Available online at www.sciencedirect.com

SCIENCE DIRECT®

Theoretical Computer Science 356 (2006) 153 – 169

Theoretical Computer Science

www.elsevier.com/locate/tcs

Pattern statistics and Vandermonde matrices☆

Massimiliano Goldwurm*, Violetta Lonati

Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano, Via Comelico 39, 20135 Milano, Italy

Abstract

In this paper, we determine some limit distributions of pattern statistics in rational stochastic models. We present a general approach to analyze these statistics in rational models having an arbitrary number of strongly connected components. We explicitly establish the limit distributions in most significant cases; they are characterized by a family of unimodal density functions defined by means of confluent Vandermonde matrices.

Keywords: Automata and formal languages; Limit distributions; Pattern statistics; Rational formal series; Vandermonde matrices

1. Introduction

This work presents some results on the limit distribution of pattern statistics. The major problem in this context is to estimate the frequency of pattern occurrences in a random text. This is a classical problem that has applications in several research areas of computer science and biology: for instance, it is considered in connection with the search of motifs in DNA sequences [17] while the earlier motivations are related to code synchronization [10,11] and approximated pattern-matching [13,22]. In a general probabilistic framework [18,16,3], given one or more patterns, defined as strings over a finite alphabet Σ, and a probabilistic source P generating words at random over Σ, one considers the number X_n of occurrences of patterns in a word of length n generated by P.[1] Typical goals are the asymptotic evaluation of the moments of X_n, in particular its mean value and variance, its limit distribution, the local limit properties and the corresponding large deviations. The results depend in particular on the stochastic model P, which is usually assumed to be a Bernoulli model [11] or a Markovian model [18,16]. For instance, in [16] Gaussian limit distributions are obtained, for any regular set of patterns and any Markovian source P, under a primitivity hypothesis on the associated stochastic matrix.

In our paper, we assume the so-called rational stochastic model, introduced in [2], which includes the traditional Markovian model as a particular case. In our framework, the pattern is reduced to the single symbol a while the text is a word of length n over the alphabet $\{a, b\}$ generated at random according to a probability distribution defined by

☆ This work includes results presented in the paper "M. Goldwurm and V. Lonati, Pattern Occurrences in Multicomponent Models, *Proceedings of the 22nd S.T.A.C.S.*, Lecture Notes in Computer Science, Vol. n. 3404, 680–692, Springer 2005". This research has been supported by Project MIUR PRIN 2006-2007 "Automata and Formal Languages: Mathematical and Applicative Aspects".

* Corresponding author. Tel.: +39 02 50316233; fax: +39 02 50316373.

E-mail addresses: goldwurm@dsi.unimi.it (M. Goldwurm), lonati@dsi.unimi.it (V. Lonati).

[1] Here, an occurrence is a position where a pattern ends in the text.

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.01.033

means of a rational formal series with nonnegative real coefficients and noncommutative variables a, b. Such a setting can simulate any Markovian source over an arbitrary finite alphabet Σ for any regular set of patterns in Σ^* [2].

Also in the rational stochastic models, Gaussian limit distributions are obtained under a primitive hypothesis, i.e. when the matrix associated with the rational formal series (counting the transitions between states) is primitive [2]. A complete study of the limit distributions is given in [5] in the bicomponent rational models, that is when the graph corresponding to the previous matrix consists of two strongly connected components.

Here, we present a general approach to the analysis of rational stochastic models with an arbitrary number of strongly connected components (called multicomponent models), explicitly establishing the limit distribution of the corresponding pattern statistics in most significant cases. The main result shows that such a limit distribution is related to the confluent Vandermonde matrices, a generalization of the classical Vandermonde matrices used in several research areas and in particular in Automatic Control Theory [4,15].

The material we present is organized as follows. In Section 3 we recall the notion of confluent Vandermonde matrix and some of its properties; in particular, we show how this is related to the convolution of a finite set of sequences. In Section 4 we introduce a family of probability distributions defined by means of confluent Vandermonde matrices and establish their main properties. We call them Vandermonde distributions. In particular, we prove that their density functions are unimodal and we compute their characteristic functions. In Section 5 we start our analysis of pattern statistics and present the rational stochastic models, discussing the natural decomposition in strongly connected components; in particular, we introduce the notions of dominant component and main chain and show their role in the analysis of multicomponent models. In Section 6 we present our main result, which concerns the simple models (those with just one main chain which in addition only has primitive dominant components); in this case, assuming a mild variability condition on the dominant components, we determine the limit distribution of our pattern statistics showing that it is a Vandermonde distribution. Finally, in Section 7, we characterize the limit distributions for all simple models and provide a natural method to determine the limit distribution in the general case.

2. Preliminary notions

Generating functions represent the main tool we use in this study (see for instance [6] or [20, Chapter 3]). We recall that the (ordinary) generating function of a sequence $\{g_n\} \subseteq \mathbb{C}$ is the analytic function $g(w)$ that admits the Taylor expansion $g(w) = \sum_0^{+\infty} g_n w^n$ for every w in an open neighbourhood of 0. In our analysis we often have to evaluate the asymptotic growth of sequences having a rational generating function. To this end we make use of the following well-known properties that allow to extract informations on the growth of a sequence from the singularities of its generating function.

Let $g(w)$ be the generating function of a sequence $\{g_n\} \subseteq \mathbb{C}$; consider the radius of convergence R of the power series $\sum_0^{+\infty} g_n w^n$ and assume R is finite. We first observe that $g_n = \mathrm{O}(r^{-n})$ for every real r such that $0 < r < R$. Moreover, let $\alpha_1, \alpha_2 \ldots, \alpha_j$ be the singularities of $g(w)$ of modulus smaller than T, for some $T > R$. If all α_i's are simple poles then $g_n = \sum_{i=1}^{j} c_i \alpha_i^{-n} + \mathrm{O}(\rho^n)$ for some $0 < \rho < R^{-1}$ and some nonnull values $c_i \in \mathbb{C}$, $i = 1, \ldots, j$. On the contrary, if each α_i is a pole of degree k_i, then $g_n = \sum_{i=1}^{j} c_i \alpha_i^{-n} n^{k_i - 1}(1 + \mathrm{O}(1/n))$ where $c_i \in \mathbb{C}$ is nonnull for every $i = 1, \ldots, j$.

We finally recall that the product of two generating functions is the generating function of the convolution of the associated sequences. More generally, if $g^{(i)}(w)$ is the generating function of the sequence $\{g_n^{(i)}\}$ for each $i = 1, 2, \ldots, k$, then $f(w) = \prod_{i=1}^{k} g^{(i)}(w)$ is the generating function of the sequence $\{f_n\}$ such that, for every $n \in \mathbb{N}$,

$$f_n = \sum_{n_1 + \cdots + n_k = n} g_{n_1}^{(1)} g_{n_2}^{(2)} \cdots g_{n_k}^{(k)}.$$

3. Confluent Vandermonde matrices

Vandermonde matrices are defined by linear systems of equations whose solution yields the coefficients of polynomials of smallest degree with a given set of distinct roots [14]. When roots are associated with a given multiplicity an analogous system of equations can be defined that leads to a generalized version of Vandermonde matrix, called confluent Vandermonde matrix. That one plays a remarkable role in Automatic Control Theory [4]; in particular, its inverse is useful to compute the solutions of linear systems of differential equations [15]. In this section we recall the main properties of such matrices; our main goal is to present Proposition 3, which shows how the inverse of a confluent Vandermonde matrix can be used to compute the terms of the convolution of a family of sequences.

Given two integers k, r such that $2 \leqslant r \leqslant k$, let $(v_1, v_2, \ldots, v_r)$ be a tuple of distinct complex numbers and let $(m_1, m_2, \ldots, m_r) \in \mathbb{N}^r$ be an associated tuple of multiplicities, such that $m_1 + m_2 + \cdots + m_r = k$ and $m_i \geqslant 1$ for each $i = 1, 2, \ldots, r$. Consider the monic polynomial

$$D(x) = \prod_{\ell=1}^{r} (x - v_\ell)^{m_\ell} = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0. \tag{1}$$

The *confluent Vandermonde matrix* associated with $D(x)$ is defined by $V = [V_1|V_2|\cdots|V_r]$ where, for each $\ell = 1, 2, \ldots, r$, V_ℓ is the matrix of size $(k \times m_\ell)$ such that

$$(V_\ell)_{hj} = \begin{cases} \binom{h-1}{j-1} v_\ell^{h-j} & \text{if } j \leqslant h \\ 0 & \text{otherwise.} \end{cases}$$

for every $h = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, m_\ell$.

For instance if $r = 2$, $m_1 = 3$ and $m_2 = 4$, then V is given by

$$V = \left(\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ v_1 & 1 & 0 & v_2 & 1 & 0 & 0 \\ v_1^2 & 2v_1 & 1 & v_2^2 & 2v_2 & 1 & 0 \\ v_1^3 & 3v_1^2 & 3v_1 & v_2^3 & 3v_2^2 & 3v_2 & 1 \\ v_1^4 & 4v_1^3 & 6v_1^2 & v_2^4 & 4v_2^3 & 6v_2^2 & 4v_2 \\ v_1^5 & 5v_1^4 & 10v_1^3 & v_2^5 & 5v_2^4 & 10v_2^3 & 10v_2^2 \\ v_1^6 & 6v_1^5 & 15v_1^4 & v_2^6 & 6v_2^5 & 15v_2^4 & 20v_2^3 \end{array}\right).$$

In the special case when $m_\ell = 1$ for every $\ell = 1, 2, \ldots r$, V reduces to the standard Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ v_1 & v_2 & \ldots & v_k \\ v_1^2 & v_2^2 & \ldots & v_k^2 \\ \ldots & \ldots & \ldots & \ldots \\ v_1^{k-1} & v_2^{k-1} & \ldots & v_k^{k-1} \end{pmatrix}. \tag{2}$$

It is well-known that V always is nonsingular and that its determinant is $\prod_{1 \leqslant i < j \leqslant r} (v_i - v_j)^{m_i m_j}$.

3.1. Inverse of a confluent Vandermonde matrix

Some identities we use in subsequent sections concern the inverse of V and especially the entries of its last column. An explicit expression for all entries of V^{-1} is presented in [4, Eq. (9)]. Here, we recall that the last column of V^{-1} is given by the vector

$$\underline{w} = (w_{11}, w_{12}, \ldots w_{1m_1}|w_{21}, w_{22}, \ldots w_{2m_2}|\ldots|w_{r1}, w_{r2}, \ldots w_{rm_r})^{\mathrm{T}} \tag{3}$$

where for every $\ell = 1, 2, \ldots r$ and $j = 1, \ldots, m_\ell$, we have

$$\frac{1}{D(x)} = \sum_{\ell=1}^{r} \sum_{j=1}^{m_\ell} \frac{w_{\ell j}}{(x - v_\ell)^j} \tag{4}$$

and the following differential formula holds

$$w_{\ell j} = \frac{1}{(m_\ell - j)!} \cdot \frac{\mathrm{d}^{m_\ell - j}}{\mathrm{d}x^{m_\ell - j}} \left[\frac{1}{\prod_{i \neq \ell} (x - v_i)^{m_i}} \right]_{|x = v_\ell} . \tag{5}$$

Notice that for $x \neq v$ we have

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} (x - v)^{-m} = (-1)^n n! \binom{m + n - 1}{m - 1} (x - v)^{-m-n}.$$

hence, by applying Leibniz differentiation rule

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} (f_1(x) \cdot f_2(x) \cdots f_r(x)) = n! \sum_{n_1 + n_2 + \cdots + n_r = n} \left(\prod_i \frac{1}{n_i!} \cdot \frac{\mathrm{d}^{n_i}}{\mathrm{d}x^{n_i}} f_i(x), \right)$$

we get the following expression for every $w_{\ell j}$

$$w_{\ell j} = (-1)^{m_\ell - j} \sum_{\sum_{i \neq \ell} n_i = m_\ell - j} \prod_{i \neq \ell} \binom{n_i + m_i - 1}{m_i - 1} (v_\ell - v_i)^{-m_i - n_i}. \tag{6}$$

Proposition 1. *Let $D(x)$ be the polynomial defined by Eq.* (1) (*with distinct v_ℓ's*). *Consider the confluent Vandermonde matrix associated with $D(x)$ and let $\underline{w}$ be the vector defined in* (3). *Then, for every $s = 1, 2, \ldots k - 1$ the following polynomial is identically null*

$$P_s(x) = \sum_{\ell=1}^{r} \sum_{j=1}^{\min(s, m_\ell)} \binom{s-1}{j-1} w_{\ell j} (v_\ell - x)^{s-j}.$$

Moreover, $P_k(0) = 1$.

Proof. First notice that $P_k(0) = 1$ and $P_s(0) = 0$ for every $s = 1, 2, \ldots, k - 1$. Indeed, such equalities can be written in matrix form as $V \cdot \underline{w} = (0, \ldots, 0, 1)_T$, which holds true by definition of V and $\underline{w}$. Now, fix an integer $1 \leqslant s \leqslant k - 1$. Replacing $(v_\ell - x)^{s-j} = \sum_{h=j}^{s} \binom{s-j}{h-j} v_\ell^{h-j} (-x)^{s-h}$ in $P_s(x)$ we get

$$P_s(x) = \sum_{\ell=1}^{r} \sum_{j=1}^{\min(s, m_\ell)} \sum_{h=j}^{s} \binom{s-1}{h-1} \binom{h-1}{j-1} w_{\ell j} \, v_\ell^{h-j} (-x)^{s-h}.$$

Since the set $\{(h, j) \in \mathbb{N}^2 \mid 1 \leqslant j \leqslant \min(s, m_\ell),\ j \leqslant h \leqslant s\}$ equals the set $\{(h, j) \in \mathbb{N}^2 \mid 1 \leqslant h \leqslant s,\ 1 \leqslant j \leqslant \min(h, m_\ell)\}$, the previous expression can be written as

$$P_s(x) = \sum_{h=1}^{s} \binom{s-1}{h-1} (-x)^{s-h} \sum_{\ell=1}^{r} \sum_{j=1}^{\min(h, m_\ell)} \binom{h-1}{j-1} w_{\ell j} \, v_\ell^{h-j} = \sum_{h=1}^{s} \binom{s-1}{h-1} (-x)^{s-h} P_h(0)$$

which is identically null by the previous reasoning. □

Corollary 2. *Let V be the Vandermonde matrix defined in* (2), *where the v_ℓ's are all distinct. Then, the entries of the last column of V^{-1} are given by $c_\ell = \prod_{i \neq \ell} (v_\ell - v_i)^{-1}$ for $\ell = 1, 2, \ldots k$ and satisfy $\sum_\ell c_\ell \, {v_\ell}^{k-1} = 1$, $\sum_\ell c_\ell (v_\ell - x)^{s-1} = 0$ for every $s = 1, 2, \ldots k - 1$.*

3.2. Multiple convolutions

Confluent Vandermonde matrices are related to the properties of convolutions of families of sequences. More precisely, consider the rational function

$$\frac{D(0)}{D(x)} = \prod_{\ell=1}^{r} \left(\frac{-v_\ell}{x - v_\ell} \right)^{m_\ell} = \prod_{\ell=1}^{r} \left(1 - \frac{x}{v_\ell} \right)^{-m_\ell}$$

and observe that each $(1 - x/v_\ell)^{-m_\ell}$ is the generating function of $\left\{\binom{n+m_\ell-1}{m_\ell-1} v_\ell^{-n}\right\}_n$. Therefore $D(0)/D(x)$ is the generating function of the sequence $\{g_D(n)\}_n$ defined by their convolution, i.e.,

$$g_D(n) = \sum_{\sum_\ell n_\ell = n} \prod_{\ell=1}^{r} \binom{n_\ell + m_\ell - 1}{m_\ell - 1} v_\ell^{-n_\ell}. \tag{7}$$

A key remark for the subsequent discussion is to notice that

$$g_D(n) = \sum (v_1^{-n_{11}} \cdots v_1^{-n_{1m_1}}) \cdot (v_2^{-n_{21}} \cdots v_2^{-n_{2m_2}}) \cdots (v_r^{-n_{r1}} \cdots v_r^{-n_{rm_r}}),$$

where the sum is extended over all the k-tuples of nonnegative exponents $(n_{11}, n_{12}, \ldots, n_{rm_r})$ whose sum equals n. In other words, $\{g_D(n)\}_n$ is the convolution of the sequences $\{v_\ell^{-n}\}_n$, each of them taken with multiplicity m_ℓ.

Proposition 3. *Let V be the confluent Vandermonde matrix associated with the polynomial $D(x)$ defined by Eq.* (1) *and assume that all roots v_ℓ's are non-null. Also, let $g_D(n)$ be defined by Eq.* (7) *for every $n \in \mathbb{N}$. Then,*

$$g_D(n) = D(0) \cdot \sum_{\ell=1}^{r} \sum_{j=1}^{m_\ell} \binom{n+j-1}{j-1} \frac{w_{\ell j}}{(-v_\ell)^j} (v_\ell)^{-n},$$

where the $w_{\ell j}$'s are the entries of the last column of V^{-1}.

Proof. The generating function of the sequence $\{g_D(n)\}_n$ is given by $D(0)/D(x)$. Then, by Eq. (4) we have

$$\sum_{n=0}^{\infty} g_D(n) x^n = \frac{D(0)}{D(x)} = D(0) \sum_{\ell=1}^{r} \sum_{j=1}^{m_\ell} \frac{w_{\ell j}}{(x - v_\ell)^j}$$

and the result follows by applying

$$\frac{1}{(x - v_\ell)^j} = \frac{1}{(-v_\ell)^j} \cdot \frac{1}{(1 - x/v_\ell)^j} = \frac{1}{(-v_\ell)^j} \sum_{n=0}^{\infty} \binom{n+j-1}{j-1} (v_\ell)^{-n} x^n. \qquad \square$$

4. Vandermonde distributions

In this section we study the properties of a family of density functions naturally associated with confluent Vandermonde matrices. Let r, k, v_ℓ, m_ℓ and $D(x)$ be defined as in Section 3. Consider the confluent Vandermonde matrix V associated with $D(x)$ and the entries $w_{\ell j}$'s of the last column of V^{-1}, for $\ell = 1, 2, \ldots, r$ and $j = 1, 2, \ldots m_\ell$. Now, assume that all v_ℓ's are real and satisfy the relation $0 \leqslant v_1 < v_2 < \cdots < v_r$. Then, we define the real function

$$f_D(x) = \begin{cases} 0 & \text{if } x < v_1, \\ (k-1) \sum_{\ell=h}^{r} \sum_{j=1}^{m_\ell} \binom{k-2}{j-1} w_{\ell j} (v_\ell - x)^{k-j-1} & \text{if } v_{h-1} \leqslant x < v_h, \text{ for some } 1 < h \leqslant r, \\ 0 & \text{if } x \geqslant v_r. \end{cases}$$

Its features mainly depend on the properties presented in Proposition 1. In particular notice that, for any $h \in \{2, 3, \ldots r\}$, if $v_{h-1} \leqslant x < v_h$, then we obtain

$$f_D(x) = -(k-1) \sum_{\ell=1}^{h-1} \sum_{j=1}^{m_\ell} \binom{k-2}{j-1} w_{\ell j} (v_\ell - x)^{k-j-1}. \tag{8}$$

Clearly, f_D is continuously differentiable till the order $k-2$ in $\mathbb{R} \backslash \{v_1, \ldots, v_r\}$ and its $(k-2)$th derivative is constant in each interval $(v_\ell, v_{\ell+1})$, $\ell = 1, \ldots, k-1$. Moreover, using Proposition 1, one can verify that, for any $\ell = 1, \ldots, k-1$, the function f_D is continuous at v_ℓ if and only if $m_\ell \leqslant k-2$ (note that this condition is true whenever $r \geqslant 3$). In general, f_D is continuously differentiable at v_ℓ till the order $k - m_\ell - 2$.

4.1. Unimodal property

Here, we prove that the function $f_D(x)$ defined above is nonnegative all over $\mathbb{R}$ and that, if $k \geqslant 3$, then f_D is *unimodal* in (v_1, v_r), that is there exists $t \in [v_1, v_r]$ such that $f_D(x)$ is strictly increasing in $[v_1, t]$ and strictly decreasing in $[t, v_r]$. Note that, if $f_D(x)$ is continuous (all over $\mathbb{R}$) then its unimodality in (v_1, v_r) implies the existence of a unique (local) maximum in (v_1, v_r).

To prove these properties, we consider two different cases: $k = r$ or $k > r$. If $k = r$, that is $m_\ell = 1$ for every $\ell = 1, 2, \ldots, r$, then f_D reduces to

$$f_D(x) = \begin{cases} 0 & \text{if } x < v_1, \\ (k-1)\sum_{j=\ell}^{k} c_j (v_j - x)^{k-2} & \text{if } v_{\ell-1} \leqslant x < v_\ell \text{ for some } 1 < \ell \leqslant k, \\ 0 & \text{if } x \geqslant v_k, \end{cases}$$

where $c_\ell = \prod_{i \neq \ell} (v_\ell - v_i)^{-1}$ for any $\ell = 1, 2, \ldots, r$. Note that if $k = 2$ then f_D is the uniform density function over the interval (v_1, v_2), while for $k = 3$ we have the triangular distribution. The next proposition shows that, if $k = r \geqslant 3$, then f_D is a unimodal function. The proof is based on Corollary 2 and makes use of the following lemma.

Lemma 4. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function admitting jth derivative all over $\mathbb{R}$ for some $j \geqslant 1$. Also assume that, for some real values $a < b$, f has m zeros in the interval (a, b) while $f(x) = 0$ for each $x \leqslant a$ and each $x \geqslant b$. Then, for every $i = 1, \ldots, j$, the ith derivative of f admits at least $m + i$ zeros in (a, b).*

Proof. We reason by induction on $i = 1, \ldots, j$. If $i = 1$, then consider the $m + 1$ intervals determined by the zeros of f in $[a, b]$. For each of them, say (x_1, x_2), Rolle's Theorem guarantees that $f'(x) = 0$ for some $x \in (x_1, x_2)$.

Now, assume $1 < i < j$ and consider the ith derivative of f, i.e. the function $g = f^{(i)}$. By the properties of f, we have $g(a) = g(b) = 0$ and by the inductive hypotheses g admits $m + i$ zeros in (a, b). Therefore, by applying the previous argument to g, one proves that $g' = f^{(i+1)}$ admits $m + i + 1$ zeros in (a, b). □

Proposition 5. *If $k = r \geqslant 3$, then f_D is unimodal in (v_1, v_k) and is nonnegative all over $\mathbb{R}$.*

Proof. Using Corollary 2, one can prove that f_D is strictly increasing in (v_1, v_2) and strictly decreasing in (v_{k-1}, v_k). In particular, this implies the property for $k = 3$.

Now, let $k \geqslant 4$. Then, f_D is continuously differentiable till the order $k - 3$. Assume by contradiction that f_D is not unimodal. Since the derivative f'_D is positive in (v_1, v_2) and negative in (v_k, v_{k-1}), this implies that f'_D necessarily vanishes in at least 3 points in the interval $[v_2, v_{k-1}]$. For $k = 4$ this leads to a contradiction because f'_D is linear in $[v_2, v_3]$. For $k > 4$, the function f'_D satisfies the hypotheses of Lemma 4 with $j = k - 4$, $m = 3$, $a = v_1$, $b = v_r$. As a consequence, the $(k - 3)$th derivative ${f_D}^{(k-3)}$ of f_D admits at least $k - 1$ zeros in (v_1, v_k), and this again leads to a contradiction. Indeed, ${f_D}^{(k-3)}(x)$ is continuous all over $\mathbb{R}$, it is nonnull for $x \in (v_1, v_2) \cup (v_{k-1}, v_k)$, it is linear with respect to x in each of the $k - 3$ intervals $(v_\ell, v_{\ell+1})$, $\ell = 2, \ldots k - 2$, and hence it has at most $k - 3$ many zeros in (v_1, v_k).

Finally, since f_D is positive in $(v_1, v_2) \cup (v_{k-1}, v_k)$ and admits a unique local maximum in (v_1, v_k), then we can conclude that $f_D(x) \geqslant 0$ for every real x. □

To prove that f_D is unimodal also when $k > r$, we use the following lemma, which can be easily proved reasoning by contradiction [12].

Lemma 6. *For every $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be a continuous function that admits a unique local maximum. If $\{f_n\}$ pointwise converges to a continuous function $f : \mathbb{R} \to \mathbb{R}$, then f admits a unique local maximum, too.*

We are now able to prove the complete property.

Proposition 7. *If $k \geqslant 3$, then f_D is nonnegative all over $\mathbb{R}$ and unimodal in (v_1, v_r).*

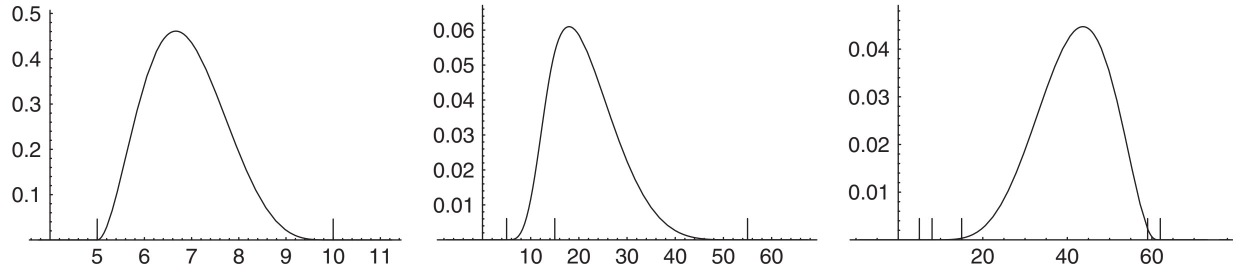

Fig. 1. Plots of function f_D for $v = (5, 10)$ and $m = (5, 3)$, $v = (5, 15, 55)$ and $m = (3, 3, 2)$, $v = (5, 8, 15, 59, 62)$ and $m = (1, 1, 1, 2, 3)$, respectively. The vertical bars indicate the values of v_j's.

Proof. If $r = 2 < k$ then, by definition of f_D and Eq. (6), one can show that

$$f_D(x) = \frac{(k-1)!}{(m_1-1)!(m_2-1)!} \frac{(v_2-x)^{m_1-1}(x-v_1)^{m_2-1}}{(v_2-v_1)^{k-1}}$$

for every $v_1 < x < v_2$. It is easy to verify that f_D is nonnegative and unimodal in (v_1, v_2). Also note that f_D is continuous unless $m_1 = 1$ or $m_2 = 1$ (and in these cases the only discontinuity point is $x = v_2$ or $x = v_1$, respectively).

If $r \geqslant 3$, then f_D is continuous and we reason by induction on the integer $k - r$. If $k - r = 0$, then the property is true by Proposition 5. Thus, consider the case $k - r > 0$. Then, there exists $\ell \in \{1, 2, \ldots, r\}$ such that $m_\ell > 1$. Recalling Eq. (8), we may assume $\ell = 1$ without loss of generality. Given $0 < \varepsilon < v_1$, set $v_0 = v_1 - \varepsilon$ and $m_0 = 1$ (if $v_1 = 0$, a similar result can be obtained by setting $v_0 = \varepsilon$ for any $0 < \varepsilon < v_2$). Now, consider the polynomial

$$D_\varepsilon(x) = \frac{1}{x - v_1} \prod_{\ell=0}^{r} (x - v_\ell)^{m_\ell}$$

and note that D_ε has $r+1$ distinct roots $v_0 < v_1 < \cdots < v_r$ with multiplicities such that $1+m_1-1+m_2+\cdots+m_r = k$. Thus, also f_{D_ε} is continuous and, by the inductive hypothesis, we know that f_{D_ε} is nonnegative in $\mathbb{R}$ and unimodal in (v_1, v_r).

Let us study the pointwise convergence of $f_{D_\varepsilon}(x)$ as ε goes to zero. If $x \geqslant v_r$, then $f_D(x) = f_{D_\varepsilon}(x) = 0$. If $x < v_1$ then, for ε small enough, $x < v_0$ and hence $f_{D_\varepsilon}(x) = 0 = f_D(x)$. Finally, for any $h \geqslant 2$ and $v_{h-1} \leqslant x \leqslant v_h$ we have

$$f_{D_\varepsilon}(x) = \sum_{\ell=h}^{r} \sum_{j=1}^{m_\ell} \binom{k-2}{j-1} (v_\ell - x)^{k-j-1} w_{\ell j}(\varepsilon),$$

where we use $w_{\ell j}(\varepsilon)$ to denote the entries of the last column of V_ε^{-1}, V_ε being the confluent Vandermonde matrix associated with $D_\varepsilon(x)$. Using Eq. (5), one can easily verify that $\lim_{\varepsilon \to 0} w_{\ell j}(\varepsilon) = w_{\ell j}$ for every $\ell \geqslant 2$. Thus, f_{D_ε} pointwise converges to f_D all over $\mathbb{R}$, and the result follows by applying Lemma 6. $\square$

In Fig. 1, we show the plots of functions f_D's for three polynomials D, which present a rather regular behaviour. In these examples the number of distinct roots of D (i.e. the value of r) is 2, 3 and 5, respectively, and for each of them $f_D(x)$ is differentiable all over $\mathbb{R}$.

In Fig. 2 four special examples are illustrated which present irregular behaviours. In the first case there are two distinct roots of D with multiplicity 4 and 1, respectively, and the corresponding function f_D is not continuous at the first root. In the other three examples D has three distinct roots with different sets of multiplicities: in the case of simple roots f_D is a triangular density function (see the second picture); if the array of multiplicities is $m = (1, 4, 1)$, then f_D is continuous but not differentiable at the second root, while the same behaviour occurs at the first root if $m = (4, 1, 1)$.

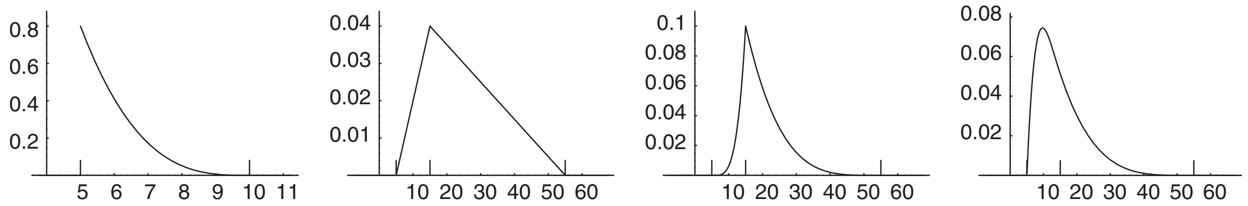

Fig. 2. Plots of function f_D for $v = (5, 10)$ and $m = (4, 1)$, $v = (5, 15, 55)$ and $m = (1, 1, 1)$, $v = (5, 15, 55)$ and $m = (1, 4, 1)$, $v = (5, 15, 55)$ and $m = (4, 1, 1)$, respectively. The vertical bars indicate the values of v_j's.

4.2. Characteristic function

Here, we prove that f_D is a density function. Since it is nonnegative all over $\mathbb{R}$, it is sufficient to prove that $\int_{-\infty}^{+\infty} f_D(x)\,\mathrm{d}x = 1$. Further, we show that the characteristic function of f_D is given by

$$\Phi_D(t) = \frac{(k-1)!}{(\mathrm{i}t)^{k-1}} \sum_{\ell=1}^{r} \mathrm{e}^{\mathrm{i}tv_\ell} \sum_{j=1}^{m_\ell} \frac{w_{\ell j}}{(j-1)!} (\mathrm{i}t)^{j-1}. \tag{9}$$

We will say that a random variable of density function $f_D(x)$ is a *Vandermonde random variable of parameter $D(x)$*.

Proposition 8. *Let D be a monic polynomial with at least two distinct roots and assume that all roots are nonnegative real. Then, the map f_D is a density function having characteristic function $\Phi_D(t)$.*

Proof. We first show that $\Phi_D(t) = \int_{-\infty}^{+\infty} f_D(x)\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}x$. Set $I(t) = \int_{-\infty}^{\infty} f_D(x)\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}x$ and observe that

$$I(t) = (k-1) \sum_{h=2}^{r} \sum_{\ell=h}^{r} \sum_{j=1}^{m_\ell} \binom{k-2}{j-1} w_{\ell j} \int_{v_{\ell-1}}^{v_\ell} (v_\ell - x)^{k-j-1} \mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}x.$$

Integrating by parts one can verify that for $t \neq 0$ the function $\mathrm{e}^{\mathrm{i}tx}(c-x)^p$ admits the antiderivative

$$\frac{\mathrm{e}^{\mathrm{i}tx}}{\mathrm{i}t} \sum_{s=0}^{p} \frac{p!(c-x)^{p-s}}{(p-s)!(\mathrm{i}t)^s}.$$

Hence we can write $I(t) = \sum_{h=2}^{r} \sum_{\ell=h}^{r} (A_{\ell,h} - A_{\ell,h-1})$ where

$$A_{\ell,h} = \mathrm{e}^{\mathrm{i}tv_h} \sum_{j=1}^{m_\ell} \frac{(k-1)!}{(j-1)!} w_{\ell j} \sum_{s=0}^{k-j-1} \frac{(v_\ell - v_h)^{k-j-1-s}}{(k-j-1-s)!(\mathrm{i}t)^{s+1}}$$

and in particular

$$A_{\ell,\ell} = \mathrm{e}^{\mathrm{i}tv_\ell} \sum_{j=1}^{m_\ell} \frac{(k-1)!}{(j-1)!} \cdot \frac{w_{\ell j}}{(\mathrm{i}t)^{k-j}}.$$

Now, set $B_h = \sum_{\ell=h}^{r} A_{\ell,h}$ and $C_h = \sum_{\ell=h}^{r} A_{\ell,h-1}$. For each $2 \leqslant h \leqslant r-1$ we have $B_h - C_{h+1} = A_{h,h}$ and moreover $B_r = A_{r,r}$. Finally, reasoning as in Proposition 1 one can prove that $C_2 = \sum_{\ell=1}^{r} A_{\ell,1} - A_{1,1} = -A_{1,1}$. As a consequence, the integral can be computed as follows

$$\int_{-\infty}^{\infty} f_D(x)\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}x = \sum_{h=2}^{r} (B_h - C_h) = \sum_{\ell=1}^{r} A_{\ell,\ell} = \sum_{\ell=1}^{r} \mathrm{e}^{\mathrm{i}tv_\ell} \sum_{j=1}^{m_\ell} \frac{(k-1)!}{(j-1)!} \cdot \frac{w_{\ell j}}{(\mathrm{i}t)^{k-j}} = \Phi_D(t).$$

The proposition is then proved if we show that $\lim_{t\to 0} \Phi_D(t) = 1$. By expanding e^{itv_ℓ}, we get

$$\Phi_D(t) = \frac{(k-1)!}{(\mathrm{i}t)^{k-1}} \sum_{\ell=1}^{r} \left(\sum_{j=0}^{\infty} \frac{v_\ell^j}{j!} (\mathrm{i}t)^j \right) \cdot \left(\sum_{j=1}^{m_\ell} \frac{w_{\ell j}}{(j-1)!} (\mathrm{i}t)^{j-1} \right)$$

$$= (k-1)! \sum_{s=1}^{\infty} \left(\sum_{\ell=1}^{r} \sum_{j=1}^{\min(s,m_\ell)} \frac{v_\ell^{s-j}}{(s-j)!} \cdot \frac{w_{\ell j}}{(j-1)!} \right) (\mathrm{i}t)^{s-k}.$$

By Proposition 1, the first non-null coefficient in the previous sum is obtained for $s = k$ and equals $1/(k-1)!$. This concludes the proof. □

5. Rational models for pattern statistics

We now turn our attention to pattern statistics. Here, we recall the definition and the main properties of the rational stochastic models introduced in [2], based on the classical notion of rational formal series [19,1].

Let $\mathbb{R}_+$ be the semiring of nonnegative real numbers and consider the finite alphabet Σ. A formal series over Σ with coefficients in $\mathbb{R}_+$ is a function $r : \Sigma^* \longrightarrow \mathbb{R}_+$, usually represented in the form $r = \sum_{\omega\in\Sigma^*} (r, \omega) \cdot \omega$, where (r, ω) denotes the value of r at $\omega \in \Sigma^*$. Moreover, r is called *rational* if it admits a *linear representation*, that is a triple (ξ, μ, η) where, for some integer $m > 0$, ξ and η are (column) vectors in $\mathbb{R}_+^m$ and $\mu : \Sigma^* \longrightarrow \mathbb{R}_+^{m\times m}$ is a monoid morphism, such that $(r, \omega) = \xi^{\mathrm{T}}\mu(\omega)\eta$ holds for each $\omega \in \sigma^*$. Observe that considering such a triple (ξ, μ, η) is equivalent to defining a (weighted) nondeterministic automaton, where the set of states is given by $\{1, 2, \dots, m\}$ and the transitions, the initial and the final states are assigned weights in $\mathbb{R}_+$ by μ, ξ and η, respectively. To avoid redundancy it is convenient to assume that (ξ, μ, η) is trim (meaning that all indices are used to define the series), i.e. for every index i there are two indices p, q and two words $x, y \in \Sigma^*$ such that $\xi_p\mu(x)_{pi} \neq 0$ and $\mu(y)_{iq}\eta_q \neq 0$. We say that (ξ, μ, η) is *primitive* if $\mathcal{M} = \sum_{\sigma\in\Sigma}\mu(\sigma)$ is a primitive matrix, that is for some $n \in \mathbb{N}$ all entries of $\mathcal{M}^n$ are strictly positive. We also recall that a matrix $\mathcal{M} \in \mathbb{R}_+^{m\times m}$ is called *irreducible* if for every pair of indices p, q there exists $n \in \mathbb{N}$ such that $\mathcal{M}^n_{pq} > 0$.

Any formal series can define a stochastic model for studying the frequency of occurrences of a letter in a word of given length. Consider the binary alphabet $\{a, b\}$ and, for any $n \in \mathbb{N}$, let $\{a, b\}^n$ denote the set of all words of length n in $\{a, b\}^*$. Consider a formal series $r : \{a, b\}^* \longrightarrow \mathbb{R}_+$ and let n be a positive integer such that $(r, x) \neq 0$ for some $x \in \{a, b\}^n$. A probability measure over $\{a, b\}^n$ can be defined by setting

$$\Pr\{\omega\} = \frac{(r, \omega)}{\sum_{x\in\{a,b\}^n} (r, x)} \quad (\omega \in \{a, b\}^n). \tag{10}$$

In particular, if r is the characteristic series χ_L of a language $L \subseteq \{a, b\}^*$, then Pr is just the uniform probability function over $L \cap \{a, b\}^n$. Then, we define the random variable (r.v. for short) $Y_n : \{a, b\}^n \to \{0, 1, \dots, n\}$ such that $Y_n(\omega) = |\omega|_a$ for every $\omega \in \{a, b\}^n$. For every $j = 0, 1, \dots, n$, we have

$$\Pr\{Y_n = j\} = \frac{\sum_{|\omega|=n, |\omega|_a=j} (r, \omega)}{\sum_{x\in\{a,b\}^n} (r, x)}. \tag{11}$$

If $r = \chi_L$ for some $L \subseteq \{a, b\}^*$, then Y_n represents the number of occurrences of a in a word chosen at random in $L\cap\{a, b\}^n$ under uniform distribution. We observe that, in this case, our results concerning Y_n are related to the analysis of additive functions over strings [9].

When r is rational, the r.v. Y_n defines a model for the study of pattern statistics we call *rational* stochastic model. This is extension of the traditional Markovian models in the following sense. Given a regular set of patterns on an arbitrary finite alphabet Σ consider a Markovian source P generating words at random over Σ and let X_n be the r.v. representing the number of occurrences of patterns in a word of length n generated by P; then, there exists a rational formal series $r : \{a, b\}^* \longrightarrow \mathbb{R}_+$ such that, for every $n \geqslant 1$, the corresponding r.v. Y_n has the same distribution as X_n [2, Section 2.1].

Let (ξ, μ, η) be a linear representation for the rational series r and set $\mathcal{A} = \mu(a)$, $\mathcal{B} = \mu(b)$, $\mathcal{M} = \mathcal{A} + \mathcal{B}$. To study the behaviour of the random variables Y_n and in particular their limit distribution, it is useful to introduce the sequence of functions $\{r_n(z)\}_n$ in the complex variable z defined by

$$r_n(z) = \sum_{x \in \{a,b\}^n} (r, x)\, \mathrm{e}^{z|x|_a} = \xi^{\mathrm{T}} (\mathcal{A}e^z + \mathcal{B})^n \eta. \tag{12}$$

Indeed, it is immediate to see that the characteristic function of Y_n satisfies the relation

$$\Psi_{Y_n}(t) = \mathbb{E}(\mathrm{e}^{\mathrm{i}tY_n}) = \frac{r_n(\mathrm{i}t)}{r_n(0)} \tag{13}$$

for $t \in \mathbb{R}$. We recall that a sequence of random variables X_n converges in distribution to a random variable X if and only if the sequence of characteristic functions $\Psi_{X_n}(t)$ pointwise converges to $\Psi_X(t)$ [7].

Now, consider the generating function of $\{r_n(z)\}_n$ and observe that

$$\sum_{n=0}^{\infty} r_n(z) w^n = \xi^{\mathrm{T}} H(z, w) \eta,$$

where $H(z, w)$ is the matrix defined by

$$H(z, w) = \sum_{n=0}^{\infty} (\mathcal{A}e^z + \mathcal{B})^n w^n = (I - w(\mathcal{A}e^z + \mathcal{B}))^{-1}. \tag{14}$$

If $\mathcal{M}$ is irreducible, by Perron–Frobenius Theorem (see [21, Theorem 1.5]) it has a nonnegative real eigenvalue λ of maximum modulus. Moreover, we know that the equation $\mathrm{Det}(yI - \mathcal{A}e^z - \mathcal{B}) = 0$ defines an implicit function $y = y(z)$ which is analytic in a neighbourhood of $z = 0$ and such that $y(0) = \lambda$.

If further $\mathcal{M}$ is primitive and $\mathcal{A} \neq 0 \neq \mathcal{B}$, then there are two constants $\beta \in (0, 1)$, $\gamma > 0$, both depending on the matrix $\mathcal{M}$ and its eigenvectors (see [2] for details), such that, as n tends to infinity, the following relations hold:

$$\mathbb{E}(Y_n) = \beta n + \mathrm{O}(1), \quad \mathbb{V}ar(Y_n) = \gamma n + \mathrm{O}(1). \tag{15}$$

Finally, under the same hypothesis, one can prove that the distribution of $(Y_n - \beta n)/\sqrt{\gamma n}$ converges to the normal distribution of mean value 0 and variance 1 [2].

In our investigation we often deal with matrices of functions. We will say that a matrix $A(w)$ is a *matrix function* if all its entries are functions of the variable w. We will also say that $A(w)$ is analytic at a point $w = a$ if all its entries are analytic at the same point; moreover, its radius of convergence at that point is the smallest radius of convergence of the power series development of its entries (with center in a).

5.1. Decomposition of a rational model

Up to now, the properties of Y_n have been studied only in the primitive models [2] and in the case of two primitive components [5]. Here, we present a general approach to deal with an arbitrary rational model. To this aim, we describe the construction of the reduced graph of the strongly connected components of the corresponding linear representation. This is a usual approach in the analysis of counting problems on regular languages (see for instance [8] for an application concerning trace languages).

Let (ξ, μ, η) be a linear representation over the alphabet $\{a, b\}$ with coefficients in $\mathbb{R}_+$. As in the previous section, set $\mathcal{A} = \mu(a)$, $\mathcal{B} = \mu(b)$, $\mathcal{M} = \mathcal{A} + \mathcal{B}$ and consider the directed graph defined by $\mathcal{M}$, where the set of nodes is $\{1, 2, \ldots, m\}$ and (p, q) is an (oriented) edge if and only if $\mathcal{M}_{pq} \neq 0$. Then, let $C_1, C_2, \ldots, C_s$ be the strongly connected components of the graph and define C_i *initial* (resp. *final*) if $\xi_p \neq 0$ (resp. $\eta_p \neq 0$) for some $p \in C_i$. The *reduced graph* of (ξ, μ, η) is then defined as the directed acyclic graph G where $C_1, C_2, \ldots, C_s$ are the vertices and any pair (C_i, C_j) is an edge if and only if $i \neq j$ and $\mathcal{M}_{pq} \neq 0$ for some $p \in C_i$ and some $q \in C_j$.

Up to a permutation of indices, the matrix $\mathcal{M}$ can be represented as a triangular block matrix of the form

$$\mathcal{M} = \begin{pmatrix} M_1 & M_{12} & M_{13} & \cdots & M_{1s} \\ 0 & M_2 & M_{23} & \cdots & M_{2s} \\ & & \cdots & & \\ 0 & 0 & 0 & \cdots & M_s \end{pmatrix}, \tag{16}$$

where each M_i corresponds to the strongly connected component C_i and every M_{ij} corresponds to the transitions from vertices of C_i to vertices of C_j in the original graph of $\mathcal{M}$. Also $\mathcal{A}$, $\mathcal{B}$, ξ and η admit similar decompositions: we define the matrices A_i, A_{ij}, B_i, B_{ij} and the vectors ξ_i, η_i in the corresponding way and we say that the component C_i is *degenerate* if $A_i = 0$ or $B_i = 0$. Since each matrix M_i is either null or irreducible, by Perron–Frobenius Theorem it has a nonnegative real eigenvalue λ_i of maximum modulus. We call *main eigenvalue* of $\mathcal{M}$ the value $\lambda = \max\{\lambda_i | i = 1, 2, \ldots, s\}$ and we say that C_i is a *dominant component* if $\lambda_i = \lambda$. Observe that $\lambda_i = 0$ only if C_i reduces to a loopless single node and hence from now on we assume $\lambda > 0$.

Further, if a matrix M_i is primitive, we say that C_i is a *primitive component*. In this case, when C_i is not degenerate (i.e. $A_i \neq 0 \neq B_i$), we may consider the constants β_i and γ_i associated with M_i defined as in (15); we have $0 < \beta_i < 1$ and $\gamma_i > 0$. On the contrary, if C_i is degenerate, it is natural to set $\gamma_i = 0$ and define $\beta_i = 0$ or $\beta_i = 1$ according whether $A_i = 0$ or $B_i = 0$ (so that (15) still holds true for a degenerate r.v.). Thus the constants β_i and γ_i are well-defined for every primitive component C_i: we say they are the *mean constant* and the *variance constant* of C_i, respectively.

The block decomposition of $\mathcal{M}$ also induces a decomposition of the matrix $H(z, w)$ defined in (14). More precisely, the blocks under the diagonal are all null, while the upper triangular part is composed by a family of matrices, say $H_{ij}(z, w)$, $1 \leqslant i \leqslant j \leqslant s$. Note that the bivariate generating function $\xi^{\mathrm{T}} H(z, w)\eta$, which is the main tool of our investigation, is now given by

$$\xi^{\mathrm{T}} H(z, w)\eta = \sum_{n=0}^{\infty} \xi^{\mathrm{T}} (\mathcal{A}e^z + \mathcal{B})^n \eta \cdot w^n = \sum_{1 \leqslant i \leqslant j \leqslant s} \xi_i^{\mathrm{T}} H_{ij}(z, w)\eta_j. \tag{17}$$

Setting $M_{ij}(z) = A_{ij}e^z + B_{ij}$ and reasoning by induction on $j - i$, one can prove that, for each $1 \leqslant i \leqslant j \leqslant s$,

$$H_{ij}(z, w) = \begin{cases} (I - w(A_i e^z + B_i))^{-1} & \text{if } j = i, \\ \sum_* H_{i_1 i_1}(z, w) M_{i_1 i_2}(z) H_{i_2 i_2}(z, w) \cdots M_{i_{\ell-1} i_\ell}(z) H_{i_\ell i_\ell}(z, w) \cdot w^{\ell-1} & \text{if } j \neq i, \end{cases} \tag{18}$$

where the sum $(*)$ is extended over all sequences of integers $(i_1, i_2, \ldots, i_\ell)$, $\ell \geqslant 2$ such that $i_1 = i$, $i_t < i_{t+1}$ for each $t = 1, \ldots, \ell - 1$ and $i_\ell = j$.

Eq. (18) suggests us to introduce the notion of chain of the reduced graph G associated with (ξ, μ, η). A *chain* is a simple path in G, i.e. a sequence of distinct components $\kappa = (C_{i_1}, C_{i_2}, \ldots, C_{i_\ell})$ where $\ell \geqslant 1$, such that $M_{i_j i_{j+1}} \neq 0$ for every $j = 1, 2, \ldots, \ell - 1$. We say that ℓ is the *length* of κ while the *order* of κ is the number of its dominant components. We also denote by Γ the family of all chains in G starting with an initial component and ending with a final component. Note that, the linear representation (ξ, μ, η) being trim, each component lies over at least one chain in Γ. We say that a chain κ is a *main chain* if $\kappa \in \Gamma$ and its order is maximal in Γ. We denote by Γ_m the set of all main chains in G.

Example 1. Here, we present a parametric example depending on two constants $\alpha \in \mathbb{R}$ and $\rho \in [0, 1]$. For such parameters let us consider the formal series $s_{\alpha,\rho}$ having linear representation (π, ν, τ) such that

$$\pi' = (1, 0), \quad \nu_{\alpha,\rho}(a) = \alpha \cdot \begin{pmatrix} 0 & 1 \\ \rho & 0 \end{pmatrix}, \quad \nu_{\alpha,\rho}(b) = \alpha \cdot \begin{pmatrix} 1 & 0 \\ 1-\rho & 0 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We note that the Perron–Frobenius eigenvalue associated with this linear representation is $\alpha(1 + \sqrt{5})/2$.

Given two families of parameters $\{\alpha_i\}_{i=1,m}$ and $\{\rho_i\}_{i=1,m}$, let us define the formal series r given by the Cauchy product $r = \prod_{i=1}^{m} s_{\alpha_i,\rho_i}$. Its linear representation of size $2m$ is given by (ξ, μ, η) where $\xi_1 = 1$ and $\xi_i = 0$

for every $i \neq 1$, $\eta_{2i} = 0$ and $\eta_{2i+1} = 1$ for every i, while for $x \in \{a, b\}$ μ is defined by

$$\mu(x) = \begin{pmatrix} v_{\alpha_1,\rho_1}(x) & M_{12}(x) & M_{13}(x) & \cdots & M_{1m}(x) \\ 0 & v_{\alpha_2,\rho_2}(x) & M_{23}(x) & \cdots & M_{2m}(x) \\ & & \cdots & & \\ 0 & 0 & 0 & \cdots & v_{\alpha_m,\rho_m}(x) \end{pmatrix} \tag{19}$$

with

$$M_{ij}(a) = \begin{pmatrix} \alpha_j & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad M_{ij}(b) = \begin{pmatrix} 0 & \alpha_j \\ 0 & 0 \end{pmatrix}.$$

Here, the entries $1, 2, \ldots, 2m$ can be gathered in strongly connected components, defined by the sets $C_i = \{2i-1, 2i\}$ for $i = 1, 2, \cdots, m$. Thus, the reduced graph of (ξ, μ, η) consists of nodes C_i's and edges (C_i, C_j) with $i < j$. The component C_1 is initial while all C_j's are final. The orders of the chains depend on the values α_i's. In particular, when all α_i's are equal, we have only one main chain of order m. If the α_i's are different, there may be several main chains. For instance, if $m = 4$, $\alpha_2 = \alpha_3$ and $\alpha_1 = \alpha_4 = 2\alpha_2$, only C_1 and C_4 are dominant; therefore we have four main chains, namely (C_1, C_2, C_3, C_4), (C_1, C_2, C_4), (C_1, C_3, C_4) and (C_1, C_4). □

5.2. *The role of main chains*

In this section we study the properties of main chains and in particular we show that they determine the limit distribution of the sequence $\{Y_n\}$ associated with the linear representation (ξ, μ, η). Intuitively, this is a consequence of two facts. First, the characteristic function of (a normalization of) Y_n depends on the sequences $\{r_n(z)\}$ for z near 0, and hence on the generating function $\xi^{\mathrm{T}} H(z, w)\eta$. Second, by (17), this function is a sum of products of the form given in (18), each of which is identified by a chain: the products corresponding to the main chains have singularities of smallest modulus with the largest degree, and hence they yield the main asymptotic contribution to the associated sequence $\{r_n(z)\}$.

So, let us take in exam the terms of the sum in the right hand side of (17). First we consider the case $i = j$ and, for every $j = 1, 2, \ldots, s$, we denote $\{r_n^{(j)}(z)\}$ the sequence given by

$$\xi_j^{\mathrm{T}} H_{jj}(z, w)\eta_j = \sum_{n=0}^{\infty} r_n^{(j)}(z) w^n.$$

By relation (18), we have

$$H_{jj}(z, w) = (I - w(A_j e^z + B_j))^{-1} = \frac{\mathrm{Adj}(I - w(A_j e^z + B_j))}{\det(I - w(A_j e^z + B_j))}$$

and hence, as z tends to 0, the singularities of each entry approach the inverses of eigenvalues of M_j. We can distinguish three cases according to the properties of M_j:

(i) M_j is primitive and dominant. Then, λ is its (sole) eigenvalue of largest modulus. The equation $\det(yI - (A_j e^z + B_j)) = 0$ implicitly defines a function $y = y_j(z)$ in a neighbourhood of $z = 0$ such that $y_j(0) = \lambda$. Such function is analytic at the point $z = 0$ and admits an expansion of the form

$$y_j(z) = \lambda\left(1 + \beta_j z + \frac{\gamma_j + \beta_j^2}{2} z^2 + \mathrm{O}(z^3)\right), \tag{20}$$

where β_j and γ_j are the mean and variance constants of C_j. Note that this equation is well-defined also when C_i is degenerate (in particular, if $A_i = 0$ then $y_i(z) = \lambda$ for all z).

Then, there exists a matrix function $R_j(z)$ analytic and nonnull at $z = 0$ such that, for every z near 0,

$$H_{jj}(z, w) - \frac{R_j(z)}{1 - y_j(z)w}$$

has a radius of convergence strictly greater than λ^{-1}. As a consequence we have

$$r_n^{(j)}(z) = \xi_j^{\mathrm{T}} R_j(z)\eta_j (y_j(z))^n + \mathrm{O}(\rho^n)$$

for some $0 < \rho < \lambda$ and every z near 0.

(ii) M_j is dominant (but not necessarily primitive). Then, we can consider the family E_j of the eigenvalues of M_j of largest modulus. By Perron–Frobenius Theorem, we know $\lambda \in E_j$ and, for every $\alpha \in E_j$, the equation $\det(yI - (A_j e^z + B_j)) = 0$ implicitly defines a function $y = y_\alpha(z)$ in a neighbourhood of $z = 0$ such that $y_\alpha(0) = \alpha$. Also $y_\alpha(z)$ is analytic at $z = 0$, where it admits an expansion of the form

$$y_\alpha(z) = \alpha(1 + m_\alpha z + s_\alpha z^2 + \mathrm{O}(z^3)) \tag{21}$$

with $m_\alpha \in \mathbb{R}_+$ and $\Re(s_\alpha) \geqslant 2m_\alpha^2$ (consequence of point (e) in [21, Theorem 1.5]). Reasoning as above this implies, for z near 0 and some $0 < \rho < \lambda$,

$$r_n^{(j)}(z) = \sum_{\alpha \in E_j} \xi_j^{\mathrm{T}} R_\alpha(z)\eta_j \ (y_\alpha(z))^n + \mathrm{O}(\rho^n)$$

where $R_\alpha(z)$ is a matrix function analytic and nonnull at $z = 0$, for each $\alpha \in E_j$.

(iii) M_j is not dominant. Then, all its eigenvalues are in modulus smaller than λ and hence, as z is near to 0 the radius of convergence of $H_{jj}(z, w)$ is greater than λ^{-1}. This implies $r_n^{(j)}(z) = \mathrm{O}(\rho^n)$ for some $0 < \rho < \lambda$ and all z near 0.

Now, let us study the behaviour of $H_{ij}(z, w)$ for $i \neq j$. Recalling (18), we consider an arbitrary chain $\kappa = (C_{i_1}, C_{i_2}, \ldots, C_{i_\ell})$ with $\ell \geqslant 2$ and we denote by $H_\kappa(z, w)$ the corresponding matrix given by

$$H_\kappa(z, w) = H_{i_1 i_1}(z, w) M_{i_1 i_2}(z) H_{i_2 i_2}(z, w) \cdots M_{i_{\ell-1} i_\ell}(z) H_{i_\ell i_\ell}(z, w) \cdot w^{\ell-1}. \tag{22}$$

We also define the sequence $\{r_n^{(\kappa)}(z)\}$ by

$$\xi_{i_1}^{\mathrm{T}} H_\kappa(z, w)\eta_{i_\ell} = \sum_{n=0}^{\infty} r_n^{(\kappa)}(z) w^n. \tag{23}$$

Then, the next proposition can be proved by applying the previous properties to (22).

Proposition 9. *Let κ be a chain in Γ of order $k \geqslant 0$. Then, as n tends to $+\infty$, the following statements hold for every $c \in \mathbb{C}$ and every $t \in \mathbb{R}$:*

(1) *If $k = 0$ then $r_n{}^{(\kappa)}(c/n) = \mathrm{O}(\rho^n)$ for some $0 < \rho < \lambda$;*

(2) *If $k \geqslant 1$ then $r_n{}^{(\kappa)}(c/n) = \mathrm{O}(\lambda^n n^{k-1})$;*

(3) *If $k \geqslant 1$ and the dominant components of κ are primitive, then $r_n{}^{(\kappa)}(c/n) = \Theta(\lambda^n n^{k-1})$;*([2])

(4) *If $k \geqslant 1$ then $r_n{}^{(\kappa)}(\mathrm{i}t/\sqrt{n}) = \mathrm{O}(\lambda^n n^{k-1})$.*

Proof. Without loss of generality, we may assume $\kappa = (C_1, C_2, \ldots, C_\ell)$. Then, we have

$$H_\kappa(z, w) = H_{11}(z, w) M_{12}(z) H_{22}(z, w) \cdots M_{\ell-1\ell}(z) H_{\ell\ell}(z, w) \cdot w^{\ell-1} \tag{24}$$

and it is clear that, for any fixed z, the singularities of $\xi_1^{\mathrm{T}} H_\kappa(z, w)\eta_\ell$ are those of the matrices $H_{jj}(z, w)$ for $j = 1, 2, \ldots, \ell$. If $k = 0$ and z near 0, the radius of convergence of each $H_{jj}(z, w)$ is greater than λ^{-1} and hence $r_n{}^{(\kappa)}(z) = \mathrm{O}(\rho^n)$ for some $0 < \rho < \lambda$, which proves point 1.

Now, set $I = \{j : C_j \text{ is dominant}\}$ and assume $k = \sharp I \geqslant 1$. Let $j \in I$ and let z be a complex value near 0. By property (ii), the dominant singularities of $H_{jj}(z, w)$ are the simple poles $y_\alpha(z)^{-1}$, where $\alpha \in E_j$. Thus, the same values are poles for $\xi_1^{\mathrm{T}} H_\kappa(z, w)\eta_\ell$ of degree k at most. Hence, $r_n{}^{(\kappa)}(z)$ is bounded by a linear combinations of terms of the form $\mathrm{O}(y_\alpha(z)^n n^{k-1})$, where $\alpha \in \bigcup_{j \in I} E_j$; setting $z = c/n$, by (21), each of them is of the order $\mathrm{O}(\lambda^n n^{k-1})$, which proves point 2.

[2] In this work, for any pair of sequences $\{f_n\}, \{g_n\} \subseteq \mathbb{C}$, the expression $f_n = \Theta(g_n)$ means that for two positive constants a, b the relation $a|g_n| \leqslant |f_n| \leqslant b|g_n|$ holds for every n large enough.

Analogously, setting $z = \mathrm{i}tn^{-1/2}$, again by (21) for every α we have

$$|y_\alpha(\mathrm{i}tn^{-1/2})^n| = \left|\alpha^n\left(1 + m_\alpha \frac{\mathrm{i}t}{\sqrt{n}} + \mathrm{O}(1/n)\right)\right|^n;$$

since $m_\alpha \in \mathbb{R}$ this implies

$$|y_\alpha(\mathrm{i}tn^{-1/2})^n| = \lambda^n\left(1 + \frac{m_\alpha^2 t^2}{n}\right)^{n/2} \mathrm{O}(1 + 1/n) = \mathrm{O}(\lambda^n)$$

and hence $r_n{}^{(\kappa)}(\mathrm{i}t/\sqrt{n}) = \mathrm{O}\left(\sum_\alpha y_\alpha(\mathrm{i}tn^{-1/2})^n n^{k-1}\right) = \mathrm{O}(\lambda^n n^{k-1})$, which proves point 4.

Finally, assume C_j primitive for every $j \in I$ and let z be a complex value near 0. Then, the main singularities of $\xi_1^{\mathrm{T}} H_\kappa(z, w)\eta_\ell$ are the values $y_j(z)^{-1}$ defined in (i). By (24) this implies

$$\xi_1^{\mathrm{T}} H_\kappa(z, w)\eta_\ell = \frac{R(z, w)}{\Pi_{j\in I}(1 - y_j(z)w)}, \tag{25}$$

where $R(z, w)$ is a function analytic in a disk $\{w \in \mathbb{C} | |w| \leqslant \lambda^{-1} + d\}$, for some $d > 0$. Thus, the leading term of $r_n{}^{(\kappa)}(z)$ is determined by the convolution of the sequences $\{y_j(z)^n\}_n$, for $j \in I$; hence, setting $z = c/n$ and using (20) we get $r_n{}^{(\kappa)}(c/n) = \Theta(\lambda^n n^{k-1})$ proving point 3. $\square$

Since by Eq. (17), we have $r_n(z) = \sum_{\kappa\in\Gamma} r_n^{(\kappa)}(z)$, we obtain the following result, which shows the key role of the main chains. Also note that the property does not hold if the main chains admit nonprimitive dominant components.

Theorem 10. *If all dominant components of the main chains are primitive then, for every constant $c \in \mathbb{C}$, we have*

$$r_n(c/n) = \sum_{\kappa\in\Gamma_m} r_n^{(\kappa)}(c/n)(1 + \mathrm{O}(1/n)) = \Theta(\lambda^n n^{k-1}),$$

where k is the order of the main chains.

6. Limit distributions in multicomponent models

Theorem 10 shows that in a multicomponent model the asymptotic behaviour of our statistics mainly depends on the main chains. This fact leads to study the relevant case when the model has just one main chain. In this case, assuming further mild conditions on the dominant components, it turns out that the limit distribution of Y_n/n coincides with a Vandermonde distribution. For this reason we introduce the notion of simple model.

Let (ξ, μ, η) be a linear representation over the alphabet $\{a, b\}$ with coefficients in $\mathbb{R}_+$. We say that (ξ, μ, η) is a *simple* linear representation, or just a *simple model*, if Γ_m contains only one chain κ and every dominant component in κ is primitive.

In simple models the limit distribution of Y_n first depends on the order k of κ, i.e. the number of its dominant components. If $k \leqslant 2$ the limit distribution is known and derives from the analysis of the bicomponent models given in [5]; in particular (if the dominant components are not degenerate) we have the following results:

- If κ has only one dominant component C_i then the limit distribution of $Y_n - \beta_i n/\sqrt{\gamma_i n}$ is a Gaussian distribution of mean value 0 and variance 1.
- If κ has two dominant components C_i, C_j then we have the following three subcases:
  (1) If $\beta_i \neq \beta_j$ then Y_n/n converges in law to a random variable uniformly distributed in the interval $[b_1, b_2]$, where $b_1 = \min\{\beta_i, \beta_j\}$ and $b_2 = \max\{\beta_i, \beta_j\}$;
  (2) If $\beta_i = \beta_j = \beta$ but $\gamma_i \neq \gamma_j$ then the limit distribution of $(Y_n - \beta n)/\sqrt{n}$ is a mixture of normal distributions of mean value 0 and variance uniformly distributed in the interval $[c_1, c_2]$, where $c_1 = \min\{\gamma_i, \gamma_j\}$ and $c_2 = \max\{\gamma_i, \gamma_j\}$. In other words, $(Y_n - \beta n)/\sqrt{n}$ converges in law to a random variable with density function

$$f(x) = \frac{1}{c_2 - c_1}\int_{c_1}^{c_2} \frac{e^{-x^2/(2v)}}{\sqrt{2\pi v}}\,\mathrm{d}v,$$

which has characteristic function

$$F(t) = 2\frac{e^{-c_1t^2/2} - e^{-c_2t^2/2}}{(c_2 - c_1)t^2}.$$

Notice that $F(t) = \Phi_P(it^2/2)$ where $P(x) = (x - c_1)(x - c_2)$;

(3) If $\beta_i = \beta_j = \beta$ and $\gamma_i = \gamma_j = \gamma$ then the distribution of $(Y_n - \beta n)/\sqrt{\gamma n}$ again converges to a Gaussian distribution of mean value 0 and variance 1.

Here, we determine the limit distribution of Y_n/n for simple models having main chain κ of order $k \geqslant 2$. We only assume that κ has at least two dominant components with different mean constants. In Section 7 we extend this result to the case when all dominant components of κ have the same mean constant.

Theorem 11. *Let Y_n be defined in a simple model with main chain κ of order $k \geqslant 2$. Let $\beta_1, \ldots, \beta_r$ denote the mean constants of the dominant components in κ in increasing order and assume $r \geqslant 2$. Also, for each $\ell = 1, 2, \ldots, r$, let m_ℓ be the multiplicity of β_ℓ, that is the number of dominant components in κ whose mean constant equals β_ℓ. Then, Y_n/n converges in law to a Vandermonde random variable associated with the polynomial $P(x) = \prod_{\ell=1}^{r}(x - \beta_\ell)^{m_\ell}$.*

Observe that in the case $k = 2$ we obtain the result stated in point (1) above. The proof of the theorem is based on the analysis of the characteristic function of Y_n/n, which by Eq. (13) is given by

$$\Psi_{Y_n/n}(t) = \frac{r_n(it/n)}{r_n(0)}. \tag{26}$$

Thus, we first present the following lemma, which provides a useful expression for $r_n(it/n)$. To this aim, as in Eq. (7), let $\{g_Q(n)\}_n$ be the sequence having generating function $Q(0)/Q(x)$, where

$$Q(x) = \prod_{\ell=1}^{r}\left(x - \frac{1}{1 + \beta_\ell it/n}\right)^{m_\ell}.$$

Lemma 12. *Assume the hypotheses of Theorem 11. Then, for every $t \in \mathbb{R}$, as n grows to $+\infty$ we have*

$$r_n\left(\frac{it}{n}\right) = \sum_{s=0}^{k-1}\lambda^{n-s}a_s\left(\frac{it}{n}\right) \cdot g_Q(n - s) \cdot (1 + O(1/n)) \tag{27}$$

and in particular

$$r_n(0) = \frac{n^{k-1}}{(k-1)!}\left(\sum_{s=0}^{k-1}\lambda^{n-s}a_s(0)\right) \cdot (1 + O(1/n)) \tag{28}$$

where, for each s, $a_s(z)$ is an analytic function at $z = 0$.

Proof. By Theorem 10, we get $r_n(it/n) = r_n^{(\kappa)}(it/n)(1 + O(1/n))$ and hence we have to show that $r_n^{(\kappa)}(it/n)$ equals the right hand side of (27). We can evaluate $r_n^{(\kappa)}(it/n)$ by refining the proof of point 3 in Proposition 9. Indeed, since $H_\kappa(z, w)$ satisfies Eq. (25), we have

$$\xi_1^{\mathrm{T}} H_\kappa(z, w)\eta_\ell = \sum_{s=0}^{k-1} a_s(z)w^s \cdot \prod_{\ell=1}^{r}(1 - f_\ell(z)w)^{-m_\ell} + G(z, w), \tag{29}$$

where each $a_s(z)$ is a polynomial in e^z, $f_\ell(z) = \lambda(1 + \beta_\ell z + O(z))$ for every $\ell = 1, 2, \ldots, r$ and, for all z near 0, the function $G(z, w)$ is analytic in a disk $\{w \in \mathbb{C} | |w| \leqslant \lambda^{-1} + d\}$, for some $d > 0$.

Clearly $\prod_{\ell=1}^{r}(1 - f_\ell(z)w)^{-m_\ell}$ is the generating function of the sequence whose nth term is

$$\sum_{\sum_\ell n_\ell = n}\prod_{\ell=1}^{r}\binom{n_\ell + m_\ell - 1}{m_\ell - 1} f_\ell(z)^{n_\ell}.$$

Setting $z = \mathrm{i}t/n$ and recalling Eq. (7), the previous expression can be rewritten as $\lambda^n g_Q(n) \cdot (1 + \mathrm{O}(1/n))$. Thus, since (29) is the generating function of $\{r_n^{(\kappa)}(z)\}_n$, the main term of $r_n^{(\kappa)}(\mathrm{i}t/n)$ is given by the convolution of $\{a_n(\mathrm{i}t/n)\}_n$ and $\{\lambda^n g_Q(n)\}_n$, which leads to Eq. (27). Eq. (28) follows by noting that if $t = 0$, then $Q(x) = (x-1)^k$ and $g_Q(n) = \binom{n+k-1}{k-1}$. □

Proof of Theorem 11. As n grows to infinity, the behaviour of $g_Q(n-s)$ does not depend on s. To prove this fact we need the following equalities, that can be proved from the definitions given in Section 3. Set $v_\ell = (1 - \beta_\ell \mathrm{i}t/n)^{-1}$, for every $\ell, j = 1, 2, \ldots, r$ we have

$$(v_\ell)^{-n} = \mathrm{e}^{-\mathrm{i}t\beta_\ell} \cdot (1 + \mathrm{O}(1/n)),$$
$$(v_\ell - v_j)^{-1} = \left(-\frac{n}{\mathrm{i}t}\right)(\beta_\ell - \beta_j)^{-1} \cdot (1 + \mathrm{O}(1/n)).$$

Moreover, if $w_{\ell j}$ and $c_{\ell j}$ are the entries of the last column of the Vandermonde matrices associated with the polynomial $Q(x)$ and $P(x) = \prod_{\ell=1}^r (x - \beta_\ell)^{m_\ell}$, respectively, then using Eq. (6) one can obtain

$$w_{\ell j} = \left(-\frac{n}{\mathrm{i}t}\right) c_{\ell j} \cdot (1 + \mathrm{O}(1/n)).$$

As a consequence, applying Proposition 3 and using the previous equalities, one proves that

$$g_Q(n-s) = \left(\frac{n}{\mathrm{i}t}\right)^{k-1} \left(\sum_{\ell=1}^r \mathrm{e}^{\mathrm{i}t\beta_\ell} \sum_{j=1}^{m_\ell} c_{\ell j} \frac{(\mathrm{i}t)^{j-1}}{(j-1)!}\right) \cdot (1 + \mathrm{O}(1/n)).$$

Replacing the previous expression into (27) we get

$$r_n\left(\frac{\mathrm{i}t}{n}\right) = \left(\frac{n}{\mathrm{i}t}\right)^{k-1} \left(\sum_{\ell=1}^r \mathrm{e}^{\mathrm{i}t\beta_\ell} \sum_{j=1}^{m_\ell} c_{\ell j} \frac{(\mathrm{i}t)^{j-1}}{(j-1)!}\right) \left(\sum_{s=0}^{k-1} \lambda^{n-s} a_s\left(\frac{\mathrm{i}t}{n}\right)\right) \cdot (1 + \mathrm{O}(1/n)).$$

Hence, applying Eq. (26) and recalling Eq. (28), one can see that the characteristic function $\Psi_{Y_n/n}(t)$ converges to $\Phi_P(t)$ for every $t \in \mathbb{R}$. This proves the result. □

7. Further results

The analysis presented in the previous section can be extended to all simple models, also when the mean constants β_j's (associated with the dominant components of the main chain) are totally coincident. In this case, clearly $Y_n^{(\kappa)}/n$ converges in distribution to such a constant, and it is natural to consider a finer normalization. With respect to this point, the following theorem holds, which can be proved as Theorem 11.

Theorem 13. *Let Y_n be defined in a simple model having main chain κ of order $k \geqslant 2$ and assume that all dominant components in κ have the same mean constant β. Let $\gamma_1, \ldots, \gamma_s$ be the distinct variance constants (in increasing order) of the dominant components in κ and let $m_1, m_2, \ldots, m_s$ denote their multiciplities. If $s = 1$ and $\gamma_1 \neq 0$, then $(Y_n - \beta n)/\sqrt{\gamma_1 n}$ converges in distribution to a normal random variable of mean 0 and variance 1. Otherwise, if $s > 1$, then $(Y_n - \beta n)/\sqrt{n}$ converges in distribution to a random variable having characteristic function $\Phi_P(\mathrm{i}t^2/2)$, where $\Phi_P(t)$ is the characteristic function of a Vandermonde random variable of parameter $P(x) = \prod_{\ell=1}^s (x - \gamma_\ell)^{m_\ell}$.*

Notice that, for $k = 2$, the previous theorem reduces to points (2) and (3) at p. 21.

The results of Theorem 11, concerning the limit distribution of $Y_n^{(\kappa)}/n$, can be further extended by a standard conditioning argument (already used in [5]) to all rational models (ξ, μ, η) whose main chains are "simple", i.e. for every $\kappa \in \Gamma_m$ all dominant components in κ are primitive. In this case, by Eqs. (22) and (23), for every $\kappa \in \Gamma_m$ one can easily see that

$$r_n^{(\kappa)}(z) = s_\kappa(z)\lambda^n n^{k-1} + \mathrm{O}(\lambda^n n^{k-2}),$$

where k is the degree of κ and $s_\kappa(z)$ is a nonnull analytic function at $z = 0$. Then, by Theorem 10, we have

$$r_n(0) = R\lambda^n n^{k-1} + \mathrm{O}(\lambda^n n^{k-2})$$

where $R = \sum_{\kappa \in \Gamma_m} s_\kappa(0)$. We can also associate each $\kappa \in \Gamma_m$ with the probability value p_κ, given by $p_\kappa = s_\kappa(0)/R$. Note that the values $\{p_\kappa\}_{\kappa\in\Gamma_m}$ define a discrete probability measure and they can be explicitly computed from the triple (ξ, μ, η).

Moreover, each $\kappa \in \Gamma_m$ defines a simple rational model in its own right, with an associate sequence of random variables $\{Y_n^{(\kappa)}\}$. The limit distribution of $Y_n^{(\kappa)}/n$ can be studied by applying Theorem 11. In particular, $Y_n^{(\kappa)}/n$ always converges in distribution to a random variable of distribution function $F_\kappa(x)$ defined according to the previous results. Note that, if all constants β_j's are here equal, then $F_\kappa(x)$ reduces to the degenerate distribution of mass point β_1. Now, it is not difficult to see that the overall statistics Y_n/n converges in distribution to a r.v. of distribution function $F(x)$ defined by $F(x) = \sum_{\kappa\in\Gamma_m} F_\kappa(x)p_\kappa$. This completes our analysis of the limit distribution of $Y_n^{(\kappa)}/n$. The only family of rational stochastic models not covered by our results consists of those models having a main chain with some nonprimitive dominant component; in those cases, periodicity phenomena occur.

References

[1] J. Berstel, C. Reutenauer, Rational Series and their Languages, Springer, Berlin, 1988.

[2] A. Bertoni, C. Choffrut, M. Goldwurm, V. Lonati, On the number of occurrences of a symbol in words of regular languages, Theoret. Comput. Sci. 302 (1–3) (2003) 431–456.

[3] J. Bourdon, B. Vallée, Generalized pattern matching statistics, Mathematics and Computer Science II: Algorithms, Trees, Combinatorics and Probabilities. Proc. Versailles Colloquium, Birkhauser, 2002, pp. 249–265.

[4] F.G. Csàki, Some notes on the inversion of confluent Vandermonde matrices, IEEE Trans. Automat. Control 20 (1975) 154–157.

[5] D. de Falco, M. Goldwurm, V. Lonati, Frequency of symbol occurrences in bicomponent stochastic models, Theoret. Comput. Sci. 327 (3) (2004) 269–300.

[6] P. Flajolet, Mathematical methods in the analysis of algorithms and data structures, in: E. Börger (Ed.), Trends in Theoretical Computer Science, Computer Science Press, Rock Ville, MD, 1988, pp. 225–304.

[7] B.V. Gnedenko, The Theory of Probability, Mir Publishers, Moscow, 1976 (translated by G. Yankovsky).

[8] M. Goldwurm, Probabilistic estimation of the number of prefixes of a trace, Theoret. Comput. Sci. 92 (1992) 249–268.

[9] P. Grabner, M. Rigo, Additive functions with respect to numeration systems on regular languages, Monatshefte für Mathematik 139 (2003) 205–219.

[10] L.J. Guibas, A.M. Odlyzko, Maximal prefix-synchronized codes, SIAM J. Appl. Math. 35 (1978) 401–418.

[11] L.J. Guibas, A.M. Odlyzko, String overlaps, pattern matching, and nontransitive games, J. Combin. Theory. Ser. A 30 (2) (1981) 183–208.

[12] P. Henrici, Applied and Computational Complex Analysis, Vol. 1, Wiley, UK, 1974.

[13] P. Jokinen, E. Ukkonen, Two algorithms for approximate string matching in static texts, Proc. MFCS'91, Lecture Notes in Computer Science, Vol. 520, Springer, Berlin, 1991, pp. 240–248.

[14] P. Lancaster, M. Tismenetsky, The Theory of Matrices, Academic Press, New York, 1985.

[15] U. Luther, K. Rost, Matrix exponentials and inversions of confluent Vandermonde matrices, Electron. Trans. Numer. Anal. 18 (2004) 91–100.

[16] P. Nicodeme, B. Salvy, P. Flajolet, Motif statistics, Theoret. Comput. Sci. 287 (2) (2002) 593–617.

[17] B. Prum, F. Rudolphe, E. Turckheim, Finding words with unexpected frequencies in deoxyribonucleic acid sequence, J. Roy. Statist. Soc. Ser. B 57 (1995) 205–220.

[18] M. Régnier, W. Szpankowski, On pattern frequency occurrences in a Markovian sequence, Algorithmica 22 (4) (1998) 621–649.

[19] A. Salomaa, M. Soittola, Automata-Theoretic Aspects of Formal Power Series, Springer, Berlin, 1978.

[20] R. Sedgewick, P. Flajolet, An Introduction to the Analysis of Algorithms, Addison-Wesley, Reading, MA, 1996.

[21] E. Seneta, Non-negative Matrices and Markov Chains, Springer, New York, Heidelberg, Berlin, 1981.

[22] E. Sutinen, W. Szpankowski, On the collapse of q-gram filtration, Proc. FUN with Algorithms, Elba, 1998, pp. 178–193.