

Local limit properties for pattern statistics and rational models^{*†}

Alberto Bertoni[◦] Christian Choffrut[‡]
Massimiliano Goldwurm[◦] Violetta Lonati[◦]

[◦] Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano
Via Comelico 39/41, 20135 Milano – Italy, {bertoni,goldwurm,lonati}@dsi.unimi.it

[‡] L.I.A.F.A., Université Paris VII, 2 Place Jussieu, 75221 Paris – France
Christian.Choffrut@liafa.jussieu.fr

September 2004

Abstract

Motivated by problems of pattern statistics, we study the limit distribution of the random variable counting the number of occurrences of the symbol a in a word of length n chosen at random in $\{a, b\}^*$, according to a probability distribution defined via a rational formal series s with positive real coefficients. Our main result is a local limit theorem of Gaussian type for these statistics under the hypothesis that s is a power of a primitive series. This result is obtained by showing a general criterion for (Gaussian) local limit laws of sequences of integer random variables. To prove our result we also introduce and analyze a notion of symbol-periodicity for irreducible matrices, whose entries are polynomials over positive semirings; the properties we prove on this topic extend the classical Perron–Frobenius theory of non–negative real matrices. As a further application we obtain some asymptotic evaluations of the maximum coefficient of monomials of given size for rational series in two commutative variables.

1 Introduction

A typical problem in pattern statistics consists of estimating the frequency of occurrences of given strings in a random text, where the set of patterns is fixed in advance and the text is a word of length n randomly generated according to a probabilistic model (for instance, a Markovian model). In this context, relevant informations are the asymptotic evaluations of the mean value and the variance of the number of occurrences of patterns in the text, as well as its limit distribution. These problems are widely studied in the literature and they are of

* Appeared in revised form in *Theory of Computing Systems*, vol. 39 (1), 209–235, 2006.

† This work includes results presented in two distinct papers appeared respectively in *Proc. STACS 2004*, 21st Symposium on Theoretical Aspects of Computer Science, V. Diekert and M. Habib editors, LNCS n. 2996, 117–128, Springer, 2004; and *Proc. DLT 2004*, 8th Int. Conference on Developments in Language Theory, C.S. Calude, E. Calude, M.J. Dinneen editors, LNCS n. 3340, 114–126, Springer, 2004.

interest for the large variety of applications in different areas of computer science, probability theory and molecular biology (see for instance [8, 17, 15, 22]). Many results show a normal limit distribution of the number of pattern occurrences in the sense of the central or local limit theorem [1]; observe that the “local” result is usually stronger since it concerns the probability of single point values, while the “central” limit refers to the cumulative distribution function. In [14] limit distributions are obtained for the number of (positions of) occurrences of words from a regular language in a random string of length n generated in a Bernoulli or a Markovian model. These results are extended in [3] to the so-called *rational stochastic model*, where the pattern is reduced to a single symbol and the random text is a word over a two-letter alphabet, generated according to a probability distribution defined via a weighted finite automaton or, equivalently, via a rational formal series. This model properly includes the Markovian models when the set of patterns is given by a regular language (as studied in [14]). The analysis presented in [3] shows that a central limit theorem holds in the rational models in the primitive case, i.e., when the matrix associated with the finite automaton (counting the transitions between states) is primitive. A Gaussian local limit theorem for a proper subclass of these primitive models is also proven.

The present paper offers a complete local limit theorem for all primitive models. Our contribution is however more general and includes several results that are of interest in their own rights. In particular we introduce and analyze a notion of periodicity for (polynomial) irreducible matrices and we prove a general criterion for local limit theorems of integer random variables. Our main results can be summarized as follows:

1. In Section 3 we introduce a notion of x -periodicity for irreducible matrices whose entries are polynomials in the variable x over an arbitrary positive semiring. Intuitively, considering the matrix as a labeled graph, its x -period is the greatest common divisor of the differences between the number of occurrences of x in (labels of) cycles of the same length. We prove several properties concerning this notion that extend the classical Perron–Frobenius theory of non-negative real matrices [21]. In particular, these results are useful to study the eigenvalues of matrices of the form $Ax + B$, where A and B are matrices with coefficients in \mathbb{R}_+ and $x \in \mathbb{C}$ lies on the circle $|x| = 1$ (see Theorem 10).
2. In Section 4 we give a general criterion for local limit laws that holds for sequences of discrete random variables $\{X_n\}$ with values in a linear progression of a fixed period d included in $\{0, 1, \dots, n\}$. Such a criterion is based on the so-called “quasi-power” conditions for Central limit theorems [10, 6] together with a uniform convergence property for the sequence of the characteristic functions of $\{X_n\}$. Under these hypotheses the probability function of X_n approximates a Gaussian density function up to the period d . If $d = 1$ the criterion includes the classical de Moivre–Laplace Theorem as a special case.
3. In Section 5 we present our main result, i.e., a local limit theorem for pattern statistics in rational stochastic models defined by powers of primitive formal series. This includes the primitive models but it also includes rational models, with an arbitrary number of strongly connected components, such that all these components are equal. The proof is based on both the criterion for local limit laws and the periodicity properties of polynomial matrices described above. We also prove a central limit theorem (in the classical sense), which holds for the same models.

4. Finally, in Section 6 we present an application of the previous analysis to the problem of estimating the maximum coefficients of rational formal series in two commutative variables. We show that, for any series s of this kind given by the power of a primitive formal series, the maximum coefficient of monomials of size n in s is $\Theta(n^{k/2}\lambda^n)$ for some integer $k \geq -1$ and some positive real λ .¹

2 Preliminaries

In this section we recall some basic notions and properties concerning rational formal series [19, 2] and weighted automata. Consider a monoid \mathcal{M} and a positive semiring \mathcal{S} , that is a semiring such that $x + y = 0$ implies $x = y = 0$ and $x \cdot y = 0$ implies $x = 0$ or $y = 0$ [13]. Examples are given by \mathbb{N} , \mathbb{R}_+ and the Boolean algebra \mathbb{B} . We call *formal series* over \mathcal{M} any application $r : \mathcal{M} \rightarrow \mathcal{S}$ that associates with each $\omega \in \mathcal{M}$ its *coefficient* $(r, \omega) \in \mathcal{S}$. Such a mapping is usually represented as an infinite formal sum $r = \sum_{\omega \in \mathcal{M}} (r, \omega) \omega$. The set of all series over \mathcal{M} with coefficients in \mathcal{S} is a monoid algebra, provided with the usual operations of sum, product and star restricted to the elements r such that $(r, 1_{\mathcal{M}}) = 0_{\mathcal{S}}$. These operations are called rational operations. A series r is called *rational* if it belongs to the smallest set closed under rational operations, containing the series 0 and all the series χ_{ω} , for $\omega \in \mathcal{M}$, such that $(\chi_{\omega}, \omega) = 1$ and $(\chi_{\omega}, x) = 0$ for each $x \neq \omega$.

In this work we consider formal series over free monoids or free commutative monoids, with coefficients in the semiring \mathbb{R}_+ of nonnegative real numbers. If \mathcal{M} is the free monoid Σ^* generated by a set symbols $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_{\ell}\}$, for each $\omega \in \Sigma^*$, we use $|\omega|$ to denote its length and $|\omega|_{\sigma}$ to denote the number of occurrences of the symbol $\sigma \in \Sigma$ in ω . Moreover, let $\mathbb{R}_+ \langle\langle \sigma_1, \dots, \sigma_{\ell} \rangle\rangle$ denote the family of all formal series over \mathcal{M} with coefficients in \mathbb{R}_+ and $\mathbb{R}_+^{Rat} \langle\langle \sigma_1, \dots, \sigma_{\ell} \rangle\rangle$ the subset of rational series.

Analogously, if \mathcal{M} is the free commutative monoid Σ^{\otimes} with the same set of generators Σ , then the family of all formal series over \mathcal{M} with coefficients in \mathbb{R}_+ and its subset of rational series are denoted by $\mathbb{R}_+[[\sigma_1, \dots, \sigma_{\ell}]]$ and $\mathbb{R}_+^{Rat}[[\sigma_1, \dots, \sigma_{\ell}]]$, respectively. In this case, any element $\sigma_1^{i_1} \dots \sigma_{\ell}^{i_{\ell}}$ of \mathcal{M} is represented in the form $\underline{\sigma}^{\underline{i}}$, where $\underline{\sigma} = (\sigma_1, \dots, \sigma_{\ell})$ and $\underline{i} = (i_1, \dots, i_{\ell}) \in \mathbb{N}^{\ell}$. Furthermore, consider the canonical morphism $\mathcal{F} : \{\sigma_1, \dots, \sigma_{\ell}\}^* \rightarrow \mathcal{M}$, that associates with each $\omega \in \{\sigma_1, \dots, \sigma_{\ell}\}^*$ the monomial $\underline{\sigma}^{\underline{i}}$ (where $i_j = |\omega|_{\sigma_j}$ for every j). It extends to the semiring of formal series: for every $r \in \mathbb{R}_+ \langle\langle \sigma_1, \dots, \sigma_{\ell} \rangle\rangle$ and every $\underline{\sigma}^{\underline{i}} \in \mathcal{M}$,

$$(\mathcal{F}(r), \underline{\sigma}^{\underline{i}}) = \sum_{\substack{|\omega|_{\sigma_j} = i_j \\ j=1,2,\dots,\ell}} (r, \omega).$$

\mathcal{F} being a morphism, for every $r \in \mathbb{R}_+^{Rat} \langle\langle \sigma_1, \dots, \sigma_{\ell} \rangle\rangle$, $\mathcal{F}(r)$ belongs to $\mathbb{R}_+^{Rat}[[\sigma_1, \dots, \sigma_{\ell}]]$.

By Kleene's Theorem [19, 18], we know that every $r \in \mathbb{R}_+^{Rat} \langle\langle \sigma_1, \dots, \sigma_{\ell} \rangle\rangle$ admits a *linear representation* over the alphabet $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_{\ell}\}$, i.e., a triple (ξ, μ, η) such that, for some integer $m > 0$, ξ and η are (nonnull column) vectors in \mathbb{R}_+^m and $\mu : \Sigma^* \rightarrow \mathbb{R}_+^{m \times m}$ is a monoid morphism, satisfying $(r, \omega) = \xi^T \mu(\omega) \eta$ for each $\omega \in \Sigma^*$. We define m as the *size* of the representation. We also say that (ξ, μ, η) is *degenerate* if for some $\sigma \in \Sigma$ we

¹Throughout this work we use the symbol Θ with the standard meaning: given two sequences $\{h_n\} \subseteq \mathbb{C}$ and $\{f_n\} \subseteq \mathbb{R}_+$, the equality $h_n = \Theta(f_n)$ means that for some pair of positive constant c_1, c_2 , the relation $c_1 f_n \leq |h_n| \leq c_2 f_n$ holds for any n large enough.

have $\mu(\rho) = 0$ for each $\rho \in \Sigma, \rho \neq \sigma$. Analogously, we say that $r \in \mathbb{R}_+ \langle\langle \sigma_1, \dots, \sigma_\ell \rangle\rangle$ is degenerate if, for some $\sigma \in \Sigma, (r, \omega) \neq 0$ implies $\omega \in \{\sigma\}^*$.

Observe that considering a linear representation is equivalent to defining a (weighted) non-deterministic finite automaton over the alphabet $\{\sigma_1, \sigma_2, \dots, \sigma_\ell\}$, where the state set 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.

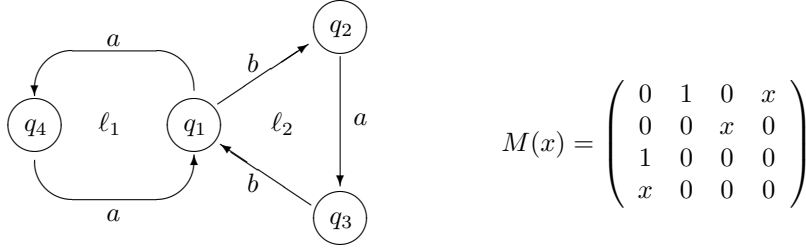


Figure 1: Example of state diagram and corresponding a -counting matrix.

It is convenient to represent the morphism μ by its *state diagram* (see Figure 1), which is a labelled directed graph where the vertices are given by the set $\{1, 2, \dots, m\}$ and there exists an edge with label $\sigma \in \Sigma$ from vertex p to vertex q if and only if $\mu(\sigma)_{pq} \neq 0$. A *path* of length n is a sequence of labeled edges of the form

$$\ell = q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \dots q_{n-1} \xrightarrow{a_n} q_n ;$$

in particular, if $q_n = q_0$ we say that ℓ is a q_0 -*cycle* and if all a_i are distinct the cycle is called *simple*. Moreover we say that $\omega = a_1 a_2 \dots a_n$ is the label of ℓ and we denote the number of occurrences of σ in ℓ by $|\ell|_\sigma = |\omega|_\sigma$.

Since we are interested in the occurrences of a particular symbol $a \in \Sigma$, we may set $A = \mu(a), B = \sum_{b \neq a} \mu(b)$ and consider the *a-counting matrix* $M(x) = Ax + B$, which can be interpreted as a matrix whose entries are polynomials in $\mathbb{R}_+[x]$ of degree lower than 2. Moreover, observe that for every $n \in \mathbb{N}$ we can write

$$\xi^T M(x)^n \eta = \sum_{|\omega|=n} (r, \omega) \cdot x^{|\omega|_a} . \quad (1)$$

Therefore $M(x)^n$ is related to the paths of length n of the associated state diagram, in the sense that the pq -entry of $M(x)^n$ is the sum of monomials of the form sx^k where $k = |\ell|_a$ for some path ℓ of length n from p to q in the state diagram.

Example 1 Consider the state diagram represented in Figure 1 and its a -counting matrix $M(x)$. The 6-th power of $M(x)$ represents the paths of length 6:

$$M(x)^6 = \begin{pmatrix} x^2 + x^6 & 2x^3 & x^5 & 2x^4 \\ x^5 & x^2 & x^4 & x^3 \\ x^3 & x^4 & x^2 & x^5 \\ 2x^4 & x^5 & x^3 & x^6 \end{pmatrix} .$$

For instance, one can easily verify that there are two q_1 -cycles of length 6 containing, respectively, 2 and 6 occurrences of a , which correspond to the monomials x^2 and x^6 .

3 The symbol periodicity of matrices

In this section we introduce the notion of x -periodicity for matrices over positive semirings of polynomials. First let us recall the classical notion of periodicity of matrices over positive semirings. Given a finite set Q and a positive semiring \mathcal{S} , consider a matrix $M : Q \times Q \rightarrow \mathcal{S}$. We say that M is *positive* whenever $M_{pq} \neq 0$ holds for all $p, q \in Q$, in which case we write $M > 0$. To avoid the use of brackets, from now on, we use the expression M^n_{pq} to denote the pq -entry of the matrix M^n . For every index q , we call *period* of q the greatest common divisor (GCD) of the positive integers h such that $M^h_{qq} \neq 0$, with the convention that $\text{GCD}(\emptyset) = +\infty$. Moreover, we recall that a matrix M is said to be *irreducible* if for every pair of indices p, q , there exists a positive integer $h = h(p, q)$ such that $M^h_{pq} \neq 0$; in this case, it turns out that all indices have the same period, which is finite and is called the *period* of M . Finally, the matrix is called *primitive* if there exists a positive integer h such that $M^h > 0$, which implies $M^n > 0$ for every $n \geq h$. It is well known that M is primitive if and only if M is irreducible and has period 1.

When \mathcal{S} is the semiring of positive real numbers a classical result is given by the following theorem (see [21]).

Theorem 1 (Perron–Frobenius) *Let M be a primitive matrix with entries in \mathbb{R}_+ . Then, M admits exactly one eigenvalue λ of maximum modulus (called the Perron–Frobenius eigenvalue of M), which is a simple root of the characteristic polynomial of M . Moreover, λ is real and positive and there exist strictly positive left and right eigenvectors u and v associated with λ such that $v^T u = 1$.*

A consequence of this theorem is that, for any primitive matrix M with entries in \mathbb{R}_+ , the relation $M^n \sim \lambda^n \cdot uv^T$ holds as n tends to $+\infty$, where λ , u and v are defined as above. A further application is given by the following proposition [21, Exercise 1.9], to be used in the next sections. Here, for any complex matrix C , $|C|$ denotes the matrix with pq -entry given by $|C_{pq}|$ for all indices p, q .

Proposition 2 *Let γ be an eigenvalue of an arbitrary complex matrix C . If M is a primitive matrix over \mathbb{R}_+ of the same size such that $|C_{pq}| \leq M_{pq}$ for every p, q and if λ is its Perron–Frobenius eigenvalue, then $|\gamma| \leq \lambda$. Moreover, if $|\gamma| = \lambda$, then necessarily $|C| = M$.*

Let us now introduce the notion of x -periodicity for matrices in the semiring $\mathcal{S}[x]$ of polynomials in the variable x with coefficients in \mathcal{S} and focus more specifically on the case of irreducible matrices.

3.1 The notion of x -period

Given a polynomial $F = \sum_k f_k x^k \in \mathcal{S}[x]$, we define the x -period of F as the integer $d(F) = \text{GCD}\{|h - k| \mid f_h \neq 0, f_k \neq 0\}$, where we assume $\text{GCD}(\{0\}) = \text{GCD}(\emptyset) = +\infty$. Observe that $d(F) = +\infty$ if and only if $F = 0$ or F is a monomial.

Now consider a finite set Q and a matrix $M : Q \times Q \rightarrow \mathcal{S}[x]$. For any index $q \in Q$ and for each integer n we set $d(q, n) = d(M^n_{qq})$ and we define the x -period of q as the integer $d(q) = \text{GCD}\{d(q, n) \mid n \geq 0\}$, assuming that any non-zero element in $\mathbb{N} \cup \{+\infty\}$ divides $+\infty$. Notice that if M is the a -counting matrix of some linear representation, this definition implies that for every index q and for every pair of q -cycles \mathcal{C}_1 and \mathcal{C}_2 of equal

length, $|\mathcal{C}_1|_a - |\mathcal{C}_2|_a$ is a multiple of $d(q)$. More precisely, $d(q)$ is the GCD of the differences of number of occurrences of a in all pairs of q -cycles of equal length.

Proposition 3 *If M is an irreducible matrix over $\mathcal{S}[x]$, then all indices have the same x -period.*

Proof. Consider an arbitrary pair of indices p, q . By symmetry, it suffices to prove that $d(p)$ divides $d(q)$, and this again can be proven by showing that $d(p)$ divides $d(q, n)$ for all $n \in \mathbb{N}$. Since M is irreducible, there exist two integers s, t such that $M^s_{pq} \neq 0$ and $M^t_{qp} \neq 0$. Then the polynomial $M^{s+t}_{pp} = \sum_r M^s_{pr} M^t_{rp} \neq 0$ and for some $k \in \mathbb{N}$ there exists a monomial in M^{s+t}_{pp} with exponent k . Therefore, for every n and every exponent h in M^n_{qq} , the integer $h + k$ appears as an exponent in M^{n+s+t}_{pp} . This proves that $d(p, n + s + t)$ divides $d(q, n)$ and since $d(p)$ divides $d(p, n + s + t)$, this establishes the result. \square

Definition *The x -period of an irreducible matrix over $\mathcal{S}[x]$ is the common x -period of its indices. The matrix is said to be x -periodic if its x -period is greater than 1.*

Example 2 We compute the x -period of the matrix $M(x)$ over $\mathbb{R}_+[x]$ corresponding to the state diagram represented in Figure 1. Consider for instance state q_1 and let \mathcal{C}_1 and \mathcal{C}_2 be two arbitrary q_1 -cycles having the same length. Clearly they can be decomposed by using the simple q_1 -cycles of the automaton, namely

$$\ell_1 = q_1 \xrightarrow{a} q_4 \xrightarrow{a} q_1 \quad \text{and} \quad \ell_2 = q_1 \xrightarrow{b} q_2 \xrightarrow{a} q_3 \xrightarrow{b} q_1 .$$

Hence, up to their order, \mathcal{C}_1 and \mathcal{C}_2 only differ in the number of cycles ℓ_1 and ℓ_2 they contain: for $k = 1, 2$, let $s_k \in \mathbb{Z}$ be the difference between the number of ℓ_k contained in \mathcal{C}_1 and the number of ℓ_k contained in \mathcal{C}_2 . Then, necessarily, $s_1|\ell_1| + s_2|\ell_2| = 0$, that is $2s_1 + 3s_2 = 0$. This implies that $s_1 = 3n$ and $s_2 = -2n$ for some $n \in \mathbb{Z}$. Hence

$$|\mathcal{C}_1|_a - |\mathcal{C}_2|_a = 3n|\ell_1|_a - 2n|\ell_2|_a = 6n - 2n = 4n$$

This proves that 4 is a divisor of the x -period of $M(x)$. Moreover, both the q_1 -cycles ℓ_1^3 and ℓ_2^2 have length 6 and the numbers of occurrences of a differ exactly by 4. Hence, in this case, the x -period of $M(x)$ is exactly 4. \square

In the particular case where the entries of the matrix are all linear in x , the matrix decomposes $M = Ax + B$, where A and B are matrices over \mathcal{S} ; this clearly happens when M is the a -counting matrix of some linear representation. If further M is primitive, the following proposition holds.

Proposition 4 *Let A and B be matrices over \mathcal{S} and set $M = Ax + B$. If M is primitive, $A \neq 0$ and $B \neq 0$, then the x -period of M is finite.*

Proof. Let q be an arbitrary index and consider the finite family of pairs $\{(n_j, k_j)\}_{j \in J}$ such that $0 \leq k_j \leq n_j \leq m$ where m is the size of M and k_j appears as an exponent in $M^{n_j}_{qq}$. Notice that since M is irreducible J is not empty. Since every cycle can be decomposed into elementary cycles all of which of length at most m , the result is proved once we show that $d(q) = +\infty$ implies either $k_j = 0$ for all $j \in J$ or $k_j = n_j$ for all $j \in J$: in the first case we get $A = 0$ while in the second case we have $B = 0$.

Because of equality $M^{\prod_j n_j} = (M^{n_i})^{\prod_{j \neq i} n_j}$, the polynomial $M^{\prod_j n_j}_{qq}$ contains the exponent $k_i \prod_{j \neq i} n_j$ for each $i \in J$. Now, suppose by contradiction that $d(q)$ is not finite. This means that all exponents in $M^{\prod_j n_j}_{qq}$ are equal to a unique integer h such that $h = k_i \prod_{j \neq i} n_j$ for all $i \in J$. Hence, h must be a multiple of the least common multiple of all products $\prod_{j \neq i} n_j$. Now we have $\text{LCM}\{\prod_{j \neq i} n_j \mid i \in J\} \cdot \text{GCD}\{n_j \mid j \in J\} = \prod_j n_j$ and by the primitivity hypothesis $\text{GCD}\{n_j \mid j \in J\} = 1$ holds. Therefore h is a multiple of $\prod_j n_j$. Thus the conditions $k_j \leq n_j$ leave the only possibilities $k_j = 0$ for all $j \in J$ or $k_j = n_j$ for all $j \in J$ and this establishes the result. \square

Observe that the previous theorem cannot be extended to the case when M is irreducible or when M is a matrix over $\mathcal{S}[x]$ that cannot be written as $Ax + B$ for some matrices A and B over \mathcal{S} .

Example 3 The matrix M with entries $M_{11} = M_{22} = 0$, $M_{12} = x$ and $M_{21} = 1$ is irreducible but it is not primitive since it has period 2. It is easy to see that the non–null entries of all its powers are monomials, thus M has infinite x -period. \square

Example 4 Consider again Figure 1 and set $M(x)_{23} = x^3$. Then we obtain a primitive matrix over $\mathbb{R}_+[x]$ that cannot be written as $Ax + B$ and does not have finite x -period. \square

3.2 Properties of x -periodic matrices

Given a positive integer d , consider the cyclic group $C_d = \{1, g, g^2, \dots, g^{d-1}\}$ of order d and the semiring $\mathcal{B}_d = \langle \mathcal{P}(C_d), +, \cdot \rangle$ (which is also called \mathbb{B} -algebra of the cyclic group) where $\mathcal{P}(C_d)$ denotes the family of all subsets of C_d and for every pair of subsets A, B of C_d we set $A + B = A \cup B$ and $A \cdot B = \{a \cdot b \mid a \in A, b \in B\}$; hence \emptyset is the unit of the sum and $\{1\}$ is the unit of the product. Now, given a positive semiring \mathcal{S} , consider the map $\varphi_d : \mathcal{S}[x] \rightarrow \mathcal{B}_d$ which associates with any polynomial $F = \sum_k f_k x^k \in \mathcal{S}[x]$ the set $\{g^k \mid f_k \neq 0\} \in \mathcal{B}_d$. Intuitively, φ_d associates F with the set of its exponents modulo the integer d . Note that since the semiring \mathcal{S} is positive, φ_d is a semiring morphism. Of course φ_d extends to the semiring of $(Q \times Q)$ -matrices over $\mathcal{S}[x]$ by setting $\varphi_d(T)_{pq} = \varphi_d(T_{pq})$, for every matrix $T : Q \times Q \rightarrow \mathcal{S}[x]$ and all $p, q \in Q$. Observe that, since φ_d is a morphism, $(\varphi_d(T)^n)_{pq} = \varphi_d(T^n)_{pq} = \varphi_d(T^n)_{pq}$.

Now, let $M : Q \times Q \rightarrow \mathcal{S}[x]$ be an irreducible matrix with finite x -period d . Simply by the definition of d and φ_d , we have that for each $n \in \mathbb{N}$ all non–empty entries $\varphi_d(M^n)_{pp}$ have cardinality 1. The following results also concern the powers of $\varphi_d(M)$.

Proposition 5 *Let M be an irreducible matrix over $\mathcal{S}[x]$ with finite x -period d . Then, for each integer n and each pair of indices p and q , the set $\varphi_d(M^n)_{pq}$ contains at most one element; moreover, if $\varphi_d(M)_{qq} \neq \emptyset$ then $\varphi_d(M^n)_{qq} = (\varphi_d(M)_{qq})^n$.*

Proof. Let n be an arbitrary integer and p, q an arbitrary pair of indices. By the remarks above we may assume $p \neq q$ and $M^n_{pq} \neq 0$. The matrix M being irreducible, there exists an integer t such that $M^t_{qp} \neq 0$. Note that if B is a non–empty subset of C_d then $|A \cdot B| \geq |A|$ holds for each $A \subseteq C_d$ and $\varphi_d(M)^{n+t}_{pp} \supseteq \varphi_d(M)^n_{pq} \cdot \varphi_d(M)^t_{qp}$. Therefore, since $|\varphi_d(M)^{n+t}_{pp}| \leq 1$, we have also $|\varphi_d(M)^n_{pq}| \leq 1$. The second statement is proved in a similar way reasoning by induction on n . \square

Proposition 6 *Let M be an irreducible matrix over $\mathcal{S}[x]$ with finite x -period d . Then, for each integer n , all non-empty diagonal elements of $\varphi_d(M)^n$ are equal.*

Proof. Let n be an arbitrary integer and let p, q be an arbitrary pair of indices such that $M^n_{pp} \neq 0$ and $M^n_{qq} \neq 0$. By the previous proposition, there exist h, k such that $\varphi_d(M)^n_{pp} = \{g^h\}$ and $\varphi_d(M)^n_{qq} = \{g^k\}$. If t is defined as in the previous proof then both $\varphi_d(M)^t_{qp} \cdot \{g^h\}$ and $\{g^k\} \cdot \varphi_d(M)^t_{qp}$ are subsets of $\varphi_d(M)^{t+n}_{qp}$; since this set contains only one element they must be equal and this completes the proof. \square

Proposition 7 *Let M be a primitive matrix over $\mathcal{S}[x]$ with finite x -period d . Then, there exists an integer $0 \leq \gamma < d$ such that for each integer n and each index q , if $M^n_{qq} \neq 0$, then $\varphi_d(M)^n_{qq} = \{g^{\gamma n}\}$. Moreover, for each pair of indices p, q and for any integer n such that $M^n_{pq} \neq 0$, we have $\varphi_d(M^n_{pq}) = \{g^{\gamma n + \delta_{pq}}\}$ for a suitable integer $0 \leq \delta_{pq} < d$ independent of n .*

Proof. Since M is primitive, there exists an integer t such that $M^n_{pq} \neq 0$ for every $n \geq t$ and for every pair of indices p and q . In particular, since $dt + 1 > t$, we have $|\varphi_d(M^{dt+1}_{qq})| = 1$ for each q and hence there exists $0 \leq \gamma < d$ such that $\varphi_d(M)^{dt+1}_{qq} = \{g^\gamma\}$. Observe that γ does not depend on q , by Proposition 6. Therefore, by Proposition 5, we have

$$\{g^{\gamma n}\} = \varphi_d(M)^{dtn+n}_{qq} \supseteq \varphi_d(M)^{dtn}_{qq} \cdot \varphi_d(M)^n_{qq} = \{1\} \cdot \varphi_d(M)^n_{qq}$$

which proves the first part of the statement.

Now, consider an arbitrary pair of indices p, q and let t be the smallest positive integer such that $M^t_{pq} \neq 0$ (the existence of such t is guaranteed by the primitivity of M). Then, for each integer n , we have

$$\varphi_d(M)^t_{qp} \cdot \varphi_d(M)^n_{pq} \subseteq \varphi_d(M)^{n+t}_{qq} = \{g^{\gamma(n+t)}\}$$

Moreover, by Proposition 5 we know that there exists an exponent k such that $\varphi_d(M)^t_{qp} = \{g^k\}$. This yields the result, by taking the integer $0 \leq \delta_{pq} < d$ such that $\delta_{pq} = \gamma t - k$ modulo d . \square

If M is the a -counting matrix of a linear representation, then all propositions in this section can be interpreted by considering its state diagram. For any pair of states p, q , all paths of the same length starting in p and ending in q have the same number of occurrences of a modulo d . Secondly, if \mathcal{C}_k is a q_k -cycle for $k = 1, 2$ and \mathcal{C}_1 and \mathcal{C}_2 have the same length, then they also have the same number of occurrences of a modulo d . Finally, if M is primitive, for each cycle \mathcal{C} we have $|\mathcal{C}|_a = \gamma|\mathcal{C}|$ modulo d for some integer γ and moreover, for every pair of states p, q there exists a constant δ_{pq} such that the number of a in any path of length n from p to q is given by $\gamma n + \delta_{pq}$.

We conclude this section with an example showing that Proposition 7 cannot be extended to the case when M is irreducible but not primitive.

Example 5 Consider the a -counting matrix $M(x)$ associated with the state diagram of Figure 2. Then $M(x)$ is irreducible with x -period 2, but it is not primitive since also its period equals 2. Consider the path $\ell = q_1 \xrightarrow{b} q_2 \xrightarrow{a} q_1$ with $|\ell| = 2$ and $|\ell|_a = 1$. For any γ , $\gamma|\ell|$ cannot be equal to $|\ell|_a$ modulo 2. Thus, Proposition 7 does not hold in this case.

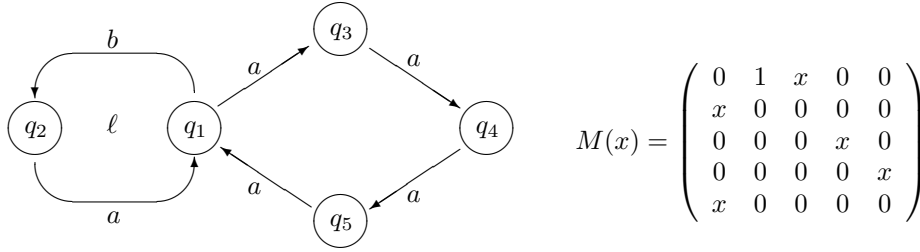


Figure 2: State diagram and matrix of Example 5.

3.3 Eigenvalues of x -periodic matrices

In this section we consider the semiring \mathbb{R}_+ of non-negative real numbers and we study the eigenvalues of primitive matrices $M(x)$ over $\mathbb{R}_+[x]$ when x assumes the complex values z such that $|z| = 1$. The next theorem shows how the eigenvalues of $M(z)$ are related to the x -period of the matrix. To this end we first give two auxiliary lemmata.

Lemma 8 *Let M be an irreducible matrix over $\mathcal{S}[x]$ with finite x -period d . Then for every index q there exist an integer n and two exponents h and k appearing in M^n_{qq} such that $h - k = d$. If further M is primitive, then the property holds for every n large enough.*

Proof. Consider an arbitrary integer q . By the definition of $d = d(q)$ there exists a finite set of integers s_j such that $d = \sum_j s_j(h_j - k_j)$, where h_j and k_j are exponents appearing in $M^{n_j}_{qq}$ for some integer n_j . Observe that, since h_j and k_j can be exchanged, we may assume positive all coefficients s_j . Now set $h = \sum_j s_j h_j$ and $k = \sum_j s_j k_j$. Then $d = h - k$ holds and both h and k are exponents in M^n_{qq} where $n = \sum_j s_j n_j$. This proves the statement if M is irreducible.

If further M is primitive, then there exists an integer t such that for each $m \geq t$, $M^m_{qq} \neq 0$ and hence M^m_{qq} has a non-null coefficient of degree l , for some $l \in \mathbb{N}$. Thus, $h + l$ and $k + l$ are exponents that appear in M^{n+m}_{qq} for each integer $m \geq t$ and this completes the proof. \square

Lemma 9 *Let $M(x)$ be a primitive matrix over $\mathbb{R}_+[x]$ with finite x -period d and set $M = M(1)$. Then, for every integer n large enough and for each $z \in \mathbb{C}$ such that $|z| = 1$, $|M(z)^n| = M^n$ if and only if z is a d -th root of unity.*

Proof. Given $n \in \mathbb{N}$ and a pair of indices p, q , let $M(x)^n_{pq} = \sum_{j=1}^l f_j x^{k_j}$. By Proposition 5, all the exponents k_1, k_2, \dots, k_l are congruent modulo d . Then, for every $z \in \mathbb{C}$ we can write

$$M(z)^n_{pq} = z^{k_1} (f_1 + \sum_{j=2}^l f_j z^{s_j})$$

where each $s_j = k_j - k_1$ is multiple of d , for $j = 2, \dots, l$. As a consequence, if $z = e^{\frac{2k\pi}{d}i}$ for some $k \in \mathbb{Z}$ then $M(z)^n_{pq} = z^{k_1} M^n_{pq}$ proving the result in one direction.

On the other hand, let n be an integer large enough to satisfy Lemma 8 for any index q and consider some diagonal entry $M(x)^n_{qq} = \sum_{j=1}^l f_j x^{k_j}$. By the previous lemma we may assume $d = k_2 - k_1$ and hence, setting $s_j = k_j - k_1$ we have $d = \text{GCD}\{s_j \mid j = 2, \dots, l\}$.

Now assume $|M(z)^n|_{qq} = \sum_j f_{k_j} = M^n|_{qq}$ for some $z = e^{i\theta}$ with $0 \leq \theta < 2\pi$. This implies that each θs_j is multiple of 2π and hence for all $j = 2, \dots, l$ we have

$$\frac{\theta}{2\pi} = \frac{p_j}{s_j} = \frac{h}{s} \quad (2)$$

where $p_j \in \mathbb{Z}$, $s = \text{LCM} \{s_j \mid j = 2, \dots, l\}$ and $h < s$ is a non-negative integer. Since h is multiple of each s/s_j it is also multiple of $s' = \text{LCM} \{s/s_j \mid j = 2, \dots, l\}$. Now $\text{GCD} \{s_j \mid j = 2, \dots, l\} = d$ and hence we have $s' = s/d$. Thus, being $\theta = 2\pi h/s$ by (2), we have that θ is a multiple of $2\pi/d$ and hence $z = e^{i\theta}$ is a d -th root of unity. \square

Theorem 10 *Let $M(x)$ be a primitive matrix over $\mathbb{R}_+[x]$ with finite x -period d , set $M = M(1)$ and let λ be the Perron–Frobenius eigenvalue of M . Then, for all $z \in \mathbb{C}$ with $|z| = 1$, the following conditions are equivalent:*

1. $M(z)$ and M have the same set of moduli of eigenvalues;
2. If $\lambda(z)$ is an eigenvalue of maximum modulus of $M(z)$, then $|\lambda(z)| = \lambda$;
3. z is a d -th root of unity in \mathbb{C} .

Proof. Clearly Condition 1) implies Condition 2). To prove that Condition 2) implies Condition 3) we reason by contradiction, that is we assume that z is not a d -th root of unity. By Lemma 9 in this case there exists an integer n such that $|M(z)^n| \neq M^n$. Therefore we can apply Proposition 2 and prove that λ^n is greater than the modulus of any eigenvalue of $M(z)^n$. In particular we have $\lambda^n > |\lambda(z)|^n$ which contradicts the hypotheses.

Finally we show that Condition 3) implies Condition 1). The case $d = 1$ is trivial; thus suppose $d > 1$ and assume that z is a d -th root of unity. It suffices to prove that if ν is an eigenvalue of M , then νz^γ is an eigenvalue of $M(z)$ with the same multiplicity, where γ is the constant introduced in Proposition 7. Indeed, set $\hat{T} = I\nu z^\gamma - M(z)$ and $T = I\nu - M$. We now verify that $\text{Det } \hat{T} = z^{\gamma m} \text{Det } T$ holds where m is the size of M . To prove this equality, recall that

$$\text{Det } \hat{T} = \sum_{\rho} (-1)^{\sigma(\rho)} \hat{T}_{1\rho(1)} \cdots \hat{T}_{m\rho(m)},$$

where the sum runs over all the permutations ρ of $\{1, 2, \dots, m\}$. By Proposition 7, since z is a d -th root of 1 in \mathbb{C} , we have $\hat{T}_{qq} = (\nu - M_{qq})z^\gamma = z^\gamma T_{qq}$ for each state q and $\hat{T}_{q_0 q_1} \cdots \hat{T}_{q_{s-1} q_0} = z^{\gamma s} T_{q_0 q_1} \cdots T_{q_{s-1} q_0}$ for each simple cycle $(q_0, q_1, \dots, q_{s-1}, q_0)$ of length $s > 1$. Therefore, for each permutation ρ , we get

$$\hat{T}_{1\rho(1)} \cdots \hat{T}_{m\rho(m)} = z^{\gamma m} \cdot T_{1\rho(1)} \cdots T_{m\rho(m)}$$

which concludes the proof. \square

Example 6 Let us consider again the primitive matrix of Figure 1. We recall that here $d = 4$; moreover it is easy to see that $\gamma = 3$. Indeed, for each $k = 1, 2$, we have that $|\ell_k| - 3|\ell_k|_a$ is equal to 0 modulo 4. Now consider the characteristic polynomial of the matrix $M(x)$, given by $\chi_x(y) = y^4 - y^2 x^2 - yx$ and let ν be a root of χ_1 . This implies that $\chi_1(\nu) = \nu^4 - \nu^2 - \nu = 0$ and hence $-i\nu$ is a root of the polynomial χ_i , $-\nu$ is a root of the polynomial χ_{-1} and $i\nu$ is a root of the polynomial χ_{-i} . This is consistent with Theorem 10, since $1, i, -1$ and $-i$ are the four roots of unity. \square

4 Global and local convergence properties

In this section we recall some basic notions in probability theory and we focus more specifically on the central and local limit theorems for sequences of discrete random variables.

Let X be a random variable (r.v.) with values in a set $\{x_0, x_1, \dots, x_k, \dots\}$ of real numbers and set $p_k = \Pr\{X = x_k\}$, for every $k \in \mathbb{N}$. We denote by F_X its distribution function, i.e., $F_X(\tau) = \Pr\{X \leq \tau\}$ for every $\tau \in \mathbb{R}$. If the set of indices $\{k \mid p_k \neq 0\}$ is finite we can consider the moment generating function of X , given by

$$\Psi_X(z) = \mathbb{E}(e^{zX}) = \sum_{k \in \mathbb{N}} p_k e^{zx_k},$$

which is well-defined for every $z \in \mathbb{C}$. This function can be used to compute the first two moments of X ,

$$\mathbb{E}(X) = \Psi'_X(0), \quad \mathbb{E}(X^2) = \Psi''_X(0),$$

and to prove convergence in distribution. We recall that, given a sequence of random variables $\{X_n\}_n$ and a random variable X , X_n converges to X *in distribution* (or *in law*) if $\lim_{n \rightarrow \infty} F_{X_n}(\tau) = F_X(\tau)$ for every point $\tau \in \mathbb{R}$ of continuity for F_X . It is well known that if Ψ_{X_n} and Ψ_X are defined all over \mathbb{C} and $\Psi_{X_n}(z)$ tends to $\Psi_X(z)$ for every $z \in \mathbb{C}$, then X_n converges to X in distribution [7]. We also recall that the restriction $\Psi_X(it)$, for $t \in \mathbb{R}$, is the characteristic function of X , which is always defined all over \mathbb{R} and completely characterizes the distribution of X .

A convenient approach to prove the convergence in law to a Gaussian random variable relies on the so called “quasi-power” theorems introduced in [10] and implicitly used in the previous literature [1] (see also [6]). For our purpose we present the following simple variant of such a theorem.

Theorem 11 *Let $\{X_n\}$ be a sequence of random variables, where each X_n takes values in $\{0, 1, \dots, n\}$ and let us assume the following conditions:*

C1 *There exist two functions $r(z)$, $y(z)$, both analytic at $z = 0$ where they take the value $r(0) = y(0) = 1$, and a positive constant c , such that for every $|z| < c$*

$$\Psi_{X_n}(z) = r(z) \cdot y(z)^n (1 + O(n^{-1})); \quad (3)$$

C2 *The constant $\sigma = y''(0) - (y'(0))^2$ is strictly positive (variability condition).*

Also set $\mu = y'(0)$. Then $\frac{X_n - \mu n}{\sqrt{\sigma n}}$ converges in distribution to a normal random variable of mean 0 and variance 1, i.e., for every $x \in \mathbb{R}$

$$\lim_{n \rightarrow +\infty} \Pr \left\{ \frac{X_n - \mu n}{\sqrt{\sigma n}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

The main advantage of this theorem, with respect to other classical statements of this kind, is that it does not require any condition of independence concerning the random variables X_n . For instance, the standard central limit theorems assume that each X_n is a partial sum of the form $X_n = \sum_{j \leq n} U_j$, where the U_j 's are independent random variables [7].

The following proposition provides a useful approximation of the function $y(z)$ defined in the previous theorem.

Proposition 12 Let $\{X_n\}$ be a sequence of random variables such that each X_n takes values in the set $\{0, 1, \dots, n\}$. Assume that Conditions **C1** and **C2** of Theorem 11 hold true and let $y(z)$, μ and σ be defined consequently. Then, for every real θ such that $|\theta| \leq n^{-5/12}$, as n grows to infinity, we have

$$\left| y(i\theta)^n - e^{-(\sigma/2)\theta^2 n + i\mu\theta n} \right| = O\left(n^{-1/2}\right).$$

Proof. First of all observe that, from our hypothesis, in a neighbourhood of $z = 0$ we have

$$y(z) = 1 + \mu z + \frac{\sigma + \mu^2}{2} z^2 + O(z^3) \quad (4)$$

This implies that, in a real neighbourhood of $\theta = 0$, the complex function $y(i\theta)$ satisfies the equalities

$$\begin{aligned} |y(i\theta)| &= \left| 1 + i\mu\theta - \frac{\sigma + \mu^2}{2} \theta^2 \right| \cdot |1 + O(\theta^3)| = \\ &= \sqrt{\left(1 - \frac{\sigma + \mu^2}{2} \theta^2\right)^2 + \mu^2 \theta^2} \cdot |1 + O(\theta^3)| = \left(1 - \frac{\sigma}{2} \theta^2 + O(\theta^4)\right) |1 + O(\theta^3)| \quad (5) \\ \arg(y(i\theta)) &= \arg\left(1 + i\mu\theta - \frac{\sigma + \mu^2}{2} \theta^2\right) + \arg(1 + O(\theta^3)) = \\ &= \operatorname{arctg}\left(\frac{\mu\theta}{1 - \frac{\sigma + \mu^2}{2} \theta^2}\right) + O(\theta^3) = \mu\theta + O(\theta^3). \end{aligned}$$

As a consequence, one has

$$y(i\theta)^n = \left(1 - (\sigma/2)\theta^2 + O(\theta^3)\right)^n \cdot e^{in(\mu\theta + O(\theta^3))} = e^{-(\sigma/2)n\theta^2 + i\mu n\theta} \cdot e^{nO(\theta^3)}$$

Now, for each $|\theta| \leq n^{-5/12}$, we have $|n\theta^3| = O(n^{-1/4})$ and the last expression yields

$$y(i\theta)^n = e^{-(\sigma/2)n\theta^2 + i\mu n\theta} \cdot (1 + O(n\theta^3))$$

Therefore

$$\left| y(i\theta)^n - e^{-(\sigma/2)n\theta^2 + i\mu n\theta} \right| = O\left(n|\theta^3|e^{-n\theta^2\sigma/2}\right) = O\left(n^{-1/2}\right),$$

the last equality being obtained by taking the maximum with respect to θ . \square

4.1 A general criterion for local convergence laws

Convergence in law of a sequence of r.v.'s $\{X_n\}$ does not yield an approximation of the probability that X_n has a specific value. Theorems concerning approximations for expressions of the form $\Pr\{X_n = x\}$ are usually called *local limit theorems* and often give a stronger property than a traditional convergence in distribution². A typical example is given by the so-called de Moivre–Laplace Local Limit Theorem [7], which intuitively states that, for certain

²For this reason, theorems showing convergence in distribution of a sequence of r.v.'s are sometimes called *global or integral limit theorems*.

sequences of binomial random variables X_n , up to a factor $\Theta(1/\sqrt{n})$ the probability that X_n takes on a value x approximates a Gaussian density at x .

In this section we present a general criterion that guarantees, for a sequence of integer random variables, the existence of a local convergence property of a Gaussian type more general than the de Moivre–Laplace theorem mentioned above. In the subsequent section, using such criterion, we show that the same local convergence property holds for certain pattern statistics.

Theorem 13 (Local Limit Criterion) *Let $\{X_n\}$ be a sequence of random variables such that, for some integer $d \geq 1$ and every $n \geq d$, X_n takes on values only in the set*

$$\{x \in \mathbb{N} \mid 0 \leq x \leq n, x \equiv \rho \pmod{d}\} \quad (6)$$

for some integer $0 \leq \rho < d$. Assume that Conditions **C1** and **C2** of Theorem 11 hold true and let μ and σ be the positive constants defined therein. Moreover assume the following property:

$$\mathbf{C3} \text{ For all } 0 < \theta_0 < \pi/d \quad \lim_{n \rightarrow +\infty} \left\{ \sqrt{n} \sup_{|\theta| \in [\theta_0, \pi/d]} |\Psi_{X_n}(i\theta)| \right\} = 0$$

Then, as n grows to $+\infty$ the following relation holds uniformly for every $j = 0, 1, \dots, n$.

$$\Pr\{X_n = j\} = \begin{cases} \frac{de^{-\frac{(j-\mu n)^2}{2\sigma n}}}{\sqrt{2\pi\sigma n}} \cdot (1 + o(1)) & \text{if } j \equiv \rho \pmod{d} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Observe that $\Psi_{X_n}(i\theta)$ is the characteristic function of X_n , it is periodic of period 2π and it assumes the value 1 at $\theta = 0$. Condition **C3** states that, for every constant $0 < \theta_0 < \pi/d$, as n grows to $+\infty$, the value $\Psi_{X_n}(i\theta)$ is of the order $o(n^{-1/2})$ uniformly with respect to $\theta \in [-\pi/d, -\theta_0] \cup [\theta_0, \pi/d]$. Note that ρ may depend on n even if $\rho = \Theta(1)$.

One can easily show that any sequence $\{X_n\}$ of binomial r.v.'s of parameters n and p , where $0 < p < 1$ (i.e., representing the number of successes over n independent trials of probability p), satisfies the hypothesis of the theorem with $d = 1$. In this case Eq. 7 coincides with the property stated in the de Moivre–Laplace Local Limit Theorem. Thus our general criterion includes the same theorem as a special case.

Periodic phenomena of the form (7) have already been studied in the literature. For instance, in [7, Section 43], Equation (7) is proved when X_n is the sum of n independent lattice r.v.'s of period d and equal distribution. Note that our theorem does not require any condition of independence of the X_n 's.

We also note that, for $d = 1$, similar criteria for local limit laws have been proposed in [6, Theorem 9.10] and [11] where, however, different conditions are assumed. In [11], a more precise evaluation of $\Pr\{X_n = j\}$ is obtained under hypotheses stronger than ours; in particular Condition **C3** is replaced by an exponential bound like $|\Psi_{X_n}(z)| = O(e^{-cn})$ where $c > 0$ is constant and z uniformly varies in a finite strip around the imaginary axis.

Here we present the proof of Theorem 13: even though we use some ideas already appeared in [3, Section 5] the present approach is much more general: we drop any rationality

hypothesis on the distribution of the r.v.'s X_n and only rely on Conditions **C1**, **C2**, **C3**, together with the assumption that each X_n takes values in a set of the form (6). Let us first prove the following proposition concerning the characteristic function of the random variable X_n .

Proposition 14 *Under the hypotheses of Theorem 13, for every real θ such that $|\theta| \in [0, \pi/d]$ we have*

$$\left| \Psi_{X_n}(i\theta) - e^{-(\sigma/2)\theta^2 n + i\mu\theta n} \right| = \Delta_n(\theta),$$

where

$$\Delta_n(\theta) = \begin{cases} O(n^{-5/12}) & \text{if } |\theta| \in [0, n^{-5/12}] \\ o(n^{-1/2}) & \text{if } |\theta| \in [n^{-5/12}, \pi/d] \end{cases} \quad (8)$$

Proof. For the sake of brevity let Ψ_n stand for Ψ_{X_n} . Let us consider the first interval given in (8), i.e., the case $|\theta| \leq n^{-5/12}$. By Condition **C1** of Theorem 11 and Proposition 12 we have

$$\Psi_n(i\theta) = r(i\theta)y(i\theta)^n(1 + O(n^{-1})) = \left(e^{-(\sigma/2)\theta^2 n + i\mu\theta n} + O(n^{-1/2}) \right) (1 + O(n^{-5/12}))$$

which proves the relation since $|e^{-(\sigma/2)\theta^2 n + i\mu\theta n}| \leq 1$ for every real θ .

As far as the second interval is concerned, let θ_0 be a constant such that $0 < \theta_0 < c$, where c is defined as in Condition **C1** of Theorem 11 and assume $|\theta| \in [n^{-5/12}, \theta_0]$. Then we have

$$\left| \Psi_n(i\theta) - e^{-(\sigma/2)\theta^2 n + i\mu\theta n} \right| \leq |\Psi_n(i\theta)| + e^{-(\sigma/2)\theta^2 n}.$$

Since $|\theta| \geq n^{-5/12}$, the second term of the right hand side is smaller than or equal to $e^{-(\sigma/2)n^{1/6}} = o(n^{-1/2})$. Let us show an analogous bound for the first term. To this end, by Equations (3) and (5) we have

$$|\Psi_n(i\theta)| = |r(i\theta)y(i\theta)^n(1 + O(n^{-1}))| \leq r \left| 1 - \frac{\sigma}{2}\theta^2 + O(\theta^3) \right|^n |1 + O(n^{-1})| \quad (9)$$

where $r = \sup_{|\theta| \leq \theta_0} |r(i\theta)|$. By the arbitrariness of θ_0 , for some constant $\varepsilon > 0$ and every $|\theta| \leq \theta_0$ we have

$$\left| 1 - \frac{\sigma}{2}\theta^2 + O(\theta^3) \right| \leq \left| 1 - \frac{\sigma}{2}\theta^2 \right| + \varepsilon\theta^3.$$

By the same reason we may assume $\theta_0 \leq \min\{\sqrt{2/\sigma}, \sigma/(4\varepsilon)\}$, which proves

$$\left| 1 - \frac{\sigma}{2}\theta^2 + O(\theta^3) \right| \leq 1 - \frac{\sigma}{2}\theta^2 + \varepsilon\theta_0\theta^2 \leq 1 - \frac{\sigma}{4}\theta^2,$$

for every $|\theta| \leq \theta_0$. Replacing this value in (9) we get

$$|\Psi_n(i\theta)| = O\left(\left|1 - (\sigma/4)\theta^2\right|^n\right) = O\left(e^{-(\sigma/4)\theta^2 n}\right),$$

which is again bounded by $O\left(e^{-(\sigma/2)n^{1/6}}\right) = o(n^{-1/2})$ because of the range of θ . This proves relation (8) for $|\theta| \in [n^{-5/12}, \theta_0]$.

Finally assume $\theta_0 \leq |\theta| \leq \pi/d$. Again, we have

$$\left| \Psi_n(i\theta) - e^{-(\sigma/2)\theta^2 n + i\mu\theta n} \right| \leq \sup_{\theta_0 \leq |\theta| \leq \pi/d} |\Psi_n(i\theta)| + e^{-(\sigma/2)\theta_0^2 n}.$$

The first term is bounded by $o(n^{-1/2})$ by Condition **C3**, while the second one is $O(\tau^n)$ for some $\tau \in (0, 1)$ and this completes the proof. \square

Proof of Theorem 13. First, we apply the Discrete Fourier Transform (see for instance [5]) to the array of probabilities of X_n . Since each X_n assumes values only in (6), set $N = \min\{h \in \mathbb{N} \mid n < \rho + hd\}$ and define $p^{(n)}$ as the array $(p_0^{(n)}, p_1^{(n)}, \dots, p_{N-1}^{(n)})$, where

$$p_h^{(n)} = \Pr\{X_n = \rho + hd\} \quad (h = 0, 1, \dots, N-1).$$

Let $f^{(n)} \in \mathbb{C}^N$ be its Discrete Fourier Transform, i.e., the array of values $f_s^{(n)}$ such that

$$f_s^{(n)} = \sum_{h=0}^{N-1} p_h^{(n)} e^{i \frac{2\pi s}{N} h} = \Psi_{X_n} \left(i \frac{2\pi s}{Nd} \right) e^{-i \frac{2\pi s \rho}{Nd}}$$

where $s = -\lceil N/2 \rceil + 1, \dots, \lfloor N/2 \rfloor$. By antitransforming, each $p_h^{(n)}$ can be obtained by

$$p_h^{(n)} = \frac{1}{N} \sum_{s=-\lceil N/2 \rceil + 1}^{\lfloor N/2 \rfloor} f_s^{(n)} e^{-i \frac{2\pi s}{N} h} = \frac{1}{N} \sum_{s=-\lceil N/2 \rceil + 1}^{\lfloor N/2 \rfloor} \Psi_{X_n} \left(i \frac{2\pi s}{Nd} \right) e^{-i \frac{2\pi s}{Nd} (\rho + hd)}. \quad (10)$$

Now, the previous proposition suggests us to define the function $F_n(\theta) = e^{-(\sigma/2)\theta^2 n + i\mu\theta n}$ for every $-\pi/d < \theta \leq \pi/d$ and to approximate $p_h^{(n)}$ with the following values:

$$\hat{p}_h^{(n)} = \frac{1}{N} \sum_{s=-\lceil N/2 \rceil + 1}^{\lfloor N/2 \rfloor} F_n \left(\frac{2\pi s}{Nd} \right) e^{-i \frac{2\pi s}{Nd} (\rho + hd)}. \quad (11)$$

Clearly, the error associated to the above approximation satisfies the inequality

$$\left| p_h^{(n)} - \hat{p}_h^{(n)} \right| \leq \frac{1}{N} \sum_{s=-\lceil N/2 \rceil + 1}^{\lfloor N/2 \rfloor} \left| \Psi_{X_n} \left(i \frac{2\pi s}{Nd} \right) - F_n \left(\frac{2\pi s}{Nd} \right) \right| \leq \frac{1}{N} \sum_{s=-\lceil N/2 \rceil + 1}^{\lfloor N/2 \rfloor} \Delta_n \left(\frac{2\pi s}{Nd} \right),$$

which can be computed by splitting the range of s in two intervals as in (8). Thus, we get

$$\left| p_h^{(n)} - \hat{p}_h^{(n)} \right| \leq \frac{2}{N} \left\{ \lceil Nd/(2\pi n^{5/12}) \rceil O(n^{-5/12}) + \lceil N/2 \rceil o(n^{-1/2}) \right\} = o(n^{-1/2}). \quad (12)$$

As n grows to $+\infty$ the right hand side of (11) tends to the integral of $F_n(x) e^{-ix(\rho+hd)} d/(2\pi)$ over the interval $x \in (-\pi/d, \pi/d)$. Thus, by standard mathematical tools (as in [3, Sec. 5.3]), one can prove that as n grows to $+\infty$ the relation

$$\hat{p}_h^{(n)} = \frac{d}{\sqrt{2\pi\sigma n}} e^{-\frac{(\rho + hd - \mu n)^2}{2\sigma n}} + o(n^{-1/2})$$

holds uniformly for every $h = 0, 1, \dots, N-1$. Hence, the result is a straightforward consequence of the previous equation, together with relation (12).

5 Pattern statistics in rational models

In this section we turn our attention to sequences of random variables defined by means of rational formal series in two noncommuting variables. We recall definitions and properties introduced in [3].

Let us consider the binary alphabet $\{a, b\}$ and, for $n \in \mathbb{N}$, let $\{a, b\}^n$ denote the set of all words of length n in $\{a, b\}^*$. Given a formal series $r \in \mathbb{R}_+ \langle\langle a, b \rangle\rangle$, let n be a positive integer such that $(r, x) \neq 0$ for some $x \in \{a, b\}^n$. Consider the probability space of all words in $\{a, b\}^n$ equipped with the probability measure given by

$$\Pr\{\omega\} = \frac{(r, \omega)}{\sum_{x \in \{a, b\}^n} (r, x)} \quad (\omega \in \{a, b\}^n). \quad (13)$$

In particular, if r is the characteristic series χ_L of a language $L \subseteq \{a, b\}^*$, then \Pr is just the uniform distribution over the set of words on length n in L : $\Pr\{\omega\} = \#(L \cap \{a, b\}^n)^{-1}$ if $\omega \in L$, while $\Pr\{\omega\} = 0$ otherwise. We define the random variable $Y_n : \{a, b\}^n \rightarrow \{0, 1, \dots, n\}$ such that $Y_n(\omega) = |\omega|_a$ for every $\omega \in \{a, b\}^n$. Then, for every $j = 0, 1, \dots, n$, we have

$$\Pr\{Y_n = j\} = \frac{\sum_{|\omega|=n, |\omega|_a=j} (r, \omega)}{\sum_{x \in \{a, b\}^n} (r, x)}. \quad (14)$$

For sake of brevity, we say that Y_n counts the occurrences of a in the stochastic model defined by r . 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.

A useful tool to study the distribution of the pattern statistics Y_n is given by the generating functions associated with formal series. Given $r \in \mathbb{R}_+ \langle\langle a, b \rangle\rangle$, for every $n, j \in \mathbb{N}$ let $r_{n,j}$ be the coefficient of $a^j b^{n-j}$ in the commutative image of r , i.e.,

$$r_{n,j} = (\mathcal{F}(r), a^j b^{n-j}) = \sum_{|x|=n, |x|_a=j} (r, x).$$

Then, we define the function $r_n(z)$ and the corresponding generating function $\mathbf{r}(z, w)$ by

$$r_n(z) = \sum_{j=0}^n r_{n,j} e^{jz} \quad \text{and} \quad \mathbf{r}(z, w) = \sum_{n=0}^{+\infty} r_n(z) w^n = \sum_{n=0}^{+\infty} \sum_{j=0}^n r_{n,j} e^{jz} w^n$$

where z, w are complex variables. Thus, from the definition of $r_{n,j}$ and from Equation (14) we have

$$\Pr\{Y_n = j\} = \frac{r_{n,j}}{r_n(0)} \quad (15)$$

and the moment generating function of Y_n is given by

$$\Psi_{Y_n}(z) = \sum_{j=0}^n \Pr\{Y_n = j\} e^{jz} = \frac{r_n(z)}{r_n(0)} \quad (16)$$

Moreover, we remark that the relation between a series r and its generating function $\mathbf{r}(z, w)$ can be expressed in terms of a semiring morphism. Consider the monoid morphism

$$\mathcal{H} : \{a, b\}^* \longrightarrow \{e^z, w\}^\otimes$$

defined by setting $\mathcal{H}(a) = e^z w$ and $\mathcal{H}(b) = w$. Then, such a map extends to a semiring morphism from $\mathbb{R}_+ \langle\langle a, b \rangle\rangle$ to $\mathbb{R}_+[[e^z, w]]$ so that

$$\mathcal{H}(r) = \mathbf{r}(z, w) \quad (17)$$

for every $r \in \mathbb{R}_+ \langle\langle a, b \rangle\rangle$. This property translate arithmetic relations among formal series into analogous relations among the corresponding generating functions.

When r is rational, the probability spaces given by (13) define a stochastic model (called *rational* stochastic model) of interest for the analysis of pattern statistics. A typical goal in that context is to estimate the limit distribution of the number of occurrences of patterns in a word of length n generated at random according to a given probabilistic model (usually a Markovian process [14]). In the rational model, the pattern is reduced to a single letter a . However, the analysis of Y_n in such a model includes as a particular case the study of the frequency of occurrences of regular patterns in words generated at random by a Markovian process [3, Sec. 2.1].

The limit distribution of Y_n is studied in [3] in the global sense, assuming that r admits a primitive linear representation (ξ, μ, η) . Set $A = \mu(a)$ and $B = \mu(b)$ and let λ denote the Perron–Frobenius eigenvalue of the matrix $A + B$. Then it is easy to see that in this case

$$\mathbf{r}(z, w) = \xi^T (I - w(Ae^z + B))^{-1} \eta. \quad (18)$$

It turns out that Y_n has a Gaussian limit distribution [3, Theorem 4], and this extends a similar result, earlier presented in [14] for pattern statistics in a Markovian model. The proof of this property is based on the quasi-power theorem and in particular one can verify that Condition **C1** and **C2** of Theorem 11 hold true.

5.1 Pattern statistics in the power model

In this section, we consider a stochastic model defined by the *power* of any primitive rational series (note that in this case the model is not primitive anymore) and we study the central and local behaviour of the associated pattern statistics Y_n . The results we obtain extend the analysis developed in [3] concerning the primitive rational stochastic models (in particular, in case $k = 1$, statement **T1** in Theorem 15 reduces to [3, Theorem 4]). They also extend some results presented in [4], where the (global) limit distribution of Y_n is determined whenever r is the product of two primitive formal series.

Theorem 15 *For any positive integer k and any primitive nondegenerate $r \in \mathbb{R}_+^{Rat} \langle\langle a, b \rangle\rangle$, let s be defined by $s = r^k$ and let Y_n count the occurrences of a in the stochastic model defined by s . Then the following properties hold true.*

T1 *There exist two constants α and β , satisfying $0 < \alpha$ and $0 < \beta < 1$, such that $\frac{Y_n - \beta n}{\sqrt{\alpha n}}$ converges in distribution to a normal random variable of mean value 0 and variance 1.*

T2 *If (ξ, μ, η) is a primitive linear representation for r and d is the x -period of $\mu(a)x + \mu(b)$, then there exist d functions $C_i : \mathbb{N} \rightarrow \mathbb{R}_+$, $i = 0, 1, \dots, d-1$, such that $\sum_i C_i(n) = 1$ for every $n \in \mathbb{N}$ and further, as n grows to $+\infty$, the relation*

$$\Pr\{Y_n = j\} = \frac{d C_{\langle j \rangle_d}(n)}{\sqrt{2\pi\alpha n}} e^{-\frac{(j - \beta n)^2}{2\alpha n}} \cdot (1 + o(1)) \quad (19)$$

holds uniformly for every $j = 0, 1, \dots, n$ (here $\langle j \rangle_d = j - \lfloor j/d \rfloor d$).

Before proving Theorem 15, let us illustrate its meaning by an example.

Example 7 Consider the series $s \in \mathbb{R}_+^{Rat}\langle\langle a, b \rangle\rangle$ defined by the linear representation (ξ_s, μ_s, η_s) where

$$A_s = \mu_s(a) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad B_s = \mu_s(b) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$\xi_s = (1, 0, 0, 0, 0, 0)_T$ and $\eta_s = (0, 0, 0, 0, 2, 1)_T$. Also, let Y_n count the occurrences of the symbol a in the stochastic model defined by s . Using Equation (18) one can compute the bivariate function $\mathbf{s}(z, w)$ associated with the series s , so obtaining

$$\mathbf{s}(z, w) = \frac{w^2(e^z w + 2)^2}{(w^2 - e^{2z} w^2 - 2w + 1)^2}.$$

Then, the coefficients in the expansion $\mathbf{s}(z, w) = \sum_{n,j} s_{n,j} e^{jz} w^n$ can be computed by decomposing the rational function $\mathbf{s}(z, w)$ in simple fractions and using basic series expansions. For each $0 \leq j \leq n$ we get

$$s_{n,j} = \begin{cases} \frac{1}{2} \binom{n}{j} \frac{(n-j)(n-j-1)}{n(j+1)(j+3)} (4n^2 + 4n + j^2 + 3j) & \text{if } j \text{ is even} \\ 2 \binom{n}{j} \frac{(n-j)(n-j-1)}{j+2} & \text{if } j \text{ is odd.} \end{cases}$$

Thus, the probability function of Y_n can be obtained via normalization, by recalling Equation (15). Such a function is represented for $n = 800$ in Figure 3, where one can recognize the superimposition of two Gaussian behaviours with common mean value and variance, multiplied by suitable constants.

This behaviour is explained by statement **T2** of Theorem 15. Indeed, one can verify that the series s is the square of a formal series in $\mathbb{R}_+^{Rat}\langle\langle a, b \rangle\rangle$ that admits a primitive linear representation with x -period $d = 2$, namely

$$\xi = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \eta = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

□

We now turn our attention to the proof of Theorem 15. We split the proof in two separate parts and we use the criteria presented in Theorem 11 and in Theorem 13, observing that by Equation (16) the characteristic function of Y_n is given by $s_n(z)/s_n(0)$. We still use the notation introduced in the previous section: set $A = \mu(a)$, $B = \mu(b)$, $M = A + B$ and let λ be the Perron Frobenius eigenvalue of the primitive matrix M .

Proof of T1. Since $s = r^k$, by applying the morphism \mathcal{H} defined in (17) we get

$$\mathbf{s}(z, w) = \mathbf{r}(z, w)^k.$$

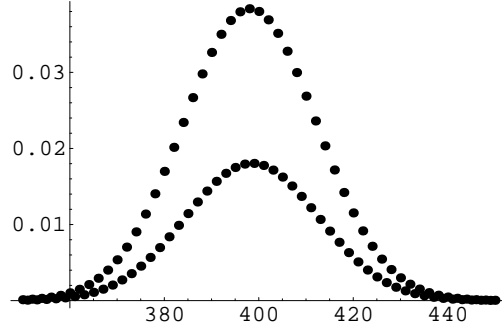


Figure 3: Plot of the probability function $P_n\{Y_n = j\}$ obtained in Example 7, for $n = 800$ and $350 \leq j \leq 450$. The limit behaviour is given by the superimposition of two alternating Gaussian densities.

From Equation (18), since $A + B$ is primitive and both A and B are non-null, one can show [3, Section 4] that near the point $(0, \lambda^{-1})$ the function $\mathbf{r}(z, w)$ admits a Laurent expansion of the form

$$\mathbf{r}(z, w) = \frac{R(z)}{1 - u(z)w} + O(1)$$

where $R(z)$ and $u(z)$ are complex functions, they are non-null and analytic at $z = 0$ and moreover $u(0) = \lambda$. As a consequence, in a neighbourhood of $(0, \lambda^{-1})$ we have

$$\mathbf{s}(z, w) = \left(\frac{R(z)}{1 - u(z)w} \right)^k + O\left(\frac{1}{1 - u(z)w} \right)^{k-1}$$

and hence the associated sequence is of the form

$$s_n(z) = R(z)^k \binom{n+k-1}{k-1} u(z)^n + O(n^{k-2} u(z)^n).$$

Hence, in a neighbourhood of $z = 0$, the characteristic function Ψ_{Y_n} can be expanded in the form

$$\Psi_{Y_n}(z) = \frac{s_n(z)}{s_n(0)} = \left(\frac{R(z)}{R(0)} \right)^k \cdot \left(\frac{u(z)}{\lambda} \right)^n \cdot (1 + O(n^{-1})).$$

Finally set $\alpha = u''(0)/\lambda - \beta^2$ and $\beta = u'(0)/\lambda$. Then α and β turn out to be strictly positive and furthermore they can be expressed as function of the matrix M and its eigenvectors [3, Sections 3 and 4]. Thus, Y_n satisfies both conditions of Theorem 11 with $r(z) = (R(z)/R(0))^k$, $y(z) = u(z)/\lambda$, $\mu = \beta$ and $\sigma = \alpha$. This proves the result. \square

Proof of T2. For every $p, q \in \{1, 2, \dots, m\}$, let $r^{(pq)}$ be the series defined by the linear representation $(\xi_p e_p, \mu, \eta_q e_q)$, where e_i is the characteristic array of entry i . Then

$$r = \sum_{p,q=1}^m r^{(pq)}$$

Thus, since $s = r^k$, we have

$$s = \sum_{*} r^{(p_1 q_1)} \cdot r^{(p_2 q_2)} \dots r^{(p_k q_k)} \quad (20)$$

where the sum is over all sequences $\ell = p_1 q_1 p_2 q_2 \dots p_k q_k \in \{1, 2, \dots, m\}^{2k}$. For sake of brevity, for every such ℓ , let $r^{(\ell)}$ be the series

$$r^{(\ell)} = r^{(p_1 q_1)} \cdot r^{(p_2 q_2)} \dots r^{(p_k q_k)}$$

By the primitivity hypothesis this series is identically null if and only if $\xi_{p_j} = 0$ or $\eta_{q_j} = 0$ for some $j \in \{1, 2, \dots, k\}$. For this reason set $Supp = \{\ell \in \{1, 2, \dots, m\}^{2k} \mid r^{(\ell)} \neq 0\}$. Then, for every $\ell \in Supp$, applying the morphism \mathcal{H} to the previous equation and recalling (18), we have

$$\mathbf{r}^{(\ell)}(0, w) = \prod_{i=1}^k \mathbf{r}^{(p_i q_i)}(0, w) = \prod_{i=1}^k \xi_{p_i} (I - wM)_{p_i q_i}^{-1} \eta_{q_i}$$

with the obvious meaning for the notations $\mathbf{r}^{(\ell)}$ and $\mathbf{r}^{(p_i q_i)}$. This implies, by the primitivity of M , that $\mathbf{r}^{(\ell)}(0, w)$ has a unique pole of smallest modulus at λ^{-1} , which has degree k . As a consequence, the sequence associated with $\mathbf{r}^{(\ell)}(0, w)$ satisfies the following relation

$$r_n^{(\ell)}(0) = c_\ell n^{k-1} \lambda^n + \mathcal{O}(n^{k-2} \lambda^n) \quad (21)$$

for some $c_\ell > 0$. Moreover, from (20) we have $s = \sum_{\ell \in Supp} r^{(\ell)}$ and hence

$$s_{n,j} = \sum_{\ell \in Supp} r_{n,j}^{(\ell)} \quad \text{and} \quad s_n(0) = \sum_{\ell \in Supp} r_n^{(\ell)}(0) = c n^{k-1} \lambda^n + \mathcal{O}(n^{k-2} \lambda^n) \quad (22)$$

for some $c > 0$. Then recalling (16) and (15), for every $j \in \{0, 1, \dots, n\}$ we have

$$\Pr\{Y_n = j\} = \frac{s_{n,j}}{s_n(0)} = \sum_{\ell \in Supp} \frac{r_{n,j}^{(\ell)}}{s_n(0)} = \sum_{\ell \in Supp} \frac{r_n^{(\ell)}(0)}{s_n(0)} \Pr\{Y_n^{(\ell)} = j\}$$

where $Y_n^{(\ell)}$ denotes the r.v. counting the occurrences of a in the model defined by $r^{(\ell)}$. Finally, from Equations (21) and (22), we get

$$\Pr\{Y_n = j\} = \sum_{\ell \in Supp} C_\ell \Pr\{Y_n^{(\ell)} = j\} + \mathcal{O}(n^{-1}) \quad (23)$$

where C_ℓ is a positive constant for every $\ell \in Supp$ and $\sum_{\ell \in Supp} C_\ell = 1$.

Thus, to determine the local behaviour of $\{Y_n\}$, we first study $\{Y_n^{(\ell)}\}$. Indeed, by the previous relation, it is sufficient to prove that the equation

$$\Pr\{Y_n^{(\ell)} = j\} = \begin{cases} \frac{d e^{-\frac{(j-\beta n)^2}{2\alpha n}}}{\sqrt{2\pi\alpha n}} \cdot (1 + o(1)) & \text{if } j \equiv \rho_\ell \pmod{d} \\ 0 & \text{otherwise} \end{cases}$$

holds uniformly for every $j = 0, 1, \dots, n$, where α and β are defined as in T1, while ρ_ℓ is a (possibly depending on n) integer such that $0 \leq \rho_\ell < d$ (in particular $C_i(n) = \sum_{\rho_\ell = i} C_\ell$)

for each i). To this aim, we simply have to show that, for every $n \in \mathbb{N}$, $Y_n^{(\ell)}$ satisfies the hypotheses of Theorem 13.

First, we prove that $Y_n^{(\ell)}$ takes on values only in a set of the form (6), where d is the period of $\mu(a)x + \mu(b)$. We provide an integer ρ_ℓ , such that, if $j \not\equiv \rho_\ell \pmod{d}$, then $r_{n,j}^{(\ell)} = 0$: by Equation (15), this implies that $\text{Prob}\{Y_n^{(\ell)} = j\}$ vanishes too. By the definition of $r^{(\ell)}$, it is clear that for any $j = 0, 1, \dots, n$, the values $r_{n,j}^{(\ell)}$ are given by the convolutions

$$r_{n,j}^{(\ell)} = \sum_{\substack{n_1+n_2+\dots+n_k=n \\ j_1+j_2+\dots+j_k=j}} r_{n_1,j_1}^{(p_1q_1)} \cdot r_{n_2,j_2}^{(p_2q_2)} \cdot \dots \cdot r_{n_k,j_k}^{(p_kq_k)}$$

Now, consider any $r_{n_i,j_i}^{(p_iq_i)}$. By Proposition 7, we know that for each pair p_i, q_i there exist an integer δ_i , $0 \leq \delta_i < d$ such that

$$r_{n_i,j_i}^{(p_iq_i)} \neq 0 \quad \text{implies} \quad j_i \equiv \gamma n_i + \delta_i \pmod{d}$$

where $0 \leq \gamma < d$ does not depend on p_i and q_i . Thus, choosing ρ_ℓ so that $0 \leq \rho_\ell < d$ and $\rho_\ell \equiv \gamma n + \sum_{i=1}^k \delta_i \pmod{d}$, we have that $r_{n,j}^{(\ell)} \neq 0$ implies $j \equiv \rho_\ell \pmod{d}$.

As far as Conditions **C1** and **C2** are concerned, we can argue (with obvious changes) as in the proof of T1 and observe that the two constants α and β are the same for all series $r^{(\ell)}$ with $\ell \in \text{Supp}$, since they depend on the matrices A and B (not on the initial and final arrays).

To prove Condition **C3** let us consider the generating function of $\{r_n^{(\ell)}(z)\}$:

$$\mathbf{r}^{(\ell)}(z, w) = \prod_{j=1}^k \xi_{p_j} (I - w(Ae^z + B))_{p_j q_j}^{-1} \eta_{q_j}.$$

For every $\theta \in \mathbb{R}$ we have

$$\mathbf{r}^{(\ell)}(i\theta, w) = \frac{\prod_{j=1}^k \xi_{p_j} \text{Adj}(I - w(Ae^{i\theta} + B))_{p_j q_j} \eta_{q_j}}{[\det(I - w(Ae^{i\theta} + B))]^k}$$

showing that the singularities of the function are inverses of eigenvalues of $Ae^{i\theta} + B$. As a consequence, by Theorem 10, for every $\theta \neq 2k\pi/d$, all singularities of $\mathbf{r}^{(\ell)}(i\theta, w)$ are in modulus greater than λ^{-1} . Hence, by Cauchy's integral formula, for any arbitrary $\theta_0 \in (0, \pi/d)$ we can choose $0 < \tau < \lambda$ such that the associated sequence $\{r_n^{(\ell)}(i\theta)\}$ is bounded by $O(\tau^n)$ for every $|\theta| \in [\theta_0, \pi/d]$. By (21) this implies

$$\Psi_{Y_n^{(\ell)}}(i\theta) = \frac{r_n^{(\ell)}(i\theta)}{r_n^{(\ell)}(0)} = \frac{O(\tau^n)}{\Theta(n^{k-1}\lambda^n)} = O(\epsilon^n)$$

for some $0 < \epsilon < 1$, which proves Condition **C3**. □

6 Estimate of the maximum coefficients

The result proved in the last section can be used to study the order of growth of the maximum coefficients of rational formal series in commutative variables. This problem was actually

among the motivations of the study presented in [3] and can be seen as an generalization of classical questions concerning the ambiguity in formal language. To introduce this subject consider a rational fraction $p(x)/q(x)$ where $p(x)$ and $q(x)$ are two polynomials with real coefficients and $q(0) \neq 0$. It is well known that the coefficient of the term x^n of its Taylor expansion is asymptotically equivalent to a linear combination of expressions of the form $n^{k-1}\lambda^n$ where λ is a root of $q(x)$ and k its multiplicity, cf. [9, Theorem 6.8] or [19, Lemma II.9.7]. It is natural to ask whether a similar evaluation holds for formal series in two variables both in the commutative and in the noncommutative case.

For rational formal series over a free monoid with integer coefficients, the growth of the coefficients was investigated in [20] (see also [16]), where it is proved that for such a series r either there exists $k \in \mathbb{N}$ such that $\max\{|(r, \omega)| : |\omega| = n\} = O(n^k)$ or $|(r, \omega_j)| \geq c^{|\omega_j|}$ for a sequence of words $\{\omega_j\}$ of increasing length and for some constant $c > 1$. In the first case, the series is the sum of products of at most $k + 1$ characteristic series of regular languages over the free monoid (see also [2, Corollary 2.11]). When the semiring of coefficients is \mathbb{N} the problem is related to the analysis of ambiguity of formal grammar (or finite automata) generating (recognizing, resp.) the support of the series; a wide literature has been devoted to this problem (see for instance [12, 24] and [25, 26] for recent results in the algebraic case).

In this section we consider the problem in the commutative case only. Formally, given a series $S \in \mathbb{R}_+[[\sigma_1, \dots, \sigma_\ell]]$, define its *maximum function* $g_S : \mathbb{N} \rightarrow \mathbb{R}_+$ as

$$g_S(n) = \max\{|(S, \omega)| : \omega \in \{\sigma_1, \dots, \sigma_\ell\}^\otimes, |\omega| = n\} \quad (\text{for every } n \in \mathbb{N}).$$

We estimate the order of magnitude of $g_S(n)$ for any formal series S that is the commutative image of a power of a primitive rational formal series $r \in \mathbb{R}_+^{Rat}\langle\langle a, b \rangle\rangle$.

Corollary 16 *For any $k \in \mathbb{N}$, $k \neq 0$ and any primitive series $r \in \mathbb{R}_+^{Rat}\langle\langle a, b \rangle\rangle$, let $s = r^k$ and consider its commutative image $S = \mathcal{F}(s) \in \mathbb{R}_+^{Rat}[[a, b]]$. Then the maximum function of S satisfies the relation*

$$g_S(n) = \begin{cases} \Theta(n^{k-(3/2)}\lambda^n) & \text{if } r \text{ is non-degenerate} \\ \Theta(n^{k-1}\lambda^n) & \text{otherwise} \end{cases}$$

where $\lambda > 0$.

Proof. Let (ξ, μ, η) be a primitive linear representation of r and let λ be the Perron–Frobenius eigenvalue of $\mu(a) + \mu(b)$. To determine $g_S(n)$ we have to compute the maximum of the values $s_{n,j} = (S, a^j b^{n-j})$ for $j = 0, 1, \dots, n$.

First consider the case when r is non-degenerate. Then, let Y_n count the occurrences of a in the model defined by $s = r^k$ and recall that $\Pr\{Y_n = j\} = s_{n,j}/s_n(0)$. Now, by (22) we have $s_n(0) = \Theta(n^{k-1}\lambda^n)$ and by Theorem 15, the set of probabilities $\{\Pr\{Y_n = j\} \mid j = 0, 1, \dots, n\}$ has the maximum at some integer $j \in [\beta n - d, \beta n + d]$, where it takes on a value of the order $\Theta(n^{-1/2})$. This proves the first equation.

On the other hand, if r is degenerate, then either $\mu(a) = 0$ or $\mu(b) = 0$. In the first case, all $r_{n,j}$ vanish except $r_{n,0}$ which is of the order $\Theta(\lambda^n)$. Hence for every n , the value $\max_j\{s_{n,j}\} = s_n(0)$ is given by the k -th convolution of $r_{n,0}$, which is of the order $\Theta(n^{k-1}\lambda^n)$. The case $\mu(b) = 0$ is similar. \square

Example 8 Consider the rational function $(1 - a - b)^{-k}$. Its Taylor expansion near the origin yields the series

$$S = \sum_{n=0}^{+\infty} \binom{n+k-1}{k-1} \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}$$

By direct computation, one can verify that

$$g_S(n) = \binom{n+k-1}{k-1} \binom{n}{\lfloor n/2 \rfloor} = \Theta(n^{k-3/2} 2^n).$$

In fact, it turns out that $S = \varphi(r^k)$ where $r = \chi_{\{a,b\}^*} \in \mathbb{R}_+ \langle\langle a, b \rangle\rangle$. □

We believe that our result is representative of the asymptotic behaviour of maximum coefficients of all rational formal series in two commutative variables. We actually think that a similar result holds for all rational formal series in commutative variables. More precisely, let us introduce the symbol $\hat{\Theta}$ with the following meaning: for any pair of sequences $\{f_n\}, \{h_n\} \subseteq \mathbb{R}_+$, we have $h_n = \hat{\Theta}(f_n)$ if $h_n = O(f_n)$ and $h_{n_j} = \Theta(f_{n_j})$ for some monotone strictly increasing sequence $\{n_j\} \subseteq \mathbb{N}$. Then we conjecture that the asymptotic behaviour of the maximum function of every rational formal series $t \in \mathbb{R}_+ [[\sigma_1, \dots, \sigma_\ell]]$, is of the form

$$g_t(n) = \hat{\Theta} \left(n^{k/2} \lambda^n \right)$$

for some integer $k \geq -\ell + 1$ and some $\lambda \in \mathbb{R}_+$.

References

- [1] E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combin. Theory Ser. A*, 15:91–111, 1973.
- [2] J. Berstel and C. Reutenauer. *Rational series and their languages*, Springer-Verlag, New York - Heidelberg - Berlin, 1988.
- [3] A. Bertoni, C. Choffrut, M. Goldwurm, and V. Lonati. On the number of occurrences of a symbol in words of regular languages. *Theoret. Comput. Sci.*, 302(1-3):431–456, 2003.
- [4] D. de Falco, M. Goldwurm, and V. Lonati. Frequency of symbol occurrences in bicomponent stochastic models. *Theoret. Comput. Sci.*, 327(3):269–300, 2004.
- [5] R.N. Bracewell. *The Fourier transform and its applications*, McGraw-Hill Book Company, 1986.
- [6] P. Flajolet and R. Sedgewick. The average case analysis of algorithms: multivariate asymptotics and limit distributions. *Rapport de recherche n. 3162*, INRIA Rocquencourt, May 1997.
- [7] B.V. Gnedenko. *The theory of probability* (translated by G. Yankovsky). Mir Publishers - Moscow, 1976.
- [8] L. J. Guibas and A. M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *J. Combin. Theory Ser. A*, 30(2):183–208, 1981.

- [9] P. Henrici. *Elements of numerical analysis*, John Wiley, 1964.
- [10] H. K.Hwang. Large deviations for combinatorial distributions. I: central limit theorems, *Ann. Appl. Probab.*, vol. 6(1):297–319, 1996.
- [11] H. K.Hwang. Large deviations for combinatorial distributions II. Local limit theorems, *Ann. Appl. Probab.*, 8(1):163–181, 1998.
- [12] W. Kuich. Finite automata and ambiguity. Technical Report n.253, I.I.G. University of Graz, 1988.
- [13] W. Kuich and A. Salomaa. *Semirings, automata, languages*. Springer–Verlag, New York Heidelberg Berlin Tokyo, 1986.
- [14] P. Nicodème, B. Salvy, and P. Flajolet. Motif statistics. *Theoret. Comput. Sci.*, 287(2):593–617, 2002.
- [15] B. Prum, F. Rudolphe and E. Turckheim. Finding words with unexpected frequencies in deoxyribonucleic acid sequence. *J. Roy. Statist. Soc. Ser. B*, 57:205–220, 1995.
- [16] C. Reutenauer. *Propriétés arithmétiques et topologiques de séries rationnelles en variables non commutatives*, These Sc. Maths, Doctorat troisieme cycle, Université Paris VI, 1977.
- [17] M. Régnier and W. Szpankowski. On pattern frequency occurrences in a Markovian sequence. *Algorithmica*, 22(4):621–649, 1998.
- [18] J. Sakarovitch. *Eléments de théorie des automates*, Vuibert Informatique, 2003.
- [19] A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series*, Springer-Verlag, 1978.
- [20] M.-P. Schützenberger. Finite counting automata. *Inform. and Control*, 5:91–107, 1962.
- [21] E. Seneta. *Non-negative matrices and Markov chains*, Springer–Verlag, New York Heidelberg Berlin, 1981.
- [22] M. Waterman. *Introduction to computational biology*, Chapman & Hall, New York, 1995.
- [23] A. Weber. On the valuedness of finite transducers. *Acta Informatica*, 27:749–780, 1990.
- [24] A. Weber and H. Seidl. On the degree of ambiguity of finite automata. *Theoret. Comput. Sci.*, 88:325–349, 1991.
- [25] K. Wich. Exponential ambiguity of context-free grammars. In *Proceedings of the 4th D.L.T. Conference*, G. Rozenberg and W. Thomas editors. World Scientific, Singapore, 2000, 125–138.
- [26] K. Wich. Sublinear ambiguity. In *Proceedings of the 25th M.F.C.S. Conference*, M. Nielsen and B. Rován editors. Lecture Notes in Comput. Sci., vol. n.1893, Springer-Verlag, 2000, 690–698.