INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Context Free Evolution of Words

F. I. Karpelevich , V. A. Malyshev , A. I. Petrov ,

S. A. Pirogov , A. N. Rybko

N° 4413

January 23 2002

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4413--FR+ENG

INRIA ROCQUENCOURT

Context Free Evolution of Words

F. I. Karpelevich [*], V. A. Malyshev [†], A. I. Petrov [‡],
S. A. Pirogov [§], A. N. Rybko [¶]

Thème 1 — Réseaux et systèmes
Projet Meval

Rapport de recherche n° 4413 — January 23 2002 — 28 pages

Abstract: Random grammars were introduced in computer science, but the study of their thermodynamic and long time behaviour started only recently. In this paper we undertake more detailed study of context free grammars in the supercritical case, that is when the word grows exponentially fast. We study and calculate the statistics of factors for large t, prove the existence of various limiting measures and study relations between them.

Key-words: random grammars, branching processes, thermodynamic limit

[*] Institute of Transport, Moscow
[†] INRIA, France
[‡] Laboratory of Large Systems, Moscow State University
[§] IITP, Russian Acad. of Sciences
[¶] IITP, Russian Acad. of Sciences

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Évolution contexte-libre des mots

Résumé : On considère une évolution markovienne de mots par substitutions aléatoires. Chaque substitutions remplace un symbole par une petite chaine. Les resultats principaux concernent la structure des mots ainsi obtenus quand le temps augmente infiniment.

Mots-clés : grammaires aléatoires, processus de branchement, limite thermodynamique

1 Definitions

Random grammars were introduced in computer science, but the study of their thermodynamic and long time behaviour started only recently, see [2]. In this paper we undertake more detailed, comparative with [2], study of context free grammars in the supercritical case, that is when the word grows exponentially fast. We study and calculate the statistics of factors for large t, prove the existence of various limiting measures and study relations between them.

A word α is a linearly-ordered sequence of symbols from S. S is a finite set which we call the alphabet. For finite words

$$\alpha = x_1 ... x_n, \beta = y_1 ... y_m$$

their concatenation $\alpha\beta$ is

$$\alpha\beta = x_1 ... x_n y_1 ... y_m$$

Let $n = \mid \alpha \mid = l(\alpha)$ be the length of α. Let $e = \emptyset$ be an empty word. We call α a subword (more exactly, a factor) of γ if $\gamma = \beta\alpha\delta$ for some words β, δ. Denote S^* to be the set of all finite words over the alphabet S, including the empty one. There are many interesting problems concerning the set S^∞ of all infinite words over the alphabet S. However, we do not study the evolution of infinite words here, see the construction of random dynamics for the general case in [2].

Consider a discrete time countable Markov chain with the state space S^*. One step transitions are defined as follows: each symbol x of the word at time t, independently of all others, is substituted by the word α_x with the probability $P(x \to \alpha_x)$. We consider the case when α_x of one or two symbols, that is only $P(x \to y)$ and $P(x \to yz)$ can be nonzero. In fact, all results of this paper are valid in the case when $K \geq 2$ exists such that $\forall x \in S$, $\forall \alpha \in S^*$, $1 \leq |\alpha| < K : P(x \to \alpha) > 0$ and $\forall x \in S : \sum_{\alpha \in S^*, |\alpha| < K} P(x \to \alpha) = 1$. This case differs longer notations only.

At time $t = 0$ the state is a word $\gamma(0) = x_1(0)...x_{N_0}(0)$ of length N_0. At time t the number of symbols is N_t and the state at time t is

$$\gamma(t) = x_1(t)...x_{N_t}(t)$$

It is clear that a.s. $N_t \to \infty$.

We are interested in a finer structure of the long random word $\gamma(t)$. For example: how many subwords of $\gamma(t)$ are identical to a given word α? We shall see that for such questions it is natural to use some terminology related to statistical physics: thermodynamic limit, large time behaviour, correlation functions etc. We should note here that one can get an impression that $\gamma(t)$ tends in some way to a stationary random field on Z with values in S. This is however a very deceptive impression based on the absence of natural choice of the origin in the long word. In fact one can get in general many limiting random fields (the notion of a local observer, see [2]), this is what we want to show.

We want to define correlation functions but at first we remark that it is useful to associate a planar tree to each trajectory of the process $\gamma(t)$. Each symbol s of the initial string produces independently a tree of descendants. This tree is connected iff the initial word $\gamma(0)$ contains one symbol only. Thus it is sufficient to consider initial word consisting of one symbol only. The tree has levels labeled by time moments $0, 1, 2, \ldots$. In fact, as we will see below, for the generic situation the considered limiting measures do not depend on $\gamma(0)$ and hence all statements which we will prove for $|\gamma(0)| = 1$ are correct also for $|\gamma(0)| > 1$. One can define the generic situation, for example, as follows

$$\forall x, y, z \in S: \; P(x \to y), P(x \to yz) > 0 \tag{1}$$

$$\forall x \in S: \; \sum_{y \in S} P(x \to y) + \sum_{y,z \in S} P(x \to yz) = 1 \tag{2}$$

In the general case there can be several limiting measures, the discussion and some results see in [2].

Define the m-point correlation functions (more exactly, the $\gamma(0)$-correlation functions) $p_j^{(m)}(y_1, \ldots, y_m; t) \equiv p_j^{(m)}(y_1, \ldots, y_m; t, \gamma(0))$ in the following way $\forall t > 0$, $\forall j > 0$, $1 \leq j \leq N_t - m + 1$:

$$p_j^{(m)}(y_1, \ldots, y_m; t) = P(x_j(t) = y_1, \ldots, x_{j+m-1}(t) = y_m),$$

where $y_1, \ldots, y_m \in S$.

We define also empirical correlation functions $p_{emp}^{(m)}(y_1, \ldots, y_m; t)$.

Let $N_t(\alpha, \gamma(t))$ be the random number of subwords identical to $\alpha = y_1, \ldots, y_m$ in the word $\gamma(t)$. Then

$$p_{emp}^{(m)}(y_1, \ldots, y_m; t) \equiv p_{emp}^{(m)}(y_1, \ldots, y_m; t, \gamma(0)) = \frac{EN_t(\alpha, \gamma(t))}{EN_t - m + 1}$$

The main objects of our study are different limiting correlation functions or probability measures.

2 Main Results

In the following theorems the initial condition $\gamma(0)$ is fixed.

Theorem 1 *For any j the limiting correlation functions*

$$\pi_j^{(m)}(y_1, \ldots, y_m) = \lim_{t \to \infty} p_j^{(m)}(y_1, \ldots, y_m; t)$$

exist.

Theorem 2 *For any $m, y_1, ..., y_m$ there exist the limiting empirical correlation functions*

$$\pi_{emp}^{(m)}(y_1, ..., y_m) = \lim_{t\to\infty} p_{emp}^{(m)}(y_1, ..., y_m; t)$$

It is clear that for fixed m

$$\sum_{y_1,...,y_m} \pi_{emp}^{(m)}(y_1, ..., y_m) = 1$$

Then by Kolmogorov theorem the correlation functions define, as finite-dimensional distributions, some probability measures μ_{emp} and $\mu_j, j = 0, 1, ...,$ on $S^{Z_+} = \{(y_1, y_2, ...) : y_i \in S\}$.

For all $t > 0$, for all $j, L > 0$ let $N_t(\alpha, [j, j + L])$ be a random number of subwords $\alpha \in S^*$ in the word $x_j(t), ..., x_{j+L}(t)$.

Theorem 3 *For all parameters, for any word $\alpha \in S^*$ and any $\varepsilon > 0$ there exist $L(\varepsilon, \alpha)$, $t(\varepsilon, \alpha)$ such that for all $L > L(\varepsilon, \alpha), t > t(\varepsilon, \alpha), j > 0$:*

$$P\left(\left|\frac{N_t(\alpha, [j, j + L])}{L - |\alpha| + 1} - \pi_{emp}^{(|\alpha|)}(\alpha)\right| > \varepsilon\right) < \varepsilon$$

Theorem 4 *For all parameters the limits*

$$\lim_{j\to\infty} \frac{1}{j} \sum_{k=1}^{j} \pi_k^{(m)}(y_1, ..., y_m) = \pi_{emp}^{(m)}(y_1, ..., y_m)$$

exist.

Theorem 5 *For almost all parameters the limits*

$$\pi^{(m)}(y_1, ..., y_m) = \lim_{j\to\infty} \pi_j^{(m)}(y_1, ..., y_m)$$

exist, i.e. there exist a limiting measure

$$\mu = \lim_{j\to\infty} \mu_j$$

For those parameters for which $\pi^{(m)}(y_1, ..., y_m)$ exists,

$$\pi^{(m)}(y_1, ..., y_m) = \pi_{emp}^{(m)}(y_1, ..., y_m)$$

For almost all parameters means for all parameters $\{P(x \to y), P(x \to yz)\}$ except some its subset of Lebesgue measure 0. An example, when the limit does not exist, is the following assertion.

Proposition 6 *If $P(x \to y) \equiv 0$ then for all parameters $\{P(x \to yz)\}$ the limits*

$$\lim_{j\to\infty} \pi_j^{(m)}(y_1, ..., y_m)$$

do not exist.

3 Long Time Limits

3.1 Boundary correlation functions

Here we will prove theorem 1.

In the proof of the existence of $\pi_j^{(m)}(y_1, ..., y_m)$ without loss of generality one can take $j = 0$. Introduce auxiliary $(\mid S \mid \times \mid S \mid)$-matrices and operator Δ:

$$Q_l(s_1, s_2) = \sum_{u \in S} P(s_1 \to s_2 u) \tag{3}$$

$$Q_r(s_1, s_2) = \sum_{u \in S} P(s_1 \to u s_2) \tag{4}$$

$$Q_1(s_1, s_2) = P(s_1 \to s_2) \tag{5}$$

If $f : S \to R$ - is any positive function then

$$\forall x, y \in S : \ (\Delta f)(x, y) = \sum_{z \in S} f(z) P(z \to xy) \tag{6}$$

Note first that

$$\pi_0(s) = \pi_0^{(1)}(s)$$

are equal to the stationary distribution of the finite Markov chain $\mathcal{L}_1$ having the state space S and transitions defined by the stochastic matrix $Q_1 + Q_l$. Similarly, one can define finite Markov chains $\mathcal{L}_m$ having the state space S^m that is the set of all words with length m. For examples, for $m = 2$ the transitions are defined by

$$p^{(2)}(t) = p^{(2)}(t-1)(Q_1 \otimes (Q_1 + Q_l)) + \Delta p^{(1)}(t-1)$$

or

$$p^{(2)}(x, y; t) = \sum_{x', y'} p^{(2)}(x', y'; t-1)[(P(x' \to x) P(y' \to y) +$$

$$+ P(x' \to x) \sum_z P(y' \to yz))] + \sum_{x'} p^{(1)}(x'; t-1) P(x' \to xy)$$

One can write down such formulas for any m but an important thing is that the first m symbols of the word at time t depend only on the first m symbols of the word at time $t-1$ and hence from (1) it follows that $\mathcal{L}_m$ is an ergodic Markov chain. ■

3.2 Empirical correlation functions

Here we will prove theorem 2.

At first we recall some facts from the theory of multi-type branching processes.

Let $N_t(s)$, $s \in S$ be a random number of symbols s in the word $\gamma(t)$. Let $\vec{N}_t = \{N_t(s_1), ..., N_t(s_{|S|})\}$ be a vector of length $|S|$, where $s_1, ..., s_{|S|} \in S$, $s_i \neq s_j$ if $i \neq j$. Then $\vec{N}_t$ is a branching process with $|S|$ types and mean matrix $M = (m_{ij})_{i,j=1,...,|S|}$, i.e. m_{ij} is the expected number of symbol s_j offspring of a single symbol s_i in one generation. It easy to see that $E\vec{N}_t = \vec{N}_0 M^t$. From (1) and (2) it follows that matrix M is positive matrix (i.e. $(M)_{ij} > 0$ for all $i, j = 1, ..., |S|$) and the process $\vec{N}_t$ is a supercritical process. (See [6], [1].) Hence the maximal eigenvalue ρ of the matrix M is greater than 1. Let $\vec{u}$ and $\vec{v}$ are respectively right and left eigenvectors of M associated with ρ. It is obvious that the extinction probability is equal to zero.

We need the following lemma.

Lemma 7 *Let $\vec{Z}(t; s_i) = (Z_1(t; s_i), ..., Z_{|S|}(t; s_i))$ be a supercritical branching process with $|S|$ types and with initial state $(0, ...0, 1, 0, ...0)$, where 1 stands on the i-th place, i.e. single symbol $s_i \in S$ is present at time $t = 0$. Let M be the mean matrix of this process. M is a positive matrix and it has a maximal eigenvalue ρ and associated positive right and left eigenvectors $\vec{u}$ and $\vec{v}$. These are normalized so that $(\vec{v} \cdot \vec{u}) = 1$ and $(\vec{1} \cdot \vec{v}) = 1$*

Then there exists a random vector $\vec{W}(s_i)$ and a one-dimensional random variable $w(s_i)$ such that

$$\lim_{t\to\infty} \frac{\vec{Z}(t; s_i)}{\varrho^t} = \vec{W}(s_i) \ a.s.$$

and

$$\vec{W}(s_i) = w(s_i) \cdot \vec{v} \ a.s.$$

If $E(Z_j(1, s_i) \log Z_j(1, s_i)) < \infty$ for all $1 \leq j \leq |S|$, $s_i \in S$ then

$$E(w(s_i)) = u_i, \ 1 \leq i \leq |S|,$$

This lemma is a special case of the theorem 1.1 in [9] and we omit proof.

Lemma 8 *Let all conditions of the previous lemma hold. Let each symbol has at least one descendant and for all $s_i \in S$*

$$E(e^{\theta_0(1\cdot\vec{Z}(1,s_i))}) < \infty \ for\ some\ \theta_0 > 0$$

then there exists constants $0 < C < \infty$ and $0 < \lambda < \infty$ such that for any $s_i \in S$, $t > 0$, $\varepsilon > 0$:

$$P(|W - \frac{\vec{Z}(t, s_i)}{\rho^t}| \geq \varepsilon) \leq Ce^{-\lambda\varepsilon^{2/3}(\rho^{1/3})^t}$$

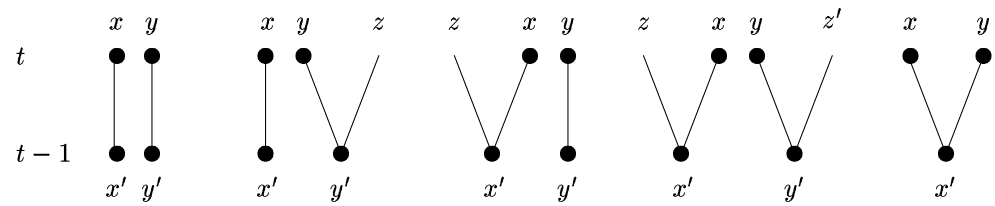

Figure 1: The different cases for $\alpha = xy$

Proof. This is theorem 2.6 from [3]. ■.
From lemmas 7 and 8 it follows that

$$\lim_{t\to\infty} \frac{E\vec{N}_t}{\rho^t} = E\lim_{t\to\infty} \frac{\vec{N}_t}{\rho^t} = u_i\vec{v}, \text{ if } \gamma(0) = s_i \in S \tag{7}$$

We will write down the recurrent equations for $e_t^{(|\alpha|)}(\alpha) = EN_t(\alpha;\gamma(t))$ for $|\alpha| = 2$

$$e_t^{(2)}(xy) = \sum_{x',y'} e_{t-1}^{(2)}(x'y')P(x' \to x)P(y' \to y) +$$

$$+\sum_{x',y'} e_{t-1}^{(2)}(x'y')P(x' \to x)\sum_z P(y' \to yz)) +$$

$$+\sum_{x',y'} e_{t-1}^{(2)}(x'y')\sum_z P(x' \to zx)P(y' \to y) +$$

$$+\sum_{x',y'} e_{t-1}^{(2)}(x',y')\sum_z P(x' \to zx)\sum_{z'} P(y' \to yz') + \sum_{x'} e_{t-1}^{(1)}(x')P(x' \to xy)$$

Let $e_t^{(2)}$ be the vector with components $e_t^{(2)}(xy)$, and $e_t^{(1)}$ be the vector with components $e_t^{(1)}(x)$, $x, y \in S$ then in the vector form the above equations look like

$$e_t^{(2)} = e_{t-1}^{(2)}A + e_{t-1}^{(1)}B, e_0^{(2)} = 0 \tag{8}$$

with the matrices

$$A = (A(x'y', xy)) = (P(x' \to x)P(y' \to y) + P(x' \to x)\sum_z P(y' \to yz)) +$$

$$+\sum_z P(x' \to zx)P(y' \to y) + \sum_z P(x' \to zx) \sum_{z'} P(y' \to yz') \qquad (9)$$

$$B = (B(x', xy)) = (P(x' \to xy))$$

We rewrite (8) in the following way:

$$e_t^{(2)} = e_{t-1}^{(2)} A + e_{t-1}^{(1)} B, e_{t-1}^{(2)} = e_{t-2}^{(2)} A + e_{t-2}^{(1)} B,$$

hence

$$e_t^{(2)} = e_{t-2}^{(2)} A^2 + e_{t-2}^{(1)} BA + e_{t-1}^{(1)} B =$$

$$= e_{t-3}^{(2)} A^3 + e_{t-3}^{(1)} BA^2 + e_{t-2}^{(1)} BA + e_{t-1}^{(1)} B = ... = \sum_{k=0}^{t-1} e_{t-k-1}^{(1)} BA^k$$

and

$$\frac{e_t^{(2)}}{\rho^t} = \frac{1}{\rho} \sum_{k=0}^{t-1} \frac{e_{t-k-1}^{(1)} B}{\rho^{t-k-1}} \frac{A^k}{\rho^k} \qquad (10)$$

It is obvious that A is a stochastic matrix. From (7) it follows that $C > 0$ exists such that $\forall t, k > 0$:

$$|\frac{e_{t-k-1}^{(1)} B}{\rho^{t-k-1}}| < C,$$

where for all vectors $\vec{a}$ the $|\vec{a}|$ is maximum modulus of its components.

Since $\rho > 1$, the sequence in the right part of (10) converges uniformly in t. Hence we can take the limit $t \to \infty$ in all terms of the right hand-side of (10). Taking into account (7) we get

$$\lim_{t \to \infty} \frac{e_t^{(2)}}{\rho^t} = \frac{u_i \vec{v} B}{\rho} \sum_{k=0}^{t-1} \frac{A^k}{\rho^k} = \frac{u_i \vec{v} B}{\rho} \frac{\rho}{\rho - A} = \frac{u_i \vec{v} B}{\rho - A},$$

where $\vec{u} = \{u_1, ..., u_{|S|}\}$ and $\vec{v}$ are defined above.

3.3 Stationary limit

This section is organized as follows. We start with the proof of proposition 6. This will show some subleties of this problem. After that we shall prove some auxiliary lemmas. Then we shall prove theorems 3 and 4. Then we again prove some auxiliary lemmas. Finally we prove theorem 5 for one-particle correlation functions and we give a scheme of the proof of theorem 5 for two-particle correlation functions. Multi-particle case differs longer notations only.

3.3.1 Proof of proposition 6.

Let us prove proposition 6. We will get an explicit formula for $p_j^{(1)}(s;t)$ and for $\pi_j(s)$. Consider all trees T at time t. For fixed T the probability of any path in T from level 0 (the root vertex) to level t is equal to the product of some elements of matrices Q_l, Q_r (see (3) and (4)) If for any $x, y \in S$ $P(x \to y) = 0$ then Q_l and Q_r are stochastic matrices. Assume that $j < 2^t$ and consider the binary decomposition of $j-1$

$$j-1 = \sum_{k=0}^{k_0} a_k 2^k, \ a_k = 0,1$$

where $k_0 = k_0(j) = \max\{k : a_k = 1\}$, $k_0 = 0$ if $a_k = 0$ for all $k \geq 0$.

Let $\vec{p}^{(1)}(0) = \{0, ..., 0, 1, 0, ..., 0\}$, where 1 stands on the i-th place, $\gamma(0) = s_i$.

Then for $t > k_0$

$$p_j^{(1)}(s;t) = \left(\vec{p}^{(1)}(0) Q_l^{t-k_0-1} Q_{u(k_0)} Q_{u(k_0-1)} ... Q_{u(0)}\right)(s)$$

where $u(i) = l$ if $a_i = 0$ and $u(i) = r$ if $a_i = 1$. It follows

$$\vec{\pi}_j = \vec{\pi}_1 Q_{u(k_0)} Q_{u(k_0-1)} ... Q_{u(0)}, \tag{11}$$

where $\vec{\pi}_j = \{\pi_j^{(1)}(s_1), ..., \pi_j^{(1)}(s_{|S|})\}$, $s_i \in S$, $i = 1, ..., |S|$ and $s_i \neq s_k$, if $i \neq k$.

It is easy to see that for almost all parameters the eigenvectors of Q_l and Q_r associated with eigenvalue 1 are different. Hence from (11) it follows that there are at least two (in fact continuum) limiting points of the sequence $\vec{\pi}_j$ - stationary distributions of finite Markov chains defined by Q_l and Q_r correspondingly, because for any k there are infinite number of j such that either $u(k) = ... = u(0)) = l$ or $u(k) = ... = u(0)) = r$. This completes the proof of proposition 6.

3.3.2 Proof of theorem 3.

Now we come to the generic situation and will prove theorem 3.

We start with some notation. Put for any $a, b > 0$:

$$[a, b] = \{n \in Z : a \leq n \leq b\}$$

Let us chose sufficient large t_0, $T > 0$ and let for any fixed $x_1(t_0), x_2(t_0), ...$:

$$D_i \equiv D_i(T, x_i(t_0)), \ i = 1, 2, 3, ... \tag{12}$$

be the set of indices of all descendants of the symbol $x_i(t_0)$ at time $t_0 + T$.

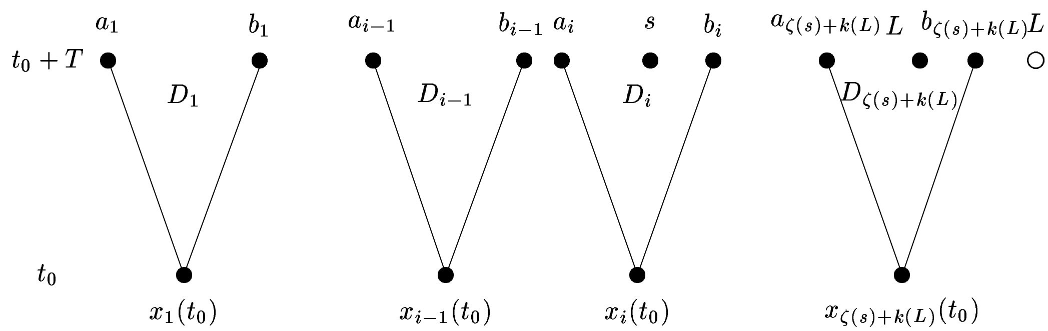

Figure 2: Picture for $i = \zeta(s)$

Put

$$a_i \equiv a_i(T, x_i(t_0)) = \min_j \{j \in D_i\},$$

$$b_i \equiv b_i(T, x_i(t_0)) = \max_j \{j \in D_i\}.$$

Then D_i is the segment $[a_i, b_i]$.

Let n_i be the random number of symbols in D_i

$$n_i \equiv n_i(T, x_i(t_0)) = b_i - a_i = \#\{D_i\} \tag{13}$$

Let $n_i(\alpha) \equiv n_i(\alpha, T, x_i(t_0))$ be the random number of subwords $\alpha \in S^*$ in D_i. Let $\vec{X}_n(t_0) = (x_1(t_0), ..., x_n(t_0))$, $n \in Z^+$. For all $s > 0$ we define $\zeta(s) \equiv \zeta(s, t_0, T)$ as follows

$$\zeta(s) = \min_{i>0} \{ \sum_{j=1}^{i} n_j \geq s - 1 \}$$

For all intervals $[s, s+L]$, $s, L > 0$, for all $\varepsilon > 0$ we choose $k(L) \equiv k(\zeta(s), L, T, \vec{X}_{s+L}(t_0)) \in Z^+$ such that

$$\left| \sum_{i=\zeta(s)+1}^{\zeta(s)+k(L)} En_i - L \right| < 2 \cdot 2^T \tag{14}$$

It is always possible since $\forall i > 0 : En_i < 2^T$.

Lemma 9 *For any fixed $T > 0$*

$$k(L) \to \infty \textit{ if } L \to \infty$$

and there exist $T_0 > 0$ asuch that for any $T > T_0$ and $L > 0$

$$\frac{C_1}{\rho^T} \leq \frac{k(L)}{L} \leq \frac{C_2}{\rho^T}$$

where $C_1, C_2 > 0$; $\rho > 1$ are defined in section 3.2.

Proof. Let for any $t_0 > 0$, $1 \leq i \leq N_{t_0}$, $1 \leq j \leq |S|$ $Z_j(T, x_i(t_0))$ be the random number of symbols $s_j \in S$ in D_i. Then $\vec{Z}(T, x_i(t_0)) = \{Z_1(T, x_i(t_0)), ..., Z_{|S|}(T, x_i(t_0))\}$ is a branching process with $|S|$ types and transitions probabilities which are defined by (1), (2), i.e. the process $\vec{Z}(T, x_i(t_0))$ is completely similar to the process $\vec{N}_T$ defined in section 3.2, but it has initial state $x_i(t_0)$. $\vec{Z}(T, x_i(t_0))$ is a supercritical process. Hence, similarly to (7), from lemmas 7 and 8 it follows that there exist T_0, $c_1, c_2 > 0$ such that for any $T > T_0$:

$$\rho^T c_1 \leq \sum_{s=1}^{|S|} EZ_s(T, x_i(t_0)) \leq \rho^T c_2, \; i = 1, ..., N_{t_0} \tag{15}$$

Since

$$En_i = \sum_{s=1}^{|S|} E\vec{Z}_s(T, x_i(t_0)), \tag{16}$$

then from (16) and (15) we get

$$c_1\rho^T \leq En_i \leq c_2\rho^T \tag{17}$$

Hence from (14) it follows that for any fixed $T > 0$

$$k(L) \to \infty \text{ if } L \to \infty \tag{18}$$

From (14) it follows

$$\sum_{i=\zeta(s)+1}^{\zeta(s)+k(L)} En_i - 2^{T+1} < L < \sum_{i=\zeta(s)+1}^{\zeta(s)+k(L)} En_i + 2^{T+1} \tag{19}$$

hence from (17) and (19) we get

$$c_1(k(L)-1)\rho^T - 2^{T+1} \leq L \leq c_2(k(L)-1)\rho^T + 2^{T+1}$$

One can rewrite it as

$$c_1(1 - \frac{1}{k(L)})\rho^T - \frac{2^{T+1}}{k(L)} \leq \frac{L}{k(L)} \leq c_2(1 - \frac{1}{k(L)})\rho^T + \frac{2^{T+1}}{k(L)}$$

By (18) this implies that

$$\frac{C_1}{\rho^T} \leq \frac{k(L)}{L} \leq \frac{C_2}{\rho^T}$$

■

Lemma 10 *For all $\alpha \in S^*$, $\varepsilon > 0$, $\delta > 0$ there exist $T > 0$, $L_0 > 0$ such that for all t_0 and $L > L_0$*

$$P(|\frac{N_{t_0+T}(\alpha, [\sum_{j=1}^{\zeta(s)} n_j, \sum_{j=1}^{\zeta(s)+k(L)} n_j])}{L} - \frac{N_{t_0+T}(\alpha, [s, s+L])}{L}| > \varepsilon) < \delta$$

Proof. It is sufficient to estimate the number of elements of

$$M := \{\left(\left[\sum_{j=1}^{\zeta(s)} n_j, \sum_{j=1}^{\zeta(s)+k(L)} n_j\right] \cup [s, s+L]\right) \setminus \left(\left[\sum_{j=1}^{\zeta(s)} n_j, \sum_{j=1}^{\zeta(s)+k(L)} n_j\right] \cap [s, s+L]\right)\}$$

From lemma 16 below and (14) it follows that

$$P(\left|\sum_{j=\zeta(s)+1}^{\zeta(s)+k(L)} n_j - L\right| > 2^{T+1} + k(L)) < 2exp\{-ck(L)\},$$

where $c > 0$. Hence

$$P(\left|\frac{\sum_{j=\zeta(s)+1}^{\zeta(s)+k(L)} n_j - L}{L}\right| > \frac{2^{T+1} + k(L)}{L}) < 2exp\{-ck(L)\}$$

On the other hand, from definition $\zeta(s)$ it is obvious that

$$|\sum_{j=1}^{\zeta(s)} n_j - s| < 2^T.$$

From lemma 9 we can chose $T > 0$ and $L > 0$ such that

$$\frac{2^T}{L} + \frac{2^{T+1} + k(L)}{L} < \varepsilon$$

and

$$2exp\{-ck(L)\} < \delta$$

Hence

$$P(\frac{\#M}{L} > \varepsilon) < \delta$$

To end the proof we apply an obvious inequality: for any $t > 0, \alpha \in S^*$ and for all sets $M \subset [1, N_t]$

$$N_t(\alpha, M) \leq \#M,$$

where $N_t(\alpha, M)$ is a number of subwords α in the set M at time t.

The lemma 10 is proved. ■

For all words $\alpha \in S^*$ let $\delta_i(\alpha) \equiv \delta_i(\alpha, T, x_i(t_0))$ be the random number of subwords α in the word

$$x_{b_{i-1}-|\alpha|+1}(t_0+T), ..., x_{a_i+|\alpha|-1}(t_0+T),$$

where $2 \le i \le N_{t_0}$, $\delta_1(\alpha) = 0$.

$$\Delta_{k(L)}(\alpha) \equiv \Delta_{k(L)}(\alpha, T, x_{k(L)}(t_0)) := \sum_{i=\zeta(s)+1}^{\zeta(s)+k(L)} \delta_i(\alpha)$$

It is obvious that

$$\Delta_{k(L)}(\alpha) \le k(L) \cdot |\alpha| \tag{20}$$

If tends t_0 to infinity then under the conditions of theorem 1 we get a time homogeneous process. We need some lemmas.

Lemma 11 *For any $L > 0$, $\alpha \in S^*$*

$$n_{\zeta(s)+1}(\alpha) + ... + n_{\zeta(s)+k(L)}(\alpha) = En_{\zeta(s)+1}(\alpha) + ... + En_{\zeta(s)+k(L)}(\alpha) + \xi_{k(L)}(\alpha)$$

where $\forall \varepsilon > 0$:

$$P(|\xi_{k(L)}(\alpha)| > \varepsilon k(L)) < \frac{2^T}{\varepsilon^2 k(L)}$$

Proof. We apply the Chebyshev inequality to $\xi_{k(L)}$:

$$P(|\xi_{k(L)}(\alpha)| > \varepsilon k(L)) \le \frac{D(n_{\zeta(s)+1}(\alpha) + ... + n_{\zeta(s)+k(L)}(\alpha))}{\varepsilon^2 k^2(L)} \tag{21}$$

Since $n_{\zeta(s)+1}(\alpha), ..., n_{\zeta(s)+k(L)}(\alpha)$ are independent random variables, we can rewrite (21) as:

$$P(|\xi_{k(L)}(\alpha)| > \varepsilon k(L)) \le \frac{D(n_{\zeta(s)+1})(\alpha) + ... + D(n_{\zeta(s)+k(L)}(\alpha))}{\varepsilon^2 k^2(L)} \le$$
$$\le \frac{2^T k(L)}{\varepsilon^2 k^2(L)} = \frac{2^T}{\varepsilon^2 k(L)} \tag{22}$$

where we use the obvious estimate: for any $i > 0$ $Dn_i(\alpha) \le 2^T$. ■

Lemma 12 *For any $t_0 > 0$, any $\varepsilon > 0$ and $i = 1, 2, ...$ There exists $T_0 > 0$ such that for any $T > T_0$:*

$$|\frac{En_i(\alpha, T, x_i(t_0))}{En_i(T, x_i(t_0))} - \pi_{emp}^{(|\alpha|)}(\alpha)| < \varepsilon$$

Proof. Let us consider the other processes, those completely similar to our process but having initial state s_i, $i = 1, ..., |S|$, $s_i \neq s_j$, if $i \neq j$.

Let $N(\alpha, t|s_i)$ be a random number of subwords α at time t in the process with initial state s_i. Let $N(t|s_i)$ be a random number of symbols at time t in the process with initial state s_i. To each of these processes we apply theorem 2 and get that $\pi_{emp}^{(|\alpha|)}(\alpha|s_i)$ exists such that: for any $\varepsilon_{\zeta(s)+1}, ..., \varepsilon_{|S|}$ there exist $t_1^0, ..., t_{|S|}^0$ such that for all $t_1 > t_1^0, ..., t_{|S|} > t_{|S|}^0$:

$$|\frac{En(\alpha, t|s_i)}{EN(t|s_i) - |\alpha| + 1} - \pi_{emp}^{(|\alpha|)}(\alpha|s_i)| < \varepsilon_i$$

From the proof of theorem 2 it follows that $\pi_{emp}^{(|\alpha|)}(\alpha|s_i)$ does not depend on the initial state s_i. Hence

$$\pi_{emp}^{(|\alpha|)}(\alpha|s_1) = ... = \pi_{emp}^{(|\alpha|)}(\alpha|s_{|S|}) = \pi_{emp}^{(|\alpha|)}(\alpha)$$

Since $|\alpha|$ is fixed then we get the proof of the lemma for any $x_i(t_0)$. ∎

Lemma 13 *For all* $\alpha \in S^*$, $\varepsilon > 0, \delta > 0$ *there exist* $T > 0, L_0 > 0$ *such that for any* $L > L_0$, $t_0 > 0$, $s > 0$:

$$P(|\frac{N_{t_0+T}(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])}{L} - \pi_{emp}^{(\alpha)}(\alpha)| < \varepsilon) > 1 - \delta \tag{23}$$

Proof. From the definitions of $n_i(\alpha)$ and $\Delta_{k(L)}(\alpha)$ we see that

$$N(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i]) \equiv N_{t_0+T}(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])$$

$$= n_{\zeta(s)+1}(\alpha) + ... + n_{\zeta(s)+k(L)}(\alpha) + \Delta_{k(L)}(\alpha) \tag{24}$$

Taking into consideration (14) and lemma 11, we estimate

$$\frac{N(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])}{L} \leq$$

$$\leq \frac{n_{\zeta(s)+1}(\alpha) + ... + n_{\zeta(s)+k(L)}(\alpha) + \Delta_{k(L)}(\alpha)}{En_{\zeta(s)+1} + En_{\zeta(s)+2} + ... + En_{\zeta(s)+k(L)} - 2^{T+1}} =$$

$$\frac{En_{\zeta(s)+1}(\alpha) + ... + En_{\zeta(s)+k(L)}(\alpha) + \xi_{k(L)}(\alpha) + \Delta_{k(L)}(\alpha)}{En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)} - 2^{T+1}} \leq$$

$$\leq \frac{En_{\zeta(s)+1}(\alpha) + ... + En_{\zeta(s)+k(L)}(\alpha)(1 + \frac{\xi_{k(L)}(\alpha)+\Delta_{k(L)}(\alpha)}{En_{\zeta(s)+1}(\alpha)+...+En_{\zeta(s)+k(L)}(\alpha)})}{En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)}(1 - \frac{2^{T+1}}{En_{\zeta(s)+1}+....+En_{\zeta(s)+k(L)}})} \tag{25}$$

From theorem 2 and (17) it follows that for any $\alpha \in S^*$ there exist $c_1, c_2 > 0$, $T_0 > 0$ such that for any $T > T_0$

$$c_1 \rho^T \leq En_i(\alpha) \leq c_2 \rho^T$$

Then by (20), (14) and lemma 11 we get that with probability greater than $1-\delta_1$, $\delta_1 = \frac{2 \cdot 2^T}{\varepsilon^2 k(L)}$ the right hand side of (25) is not greater than

$$\frac{En_{\zeta(s)+1}(\alpha) + ... + En_{\zeta(s)+k(L)}(\alpha)(1 + \frac{\varepsilon k(L) + k(L)|\alpha|}{c_1 \rho^T (k(L)-1)})}{En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)}(1 - \frac{2^{T+1}}{L - 2^{T+1}})} \tag{26}$$

From lemma 9 we obtain thjat for any $\varepsilon' > 0, \delta_2 > 0$ there exist $T > 0, L_0 > 0$ such that for any $L > L_0$

$$\frac{\varepsilon k(L) + k(L)|\alpha|}{c_1 \rho^T (k(L) - 1)} < \varepsilon', \frac{2^{T+1}}{L - 2^{T+1}} < \varepsilon'$$

with probability greater than $1 - \delta_2$.

From (26) and lemma 12 we obtain

$$\frac{N(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])}{L} \leq \frac{En_{\zeta(s)+1}(\alpha) + ... + En_{\zeta(s)+k(L)}(\alpha)(1 + \varepsilon')}{En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)}(1 - \varepsilon')} \leq$$

$$\leq \frac{((\pi_{emp}^{(|\alpha|)}(\alpha) + \varepsilon)(En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)}))(1 + \varepsilon')}{En_{\zeta(s)+1} + ... + En_{\zeta(s)+k(L)}(1 - \varepsilon')} =$$

$$= \frac{(\pi_{emp}^{(|\alpha|)}(\alpha) + \varepsilon)(1 + \varepsilon')}{(1 - \varepsilon')} = \pi_{emp}^{(|\alpha|)}(\alpha) + \varepsilon'', \tag{27}$$

where ε'' is a function of $\varepsilon, \varepsilon'$ and $\varepsilon, \varepsilon' \to 0$ implies that $\varepsilon'' \to 0$.

The estimate

$$\pi_{emp}^{(|\alpha|)}(\alpha) - \varepsilon'' \leq \frac{N(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])}{L} \tag{28}$$

is proved similarly.

From (27) and (28) we get

$$\pi_{emp}^{(|\alpha|)}(\alpha) - \varepsilon'' \leq \frac{N(\alpha, [\sum_{i=1}^{\zeta(s)} n_i, \sum_{i=1}^{\zeta(s)+k(L)} n_i])}{L} \leq \pi_{emp}^{(|\alpha|)}(\alpha) + \varepsilon''$$

with probability $1 - \delta$, $\delta = \delta_1 + \delta_2$.

Lemma 13 is completely proved. ∎

From the lemmas 10 and 13 theorem 3 obviously follows.

3.3.3 Proof of the theorem 4

In fact we shall prove more.

Lemma 14 *For all parameters* $\{P(x \to y), P(x \to yz)\}$, $x, y, z \in S$, *for any word* $\alpha = y_1 ... y_m \in S^*$ *and for all* $\varepsilon > 0$ *there exists* L_0 *such that for all* $L > L_0$, *for all* $s > 0$:

$$|\frac{1}{L}\sum_{k=s}^{s+L} \pi_k^{(m)}(y_1, ..., y_m) - \pi_{emp}^{(m)}(y_1, ..., y_m)| < \varepsilon$$

Proof. From theorem 3 it follows that for any word $\alpha \in S^*$, for all $\varepsilon > 0$ there exist $L(\varepsilon, \alpha)$, $t(\varepsilon, \alpha)$ such that for any $L > L(\varepsilon, \alpha)$, $t > t(\varepsilon, \alpha)$, $s > 0$:

$$|\frac{EN_t(\alpha, [s, s+L])}{L} - \pi_{emp}^{(\alpha)}(\alpha)| = |\frac{\sum_{i=s}^{s+L-|\alpha|} p_i^{(|\alpha|)}(\alpha; t)}{L} - \pi_{emp}^{(|\alpha|)}(\alpha)| < 2\varepsilon$$

We pass to the limit $t \to \infty$ and get the result of this lemma. The lemma 14 is completely proved, hence the theorem 4 is also proved. ∎

3.3.4 Proof of theorem 5

We come to the proof of theorem 5 now. This proof is based on the lemma 14 and the main lemma 22 below.

We give the proof for the one-particle correlation functions and the scheme of the proof for the two-particle correlation functions. We start with some auxiliary lemmas.

Lemma 15 *Let* $\xi_1, ..., \xi_n$ *are independent random variables,* $a_k \le \xi_k \le b_k$, $S_n = \sum_{i=1}^n \xi_i$, $k = 1, ..., n$ *then*

$$P\{S_n - ES_n \ge nx\} \le exp\{-\frac{2n^2x^2}{\sum_{k=1}^n (b_k - a_k)^2}\} \tag{29}$$

for all $x > 0$.

Proof. It can be found in [10] or in [7] and we omit it. ∎

From this the following lemma 16 follows.

Lemma 16 *Let* $\xi_1, ..., \xi_n$ *be independent random variables,* $0 < \delta < \frac{1}{2}$
$0 \le \xi_k \le 2$, $S_n = \sum_{i=1}^n \xi_i$, $k = 1, ..., n$, *then for any* $c > 0$:

$$P\{|S_n - ES_n| \ge cn^{\frac{1}{2}+\delta}\} \le 2exp\{-c^2 n^{2\delta}\} \tag{30}$$

Proof. It is sufficient to apply the previous lemma with $x = cn^{-\frac{1}{2}+\delta}$ to the proof of

$$P\{S_n - ES_n \geq cn^{\frac{1}{2}+\delta}\} \leq exp\{-c^2 n^{2\delta}\}$$

If we apply the previous lemma to $\xi' = 2 - \xi$ with $x = cn^{-\frac{1}{2}+\delta}$ then we obtain the following estimate

$$P\{S_n - ES_n \leq -cn^{\frac{1}{2}+\delta}\} \leq exp\{-c^2 n^{2\delta}\}$$

■

Lemma 17 *Let $\xi_1, \xi_2, ...$ be a sequence of independent random variable with values in Z. Let $j_0 > 0$ exists such that $P\{\xi_m = j_0\} \geq P\{\xi_m = j\}$ for all $m > 1$, $j \in Z$. Assume that the greatest common divisor of those $j - j_0$ for which*

$$\frac{1}{logn} \sum_{m=1}^{n} P\{\xi_m = j_0\} P\{\xi_m = j\} \to \infty \ if \, n \to \infty \tag{31}$$

is equal to 1. *Denote*

$$a_i = E\xi_i, \ \sigma_i^2 = D\xi_i, \ s_n^2 = \sum_{i=1}^{n} \sigma_i^2$$

$$A_n = \sum_{i=1}^{n} a_i, \ P_n(z) = P(\xi_n = z), \ i = 1, 2; z > 0$$

Assume also that for some $g, G > 0$:

$$s_n^2 \geq ng \ \sum_{m=1}^{n} E|\xi_m - a_m|^3 \leq nG \tag{32}$$

Then for all $N, n > 0$

$$|s_n P_n(N) - \frac{1}{\sqrt{2\pi}} e^{\frac{-(N-A_n)^2}{2s_n^2}}| \leq \frac{C_0}{\sqrt{n}}, \ C_0 > 0 \tag{33}$$

Proof. It can be found in [10] and we omit it. ■

From this we have the following lemma:

Lemma 18 *Let $\xi_1, \xi_2, ...$ be a sequence of independent random variable with the distribution functions $V_1(x), V_2(x), ...$, where $V_i(x)$ is one of the following distribution functions*

$$F^1(x), ..., F^{|S|}(x), \ F^j(x) = \sum_{k \leq x} p_k^j, j = 1, ..., |S|,$$

where for any $j = 1, ..., |S|$ *the numbers* $p_k^j > 0$, *if* $k = 1, 2$ *and* $p_k^j = 0$ *otherwise. Moreover for any* $1 \leq j \leq |S|$

$$\sum_{k=1}^{2} p_k^j = 1$$

Let

$$S_n = \sum_{i=1}^{n} \xi_i, \ n > 0$$

and

$$P_n(N) = P(S_n = N), n, N > 0$$

Then there exists $N_0 > 0$ *such that for all* $n > N_0, N > 0, \delta,\ 0 < \delta < \frac{1}{2}$:

$$|P_n(N) - P_n(N+1)| < \frac{C}{n^{1-\delta}}, \ C > 0 \tag{34}$$

Proof. Let us use the notation of the previous lemma. All conditions of the lemma 17 hold. Remark also that there exist $c_1, c_2 > 0$ such that for any $n > 0$:

$$c_1 n \leq s_n^2 \leq c_2 n \tag{35}$$

Hence we get from (33)

$$|P_n(N) - \frac{1}{s_n\sqrt{2\pi}} e^{\frac{-(N-A_n)^2}{2s_n^2}}| \leq \frac{C_0}{n}, \ C_0 > 0$$

Applying the triangle inequality, we obtain:

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{1}{s_n\sqrt{2\pi}} e^{\frac{-(N+1-A_n)^2}{2s_n^2}} |1 - e^{\frac{-(N-A_n)^2+(N+1-A_n)^2}{2s_n^2}}|, \tag{36}$$

$C > 0$.

Let us consider

$$-(N - A_n)^2 + (N + 1 - A_n)^2 =$$

$$(-N^2 - A_n^2 + 2NA_n + N^2 + 1 + A_n^2 + 2N - 2NA_n - 2A_n) = 2N - 2A_n + 1$$

From (35) and (36) we obtain

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{c}{\sqrt{n}} e^{\frac{-(N+1-A_n)^2}{2s_n^2}} |1 - e^{\frac{N-A_n+1/2}{s_n^2}}|, \tag{37}$$

$c > 0$. Let us consider two cases.

1) The first case:

$$|N - A_n| > n^{\frac{1}{2}+\delta} \text{ for some } \frac{1}{2} > \delta > 0 \tag{38}$$

From (37) we get

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{c'}{\sqrt{n}} e^{\frac{-(N+1-A_n)^2+(N-A_n+1/2)}{2s_n^2}}, \tag{39}$$

where $c' > 0$. We have

$$-(N+1-A_n)^2 + (N - A_n + 1/2) = -(N-A_n)^2 - 2(N-A_n) - 1 + (N-A_n) + 1/2 =$$

$$-(N-A_n)^2 - (N-A_n) - 1/2 \leq -(N-A_n)^2 - (N-A_n) =$$

$$-(N-A_n)((N-A_n)+1) \leq -n^{\frac{1}{2}+\delta}(n^{\frac{1}{2}+\delta}+1) \leq -n^{1+2\delta}$$

where we applied (38).

From (39) and (35) we have

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{c'}{\sqrt{n}} e^{\frac{-n^{1+2\delta}}{2s_n^2}} \leq \frac{C}{n} + \frac{c'}{\sqrt{n}} e^{\frac{-c'' n^{1+2\delta}}{n}} \leq \tag{40}$$

$$\frac{C}{n} + \frac{c'}{\sqrt{n}} e^{-c'' n^{2\delta}} \leq \frac{c}{n^{1-\delta}}, \text{ for } n > N_0, N_0 > 0$$

where $C, c, c', c'' > 0$

2) The second case:

$$|N - A_n| < n^{\frac{1}{2}+\delta} \text{ for } \delta \text{ as in (38)} \tag{41}$$

Then from (37) we have

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{c'}{\sqrt{n}} |1 - e^{\frac{N-A_n+1/2}{s_n^2}}| \tag{42}$$

since

$$\frac{|N-A_n|}{s_n^2} \leq \frac{n^{\frac{1}{2}+\delta}}{s_n^2} \to 0 \text{ as } n \to \infty$$

Taking into account (35) we rewrite (42) as

$$|P_n(N) - P_n(N+1)| \leq \frac{C}{n} + \frac{c'}{\sqrt{n}}|1 - 1 - \frac{N - A_n + 1/2}{s_n^2} + o(\frac{N - A_n + 1/2}{s_n^2})| \leq \tag{43}$$

$$\frac{C}{n} + \frac{c''}{\sqrt{n}} \frac{n^{\frac{1}{2}+\delta}}{n} \leq \frac{c}{n^{1-\delta}},\ C, c', c'', c > 0$$

Lemma 18 is completely proved. ■

We shall prove the main lemma 19 now. We begin with some notation. In the following it is convenient to consider the process on half-infinite words $\alpha = x_1 x_2 ...$, so $N_t = \infty$. This porcess is defined quite similarly as the process on finite words.

For all $t > 0$, $j > 0$ define

$$\overline{\pi}_j(t) := \left(P(x_j(t) = s_0), ..., P\left(x_j(t) = s_{|S|}\right)\right) \equiv \left(p_j^{(1)}(s_0; t), ..., p_j^{(1)}(s_{|S|}; t)\right),$$

where $s_i \in S$, $i = 1, .., |S|$, $s_i \neq s_j$, $i \neq j$ and

$$\sum_{s \in S} P(x_j(t) = s) = 1$$

Let for all vectors $\vec{a} = \{a(s), s \in S\}$

$$|\vec{a}(s)| := \max_{s \in S} |a(s)|$$

For all vectors $\overline{a} = (a_1, ..., a_{|S|})$ and $\overline{b} = (b_1, ..., b_{|S|})$ we write $\overline{a} \leq \overline{b}$ iff $a_i \leq b_i$ for $i = 1, ..., |S|$

Lemma 19 *For almost all parameters there exist* $T_0 > 0, J_0 > 0$ *such that for all* $t > T_0, j > J_0, \delta,\ 0 < \delta < \frac{1}{4}$

$$|\overline{\pi}_j(t) - \overline{\pi}_{j+1}(t)| < Cj^{-\frac{1}{2}+2\delta},\ C > 0 \tag{44}$$

Remark 20 *This lemma does not hold if for any* $x, y \in S$ $P(x \to y) = 0$. *In this case we can not apply lemmas 17 and 18*

Proof. For all $t > 0$, $1 \leq i \leq N_t$ we define

$$\xi(x_i(t)) = \begin{cases} 2 & \text{, with probability } p(x_i(t)) = \sum_{x,y \in S} P(x_i(t) \to xy) \\ 1 & \text{, with probability } q(x_i(t)) = 1 - p(x_i(t)). \end{cases} \tag{45}$$

$$S_n(t) := \sum_{i=1}^{n} \xi(x_i(t))$$

i.e. $S_n(t)$ is the number of descendants from symbols $x_1(t),...,x_n(t)$. We remark that if we fix $x_1(t), x_2(t),...$ then $\xi(x_1(t)), \xi(x_2(t)),...$ are independent random variables.

Let us assume that the symbol $x_j(t+1)$ appeared from the symbol $x_n(t)$ and write this the following way: $x_n(t) \Rightarrow x_j(t+1)$. It is possible only if $\lceil j/2 \rceil \le n \le j$. It is obvious that

$$\{x_n(t) \Rightarrow x_j(t+1)\} = \{x_n(t) \to x_{j-1}(t+1)\, x_j(t+1)\} \sqcup$$
$$\sqcup\{x_n(t) \to x_j(t+1)\, x_{j+1}(t+1)\} \sqcup \{x_n(t) \to x_j(t+1)\}$$

It is clear that

$$\{x_n(t) \to x_{j-1}(t+1)\, x_j(t+1)\} = \{S_{n-1}(t) = j-2, \xi(x_n(t)) = 2\}$$

$$\{x_n(t) \to x_j(t+1)\, x_{j+1}(t+1)\} = \{S_{n-1}(t) = j-1, \xi(x_n(t)) = 2\}$$

$$\{x_n(t) \to x_j(t+1)\} = \{S_{n-1}(t) = j-1, \xi(x_n(t)) = 1\}$$

Hence

$$\{x_n(t) \Rightarrow x_j(t+1)\} = \{S_{n-1}(t) = j-2, \xi(x_n(t)) = 2\} \sqcup$$

$$\{S_{n-1}(t) = j-1, \xi(x_n(t)) = 2\} \sqcup \{S_{n-1}(t) = j-1, \xi(x_n(t)) = 1\} \tag{46}$$

Let for any $t > 0$, $1 \le j$, and for any events A and B

$$\overline{\pi}_j(t|A) := \left(P(x_j(t) = s_1|A), ..., P\left(x_j(t) = s_{|S|}|A\right)\right),$$

$$\overline{\pi}_j(t;B) := \left(P(\{x_j(t) = s_1\} \cap B), ..., P\left(\{x_j(t) = s_{|S|}\} \cap B\right)\right),$$

$s_k \in S$, $k = 1, ..., |S|$ $s_i \neq s_j$, $i \neq j$.

It is not hard to see that:

$$\overline{\pi}_j(t+1|S_{n-1}(t) = j-2, \xi(x_n(t)) = 2) = \overline{\pi}_n(t)\, Q_r \tag{47}$$

$$\overline{\pi}_j(t+1|S_{n-1}(t) = j-1, \xi(x_n(t)) = 2) = \overline{\pi}_n(t)\, Q_l \tag{48}$$

$$\overline{\pi}_j(t+1|S_{n-1}(t) = j-1, \xi(x_n(t)) = 1) = \overline{\pi}_n(t)\, Q_1 \tag{49}$$

where the matrices Q_r, Q_l, Q_1 were defined in section 3.1.

It was mentioned above $\xi(x_1(t)), \xi(x_2(t)), ...$ are independent for any fixed $x_1(t), x_2(t), ...$. Then from (46) - (49) and the formula of complete probability we obtain $\forall t > 0$, $\forall j, n > 2$

$$\begin{aligned} &\overline{\pi}_j\left((t+1)\,; x_n\left(t\right) \Rightarrow x_j\left(t+1\right)\right) = \\ &\overline{\pi}_n\left(t\right) Q_r \cdot P\left(\xi(x_n(t)) = 2\right) \cdot P\left(S_{n-1}(t) = j-2\right) + \\ &\overline{\pi}_n\left(t\right) Q_l \cdot P\left(\xi(x_n(t)) = 2\right) \cdot P\left(S_{n-1}(t) = j-1\right) + \\ &\overline{\pi}_n\left(t\right) Q_1 \cdot P\left(\xi(x_n(t)) = 1\right) \cdot P\left(S_{n-1}(t) = j-1\right) \end{aligned} \quad (50)$$

From lemma 18 it follows that there exists $N_0 > 0$ such that for any $t > 0, n > N_0, j > 2, \delta$, $0 < \delta < \frac{1}{4}$

$$\left|P\left(S_{n-1}(t) = j-1\right) - P\left(S_{n-1}(t) = j-2\right)\right| \leq \frac{C}{n^{1-\delta}},\ C > 0 \quad (51)$$

Hence from (50) we obtain

$$P(S_{n-1}(t) = j-1)\overline{\pi}_n\left(t\right)\left(Q_1 q\left(x_n\left(t\right)\right) + Q_l p\left(x_n\left(t\right)\right) + Q_r p\left(x_n\left(t\right)\right)\right) - \frac{C_1}{n^{1-\delta}}\overline{\pi}_n\left(t\right) Q_r \leq$$

$$\overline{\pi}_j\left(t+1; x_n(t) \Rightarrow x_j(t+1)\right) \leq$$

$$P(S_{n-1}(t) = j-1)\overline{\pi}_n\left(t\right)\left(Q_1 q\left(x_n\left(t\right)\right) + Q_l p\left(x_n\left(t\right)\right) + Q_r p\left(x_n\left(t\right)\right)\right) + \frac{C_1}{n^{1-\delta}}\overline{\pi}_n\left(t\right) Q_r \quad (52)$$

where $C_1 > 0$, $p(x_n(t))$, $q(x_n(t))$ were defined above in (45).

It is clear that

$$\overline{\pi}_j\left(t+1\right) = \sum_{n=\lceil j/2 \rceil}^{j} \overline{\pi}_j\left(t+1; x_n\left(t\right) \Rightarrow x_j\left(t+1\right)\right) \quad (53)$$

We would like to estimate $\overline{\pi}_j\left(t+1\right)$. Let for any $j > 0, n > 1$:

$$A(j) := \{n : |ES_{n-1}(t) - j| \leq (n-1)^{\frac{1}{2}+\delta}\} \quad (54)$$

and

$$B(j) := \{n : |ES_{n-1}(t) - j| > (n-1)^{\frac{1}{2}+\delta}\}$$

We need the following auxiliary lemma.

Lemma 21 *There exists J_0 such that for all $t, j > J_0$:*

$$\#A(j) \leq cj^{\frac{1}{2}+\delta},\ c > 0 \quad (55)$$

Proof. It is sufficient to prove that

$$\#\{n : |ES_n(t) - j| \le n^{\frac{1}{2}+\delta}\} \le cj^{\frac{1}{2}+\delta},\ c > 0 \tag{56}$$

It is obvious that fro any $n > 0$ there exists $C \equiv C(n)$, $1 \le C \le 2$ such that $S_n(t) = Cn$. Let us find from (56) estimates on n. If (56) holds then $|ES_n(t) - j| \le n^{\frac{1}{2}+\delta}$, hence $|Cn - j| \le n^{\frac{1}{2}+\delta}$ and

$$Cn - n^{\frac{1}{2}+\delta} \le j \le Cn + n^{\frac{1}{2}+\delta}$$

Let us consider

$$Cn - n^{\frac{1}{2}+\delta} \le j$$

hence

$$C - n^{-\frac{1}{2}+\delta} \le \frac{j}{n} \Rightarrow n \le \frac{j}{C - n^{-\frac{1}{2}+\delta}} = \frac{j}{C}(1 + \frac{n^{-\frac{1}{2}+\delta}}{C} + o(n^{-\frac{1}{2}+\delta}))$$

since $n \le j$ we get

$$n \le \frac{j}{C} + C_1 j^{\frac{1}{2}+\delta},\ C_1 > 0 \tag{57}$$

The estimate

$$n \ge \frac{j}{C} - C_1 j^{\frac{1}{2}+\delta} \tag{58}$$

can be obtained quite similarly. From (57) and (58) we get

$$\#\{n : |ES_n(t) - j| \le n^{\frac{1}{2}+\delta}\} \le 2C_1 j^{\frac{1}{2}+\delta}$$

■

From lemma 16 and the inequality $\frac{j}{2} \le n \le j$ it follows

$$\left| \sum_{n \in B(j)} \overline{\pi}_j \left(t+1; x_n\left(t\right) \Rightarrow x_j\left(t+1\right)\right) \right| \le \varepsilon'_j \tag{59}$$

where

$$\varepsilon'_j = cj \cdot exp\{-\left(\frac{j}{2}\right)^{2\delta}\},\ c > 0 \tag{60}$$

From (55) and the inequality $\frac{j}{2} \le n \le j$ it is obvious that

$$\left| \sum_{n \in A(j)} \frac{C_1}{n^{1-\delta}} \overline{\pi}_n\left(t\right) Q_r \right| \le \varepsilon''_j \tag{61}$$

where

$$\varepsilon_j'' = c' j^{-\frac{1}{2}+2\delta},\ c' > 0 \tag{62}$$

Put

$$\varepsilon_j = \varepsilon_j' + \varepsilon_j'' \tag{63}$$

and let $\overline{\varepsilon}_j$ ve the vector of length $|S|$ with components ε_j. From (60)-(63) it is clear that

$$\overline{\varepsilon}_j \to 0 \text{ if } j \to \infty, \delta < \frac{1}{4} \tag{64}$$

From (53) , (52) , (59), (61) and (63) we get

$$\sum_{n\in A(j)} (P(S_{n-1}(t) = j-1)\overline{\pi}_n(t)(Q_1 q(x_n(t)) + Q_l p(x_n(t)) + Q_r p(x_n(t)))) - \overline{\varepsilon}_j \leq$$

$$\leq \overline{\pi}_j(t+1) \leq \tag{65}$$

$$\leq \sum_{n\in A(j)} (P(S_{n-1}(t) = j-1)\overline{\pi}_n(t)(Q_1 q(x_n(t)) + Q_l p(x_n(t)) + Q_r p(x_n(t)))) + \overline{\varepsilon}_j$$

And by analogy

$$\sum_{n\in A(j-1)} (P(S_{n-1}(t) = j-2)\overline{\pi}_n(t)(Q_1 q(x_n(t)) + Q_l p(x_n(t)) + Q_r p(x_n(t)))) - \overline{\varepsilon}_{j-1} \leq$$

$$\leq \overline{\pi}_{j-1}(t+1) \leq \tag{66}$$

$$\leq \sum_{n\in A(j-1)} (P(S_{n-1}(t) = j-2)\overline{\pi}_n(t)(Q_1 q(x_n(t)) + Q_l p(x_n(t)) + Q_r p(x_n(t)))) + \overline{\varepsilon}_{j-1}$$

From (51) and (66) it follows

$$\sum_{n\in A(j-1)} (P(S_{n-1}(t) = j-1)\overline{\pi}_n(t)(Q_1 q(x_n(t)) + Q_l p(x_n(t)) + Q_r p(x_n(t)))) - \overline{\varepsilon'}_{j-1} \leq$$

$$\leq \overline{\pi}_{j-1}(t+1) \leq \tag{67}$$

$$\leq \sum_{n\in A(j-1)} (P(S_{n-1}(t)=j-1)\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right))+\overline{\varepsilon'}_{j-1}$$

where

$$\overline{\varepsilon'}_{j-1}=\sum_{n\in A(j-1)} \frac{C}{n^{1-\delta}}\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right)+\overline{\varepsilon}_{j-1} \tag{68}$$

From (55), (60), (62), (63) and (68) it is evident that

$$\overline{\varepsilon'}_{j-1}\leq C_2 j^{-\frac{1}{2}+2\delta}\rightarrow 0 \text{ when } j\rightarrow\infty \quad C_2>0 \tag{69}$$

Let

$$\varepsilon'_{j-1}=|\overline{\varepsilon'}_{j-1}| \tag{70}$$

Let us estimate $|\overline{\pi}_j(t)-\overline{\pi}_{j-1}(t)|$ taking into consideration (65) and (67)

$$|\overline{\pi}_j(t)-\overline{\pi}_{j-1}(t)|\leq\varepsilon_j+\varepsilon'_{j-1}+$$

$$|\sum_{n\in A(j)} (P(S_{n-1}(t)=j-1)\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right))-$$

$$\sum_{n\in A(j-1)} (P(S_{n-1}(t)=j-1)\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right))|=$$

$$|\sum_{n\in A(j)\backslash A(j-1)} (P(S_{n-1}(t)=j-1)\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right))-$$

$$\sum_{n\in A(j-1)\backslash A(j)} (P(S_{n-1}(t)=j-1)\overline{\pi}_n(t)\left(Q_1 q\left(x_n\left(t\right)\right)+Q_l p\left(x_n\left(t\right)\right)+Q_r p\left(x_n\left(t\right)\right)\right))|+$$

$$+\varepsilon_j+\varepsilon'_{j-1} \tag{71}$$

If either $n\notin A(j)$ or $n\notin A(j-1)$ then $|ES_{n-1}(t)-j|\geq cn^{\frac{1}{2}+\delta}$, $c>0$ and we can apply lemma 16. Then from (71) and the obvious inequality $\frac{j}{2}\leq n\leq j$ we obtain:

$$|\overline{\pi}_j(t)-\overline{\pi}_{j-1}(t)|\leq c\cdot j\cdot exp\{-c_1 j^{2\delta}\}+\varepsilon_j+\varepsilon'_{j-1},\ c,c_1>0$$

Taking into consideration (69), (70) and (60), (62), (63) we get from (71): there exists $J_0 > 0$ such that fro any $j > J_0$:

$$|\overline{\pi}_j(t) - \overline{\pi}_{j-1}(t)| \leq Cj^{-\frac{1}{2}+2\delta},\ C > 0$$

The lemma 19 is proved.

Lemma 22 *For almost all parameters, for any $\alpha \in S^*$ there exist $T_0,\ J_0 > 0$ such that for all $t > T_0$, all $j,\ J_0 \leq j,\ \delta,\ 0 < \delta < \frac{1}{4}$:*

$$|p_j^{(|\alpha|)}(\alpha;t) - p_{j+1}^{(|\alpha|)}(\alpha;t)| < Cj^{-\frac{1}{2}+2\delta},\ C > 0 \tag{72}$$

Proof. The case $|\alpha| = 1$ was considered in lemma 19. The proof for the case $|\alpha| > 1$ is completely similar to the proof for the case $|\alpha| = 1$, but it needs a new notation. We give only a schema of the proof for the case $|\alpha| = 2$.

Recall that for any $x, y \in S$ and any functions $f : S \to R, g : S \times S \to R$ we put

$$(\Delta f)(x,y) = \sum_{z \in S} f(z) P(z \to xy) \tag{73}$$

$$(g(\cdot,\cdot) A \otimes B)(x,y) := \sum_{m,n \in S} g(m,n) A(m,x) B(n,y) \tag{74}$$

where A, B are $S \times S$-matrices. We need also the following notations. Let $\alpha = s_1 s_2$, where $s_1, s_2 \in S$, then for any $t > 0, j \in Z^+$ put

$$\pi_{j,j+1}^t(s_1,s_2) := P(x_n(t) = s_1, x_{n+1}(t) = s_2) \equiv p_j^{(2)}(\alpha;t),$$

We can write the recurrent equation on $\pi_{j,j+1}^{t+1}$ now. We shall use the notation of the previous lemma $(S_n(t), \xi(x_i(t))$, etc.)

We define $\pi_{j,j+1}^{t+1}(s_1, s_2; x_n(t) \Rightarrow x_j(t+1))$ equal to

$$P(x_j(t+1) = s_1, x_{j+1}(t+1) = s_2; x_n(t) \Rightarrow x_j(t+1)),$$

i.e. it is the distribution of $(x_j(t+1), x_{j+1}(t+1))$ when $x_j(t+1)$ was born from $x_n(t)$.

At first, we find $\pi_{j,j+1}^{t+1}(\cdot,\cdot; x_n(t) \Rightarrow x_j(t+1))$ (see Fig.1):

$$\pi_{j,j+1}^{t+1}(\cdot,\cdot; x_n(t) \Rightarrow x_j(t+1)) =$$

$$\pi_{n,n+1}^t Q_1 \otimes Q_1 P(\xi(x_n(t)) = 1) P(\xi(x_{n+1}(t)) = 1) P(S_{n-1}(t) = j-1) +$$

$$\pi_{n,n+1}^t Q_1 \otimes Q_l P(\xi(x_n(t)) = 1) P(\xi(x_{n+1}(t)) = 2) P(S_{n-1}(t) = j-1) +$$

$$\pi^t_{n,n+1} Q_r \otimes Q_1 P\left(\xi(x_n(t)) = 2\right) P\left(\xi(x_{n+1}(t)) = 1\right) P\left(S_{n-1}(t) = j-2\right) +$$

$$\pi^t_{n,n+1} Q_r \otimes Q_l P\left(\xi(x_n(t)) = 2\right) P\left(\xi(x_{n+1}(t)) = 2\right) P\left(S_{n-1}(t) = j-2\right) +$$

$$\left(\Delta\pi_n(t)\right) P\left(\xi(x_n(t)) = 2\right) P\left(S_{n-1}(t) = j-1\right) \sim$$

$$\sim \pi^t_{n,n+1} P\left(S_{n-1}(t) = j-1\right) \left[Q_1 q\left(x_n(t)\right) \otimes Q_1 q\left(x_{n+1}(t)\right) +\right.$$

$$Q_1 q\left(x_n(t)\right) \otimes Q_l p\left(x_{n+1}(t)\right) + Q_r p\left(x_n(t)\right) \otimes Q_1 q\left(x_{n+1}(t)\right) +$$

$$\left. Q_r p\left(x_n(t)\right) \otimes Q_l p\left(x_{n+1}(t)\right)\right] + \Delta\pi_n(t) P\left(S_{n-1}(t) = j-1\right) \tag{75}$$

where we applied (51) and the obvious inequality $\frac{j}{2} \leq n \leq j$.

$$\pi^{t+1}_{j,j+1}(\cdot,\cdot) = \sum_{n=\lceil j/2 \rceil}^{j} \pi^{t+1}_{j,j+1}\left(\cdot,\cdot; x_n(t) \Rightarrow x_j(t+1)\right) \tag{76}$$

Further on, we should estimate $|\pi^{t+1}_{j,j+1} - \pi^{t+1}_{j+1,j+2}|$, where for all matrices $A = \{A(s_1,s_2), s_1, s_2 \in S\}$ $|A| = \max_{s_1,s_2 \in S} |A(s_1,s_2)|$. This part of the proof completely repeats the proof of the previous lemma and we omit it. ■

We finish the proof of the theorem 5 now. This proof is based on the lemmas 14 and 22.

From lemma 14 it follows that for all $\varepsilon > 0$ there exists $L_0 > 0$ such that for any $L > L_0, s > 0:$

$$\left|\frac{\sum_{i=s}^{s+L-|\alpha|} \pi_i^{(|\alpha|)}(\alpha)}{L} - \pi_{emp}^{(|\alpha|)}(\alpha)\right| < \varepsilon \tag{77}$$

Hence for all $\varepsilon > 0$ there exists $L_0 > 0$ such that for any $L > L_0, s > 0:$

$$\left|\frac{\sum_{i=s}^{s+L-|\alpha|} \pi_i^{(|\alpha|)}(\alpha)}{L+1-|\alpha|} - \pi_{emp}^{(|\alpha|)}(\alpha)\right| < \varepsilon \tag{78}$$

Passing to the limit $t \to \infty$ in lemma 22, and taking into account the theorem 1, we obtain that there exists $I \equiv I(\varepsilon, L)$ such that for any $i > I:$

$$|\pi_i^{(|\alpha|)}(\alpha) - \pi_{i+1}^{(|\alpha|)}(\alpha)| < \frac{\varepsilon}{L} \tag{79}$$

Let us choose $s > I$, then for any $m, n \in [s, s+L-|\alpha|]:$

$$|\pi_m^{(|\alpha|)}(\alpha) - \pi_n^{(|\alpha|)}(\alpha)| < \varepsilon \tag{80}$$

Hence

$$\left| \frac{\sum_{i=s}^{s+L-|\alpha|} \pi_i^{(|\alpha|)}(\alpha)}{L+1-|\alpha|} - \pi_s^{(|\alpha|)}(\alpha) \right| < \varepsilon \tag{81}$$

From (78) and (81) we get

$$\left| \pi_s^{(|\alpha|)}(\alpha) - \pi_{emp}^{(|\alpha|)}(\alpha) \right| < 2\varepsilon$$

Thus $\lim_{s\to\infty} \pi_s^{(|\alpha|)}(\alpha)$ exists and the theorem is proved. ■

References

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

[2] V.A.Malyshev. Random Grammars. Russian Math. Surveys, 1998, v. 53, No. 2, 107-134.

[3] K.B. Athreya, A.N.Vidyashankar. Large deviation rates for supercritical and critical branching processses. In IMA, v. 74 "Classical and Modern Branching Processes", 1997, pp. 3-18.

[4] F R. Gantmacher (1958) Theory of Matrices.

[5] G. Hansel, D. Perrin. Mesures de probabilites rationnelles. In "Mots", 1990, pp. 335-357. Editions Hermes, Paris.

[6] T.E.Harris (1963), The theory of branching processes, Springer-Verlag, Berlin.

[7] Hoeffding W.(1963) Probability inequalities for sums of bounded random variables, J.Amer. Stat. Assoc. 58, 301, 13-20.

[8] Kesten H. Stigum B.P.(1966) A limit theorem for multidimensional Galton-Watson processes, Ann.Math.Stat. v. 37, No. 5, pp.1211-1223.

[9] H. Kesten, B.P. Stigum (1968) Additional limit theorems for indecomposable multidimensional Galton-Watson processes, Ann. Math. Stat. v. 37, No. 6, pp.1463-1481

[10] V.V.Petrov (1972) Sums of independent random variables. "Nauka",Moscow

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399