes. Theory and applications." Modern analytic and computational methods in science and mathematics, Volume 34. Elsevier, Amsterdam, 1971.

[14] R. Neininger and L. Rueschendorf, A general limit theorem for recursive algorithms and combinatorial structures, Ann Appl Probab (2003), to appear, http://www.stochastik.uni-freiburg.de/homepages/rueschendorf/papers/general.ps.

[15] J. Neveu, Arbres et processus de Galton-Watson. Ann Inst Henri Poincaré 22(2) (1986), 199–207.

[16] R. T. Smythe, Central limit theorems for urns models, Stochastic Process Appl 65 (1996), 115–137.

10982418, 2004, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.10108 by Université de Versailles-Saint-Quentin-en-Yvelines, Wiley Online Library on [23/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License