INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

New analysis of the asymptotic behavior of the Lempel-Ziv compression algorithm

Philippe Jacquet — Wojciech Szpankowski

N° ????

April 2010

Thème COM

Rapport de recherche

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

New analysis of the asymptotic behavior of the Lempel-Ziv compression algorithm

Philippe Jacquet , Wojciech Szpankowski

Thème COM — Systèmes communicants
Projets Hipercom

Rapport de recherche n° ???? — April 2010 — 18 pages

Abstract: We give a new analysis and proof of the Normal limiting distribution of the number of phrases in the 1978 Lempel-Ziv compression algorithm on random sequences built from a memoriless source. This work is a follow-up of our last paper on this subject in 1995. The analysis stands on the asymptotic behavior of a DST obtained by the insertion of random sequences. Our proofs are augmented of new results on moment convergence, moderate and large deviations, redundancy analysis.

Key-words: data compression, digital search tree, generating functions, depoissonization, renewal processes

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

Nouvelle analyse du comportement asymptotique de l'algorithme de compression de Lempel-Ziv

Résumé : Nous donnons une nouvelle analyse et preuve de la convergence vers la loi normale des performances de l'algorithme de compression de Lempel et Ziv 1978 sur des séquences aléatoires construites à partir d'une source sans mémoire. Ce travail est une continuation de notre papier sur le même sujet de 1995. L'analyse repose sur le comportement asymptotique d'un arbre digital de recherche obtenu par l'insertion de séquences aléatoires. Nos preuves sont augmentées de nouveaux résultats sur la convergence des moments, les moyennes et grandes déviations et l'analyse de la redondance.

Mots-clés : compression de données, arbre digital de recherche, fonctions génératrices, dépoissonisation, processus à renouvellement

1 Ziv Lempel compression algorithm

We investigate the number of phrases generated by the compression algorithm Lempel-Ziv 1978 [**?**]. The algorithm consists in arranging the text to be compressed in consecutive phrases such that the next phrase is the unique prefix of the remaining text that has a previous phrase as largest prefix. In other words the next phrase can be identified by its prefix in the previously listed phrases plus one symbol: namely one pointer and a symbol.

It is convenient to see the phrases organized in a digital search tree. Imagine a root. The first phrase is the first symbol, say a pending on the root. Depending if the next symbol is again a a, then the next phrase will be with two symbols, the last symbol being appended to the node with a. The new phrase can be read from the root to the new symbol. If the next symbol was different of a, say b, then the new phrase is simply b (the largest phrase prefix is therefore empty) and is directly appended to root.

Let a text w written on alphabet $\mathcal{A}$, and $\mathcal{T}(w)$ is the digital tree created by the algorithm. To each node in $\mathcal{T}(w)$ corresponds a phrase in the parsing algorithm. Let $L(w)$ be the path length of $\mathcal{T}(w)$: the sum of all the node distances to the root. We should have $L(w) = |w|$. To make this rigorous we have to assume that the text w ends with an extra symbol that never occurs before in the text. The knowledge of $\mathcal{T}(w)$ without the node sequence order is not sufficient to recover the original text w. But if we know the sequence order of node creation in the tree, then we can reconstruct the original text w.

The compression code $C(w)$ is in fact a description of $\mathcal{T}(w)$, node by node in their order of creation, each node being identified by a pointer to its parent node in the tree and the symbol that label the edge that link the parent to the node. Since this description contains the sequence order of node creation, then the original text can be recovered. Every node points to parents that have been inserted *before*, therefore pointer for the kth node costs at most $\lceil \log_2 k \rceil$, and the next symbol costs $\lceil \log_2 |\mathcal{A}| \rceil$. A negligible economy can be done by fusionning the two informations in a pointer of cost $\lceil \log_2(k|\mathcal{A}|) \rceil$. The compressed code has therefore length

$$|C(w)| = \sum_{k=1}^{k=M(w)} \lceil \log_2(k|\mathcal{A}|) \rceil \ . \tag{1}$$

Notice that the code is self-consistent and does not need the *a priori* knowledge of the text length, since the pointer field is length is a simple function of node sequence order. Formula 1 gives the estimate

$$|C(w)| \leq M(w)(\log_2(M(w)) + 1)\lceil \log_2 |\mathcal{A}|) \rceil \ . \tag{2}$$

This modification does not change the asymptotic analysis, since $\log_2(M(w)!) \sim M(w)(\log_2(M_w) - 1)$ and $\log_2(M(w)) = \log_2(|w|) + O(1)$ Let n be an integer, we denote M_n be the number of phrases $M(w)$ when the original text w is random with fixed length n. Let $H_n = nh$ be the entropy of the text w, $h = -\sum_{a\in\mathcal{A}} p_a \log p_a$ being the entropy rate of the memoriless source (p_a is the probability of symbol a). Let $\Phi(x) = \int_x^\infty e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}}$ denotes the normal error function.

We denote $h_2 = \sum_{a\in\mathcal{A}} p_a(\log p_a)^2$ and

$$\eta = -\sum_{k\geq 2} \frac{\sum_{a\in\mathcal{A}} p_a^k \log p_a}{1-\sum_{a\in\mathcal{A}} p_a^k} \ . \tag{3}$$

We also introduce the functions

$$\begin{aligned}
\beta(m) &= \frac{h_2}{2h} + \gamma - 1 - \eta + \Delta_1(\log m) && (4)\\
&\quad + \frac{1}{m}\left(\log m + \frac{h_2}{2h} - \gamma - \sum_{a\in\mathcal{A}} \log p_a + \eta\right) && (5)\\
v(m) &= \frac{m}{h}\left((\frac{h_2}{h^2} - 1)\log m + c_2 + \Delta_2(\log m)\right) && (6)
\end{aligned}$$

We prove the following theorem

Theorem 1. *Let $\ell(m) = \frac{m}{h}(\log m + \beta(m))$. The random variable M_n has a mean $\mathbf{E}(M_n)$ and variance $\mathrm{Var}(M_n)$ which and for all $\delta > \frac{1}{2}$ satisfy the following estimates for all $\delta > \frac{1}{2}$, when $n \to \infty$:*

$$\begin{aligned}
\mathbf{E}(M_n) &= \ell^{-1}(n) + O(n^\delta) \sim \frac{H_n}{\log_2 n} && (7)\\
\mathrm{Var}(M_n) &\sim \frac{v(\ell^{-1}(n))}{(\ell'(\ell^{-1}(n)))^2} = O(n) && (8)
\end{aligned}$$

When $n \to \infty$ the distribution random variable $\frac{M_n - \mathbf{E}(M_n)}{\sqrt{\mathrm{Var}(M_m)}}$ tends to a normal distribution with mean 0 and variance 1. that is for any given real number x:

$$\lim_{n\to\infty} P(M_n > \mathbf{E}(M_n) + x\sqrt{\mathrm{Var}(M_n)}) = \Phi(x) \ . \tag{9}$$

and for all integer k:

$$\lim_{n\to\infty} \mathbf{E}\left((\frac{M_n - \mathbf{E}(M_n)}{\sqrt{\mathrm{Var}(M_n)}})^k \right) = \mu_k \ . \tag{10}$$

where $\mu_k = 0$ when k is odd and $\mu_k = \frac{k!}{2^{k/2}(\frac{k}{2})!}$.

This theorem is slightly augmented from [2] since it contains the convergence of M_n in moment. We also have large and moderate deviation new results

Theorem 2.

- **Large deviation:** *For all $\delta > \frac{1}{2}$ there exists $\varepsilon > 0$, $B > 0$ and $\beta > 0$ such that for all $y > 0$ we have*

$$P(|M_n - \mathbf{E}(M_n)| > +yn^\delta) \leq A\exp(-\beta n^\varepsilon \frac{y}{(1+n^{-\varepsilon}y)^\delta}) \tag{11}$$

- **Moderate deviation** *Let $\delta < \frac{1}{6}$ and $A > 0$, there exists $B > 0$ such that for all integer m and for all non-negative real number $x < An^\delta$:*

$$P(|M_n - \mathbf{E}(M_n)| \geq x\sqrt{\mathrm{Var}(M_n)}) \leq Be^{-\frac{x^2}{2}} \ . \tag{12}$$

This prove that the average compression rate converge to the entropy rate:

Corollary 1. *The average compression rate* $\frac{|C(w)|}{|w|\log_2(|\mathcal{A}|)}$ *converges to the entropy rate that is*

$$\lim_{n\to\infty} \frac{\mathbf{E}(M_n \log_2 M_n + M_n)}{n} = h \ . \tag{13}$$

The result can also be extended to the redundancy estimate:

Corollary 2. *The average redundancy rate* $\frac{|C(w)|}{|w|\log_2(|\mathcal{A}|)} - h$ *satisfies the estimate for all* $\delta > \frac{1}{2}$*:*

$$\frac{\mathbf{E}(M_n \log_2 M_n + M_n)}{n} - h = \frac{\beta(\ell^{-1}(n))}{\log \ell^{-1}(n) + \beta(\ell^{-1}(n))} + O(n^{\delta-1}) \sim \frac{\beta\left(h\frac{n}{\log n}\right)}{\log n}) \ . \tag{14}$$

This theorem was proven in [3] but we will give a new proof. Notice that this estimate of the redundancy proven in [3] is smaller than the previously obtained estimates of $O(\frac{\log\log n}{\log n})$ obtained via probabilistic methods.

2 Equivalence phrase number and path length in Digital Search Tree

2.1 Renewal argument

Our aim is to derive an estimate of probability distribution of M_n. We take the same points as in [2]. In this section we assume that our original text is in fact prefix of an infinite sequence X built from a memoriless source on alphabet $\mathcal{A}$. For $a \in \mathcal{A}$ we denote p_a the probability of symbol a. We assume that for all $a \in \mathcal{A}$, we have $p_a \neq 0$, otherwise this would be equivalent to restrict our analysis on the sub-alphabet with non zero probabilities. We build the Digital Search Tree by parsing the infinite sequence X up to the mth phrase.

Let L_m be the random variable that denotes the path obtained with the mth phrase. The quantity M_n is in fact the number of phrase needed to get the DST absorbing the nth symbol of X. This observation leads to the identy valid for all integers n and m:

$$P(M_n = m) = P(L_{m-1} < n \ \& \ L_m \geq n) \ , \tag{15}$$

or simpler

$$P(M_n \leq m) = P(L_m \geq n) \ , \tag{16}$$

and equivalently

$$P(M_n > m) = P(L_m < n) \ . \tag{17}$$

Proof of theorem 2. **Large deviation:** With equation (17) we have

$$\begin{aligned} P(M_n > \ell^{-1}(n) + yn^\delta) &= P(M_n > \lfloor \ell^{-1}(n) + yn^\delta \rfloor) \qquad (18) \\ &= P(L_{\lfloor \ell^{-1}(n)+yn^\delta \rfloor} < n) \ . \qquad (19) \end{aligned}$$

We take the estimate $\mathbf{E}(L_m) = \ell(m) + O(1)$.

$$\mathbf{E}(L_{\lfloor \ell^{-1}(n)+yn^\delta \rfloor}) = \ell(\ell^{-1}(n) + yn^\delta) + O(1) \tag{20}$$

Since function $\ell(.)$ is convex and $\ell(0) = 0$, we have for all real numbers $a > 0$ and $b > 0$

$$\ell(a+b) \geq \ell(a) + \frac{\ell(a)}{a}b \tag{21}$$

$$\ell(a-b) \leq \ell(a) - \frac{\ell(a)}{a}b \tag{22}$$

Applying inequality (21) to $a = \ell^{-1}(n)$ and $b = yn^\delta$ yields

$$n - \mathbf{E}(L_{\lfloor \ell^{-1}(n)+yn^\delta \rfloor}) \leq -y\frac{n}{\ell^{-1}(n)}n^\delta + O(1) \tag{23}$$

And thus

$$P(L_{\lfloor \ell^{-1}(n)+yn^\delta \rfloor} < n) \leq P(L_m - \mathbf{E}(L_m) < -xm^\delta + O(1)) \ , \tag{24}$$

identifying $m = \lfloor \ell^{-1}(n) + yn^\delta \rfloor$ and $x = \frac{n}{\ell^{-1}(n)}\frac{n^\delta}{m^\delta}y$ Using theorem 6: that is for all $x > 0$ and for all m, there exists $\varepsilon > 0$ A:

$$P(L_m - \mathbf{E}(L_m) < xm^\delta) < Ae^{-\beta x m^\varepsilon} \ . \tag{25}$$

or $P(L_m - \mathbf{E}(L_m) < xm^\delta + O(1)) \leq Ae^{-\beta x m^\varepsilon + O(n^{\varepsilon-\delta})} \leq A'e^{-\beta x m^\varepsilon}$ for some $A' > A$ we get

$$P(M_n > \ell^{-1}(n) + yn^\delta) \leq A' \exp(-\beta x m^\varepsilon) \tag{26}$$

With little effort, since $\ell^{-1}(n) = \Omega(\frac{n}{\log n})$ it turns that $\frac{n^\delta y}{m^\delta} \leq \beta' \frac{y}{1+n^{-\varepsilon'}}$ for some $\varepsilon' > 0$ and $\beta' > 0$.

For the opposite case we have

$$P(M_n < \ell^{-1}(n) - yn^\delta) = P(L_{\lfloor \ell^{-1}(n)-yn^\delta \rfloor} > n) \ . \tag{27}$$

and

$$\mathbf{E}(L_{\lfloor \ell^{-1}(n)-yn^\delta \rfloor}) = \ell(\ell^{-1}(n) - yn^\delta) + O(1) \tag{28}$$

using inequality (22) we have

$$n - \mathbf{E}(L_{\lfloor \ell^{-1}(n)-yn^\delta \rfloor}) \geq y\frac{n}{\ell^{-1}(n)}n^\delta + O(1) \tag{29}$$

And thus

$$P(L_{\lfloor \ell^{-1}(n)-yn^\delta \rfloor} > n) \leq P(L_m - \mathbf{E}(L_m) > xm^\delta + O(1)) \ , \tag{30}$$

identifying $m = \lfloor \ell^{-1}(n) - yn^\delta \rfloor$ and $x = \frac{n}{\ell^{-1}(n)}\frac{n^\delta}{m^\delta}y$. This case is easier than the previous case since we have now $m < \ell^{-1}(n)$ and we don't need the correcting term $\frac{1}{(1+yn^\varepsilon)^\delta}$.

Moderate deviation: It is essentially the same proof except that we consider $y\frac{s_n}{\ell'(\ell^{-1}(n))}$ with $s_n = \sqrt{v(\ell^{-1}(n))}$ instead of yn^δ and $y = O(n^{\delta'})$ for some $\delta' < \frac{1}{6}$. Thus $y\frac{s_n}{\ell'(\ell^{-1}(n))} = O(n^{\frac{1}{2}+\delta}) = o(n)$. If $y > 0$

$$P(M_n > \ell^{-1}(n) + y\frac{s_n}{\ell'(\ell^{-1}(n))}) = P(L_m < n) \tag{31}$$

with $m = \lfloor \ell^{-1}(n) + y\frac{s_n}{\ell'(\ell^{-1}(n))} \rfloor$. We use the estimate

$$\ell(a+b) = \ell(a) + \ell'(a)b + o(1) \tag{32}$$

when $b = o(a)$. Thus

$$\ell(\ell^{-1}(n) + y\frac{s_n}{\ell'(\ell^{-1}(n))}) = n + yv(\ell^{-1}(n)) + o(1) \ . \tag{33}$$

and

$$n = \mathbf{E}(L_m) - yv(m) + O(1) \tag{34}$$

Refering again to theorem 6 we use the fact that $P(L_m < E(L_m) - yv(m) + O(1)) \leq A\exp(\frac{y^2}{2})$, the term $O(1)$ inducing a term $\exp(O(\frac{y^2}{v(m)}) = \exp(o(1))$ absorbed in factor A. The prrof for $y < 0$ follows a similar path. □

Proof of theorem 1. The previous result implies that for all $1 > \delta > \frac{1}{2}$ $\mathbf{E}(L_m) = \ell^{-1}(n) + O(n^\delta)$. Indeed

$$|\mathbf{E}(M_n) - \ell^{-1}(n)| \leq n^\delta \int_0^\infty P(|M_n - \ell^{-1}(n)| > yn^\delta)dy \tag{35}$$

$$\leq A\int_0^\infty \exp(-\beta\frac{yn^\varepsilon}{(1+yn^{-\varepsilon})^\delta})dy = O(n^\delta) \tag{36}$$

For a given y fixed taking the same proof as for moderate devation we have

$$P(M_n > \ell^{-1}(n) + y\frac{s_n}{\ell'(\ell^{-1}(n))}) = P(L_{\lfloor \ell^{-1}(n) - yn^\delta \rfloor} < n) \tag{37}$$

Let $m = \lfloor \ell^{-1}(n) - yn^\delta \rfloor$. We know that $n - E(L_m) = yv(\ell^{-1}(n)) + O(1)$ and $v(\ell^{-1}(n)) = \sqrt{\mathrm{Var}(L_m)} + o(1)$. Therefore

$$P(M_n > \ell^{-1} + y\frac{s_n}{\ell'(\ell^{-1}(n))}) = P(L_m < E(L_m) + y\sqrt{\mathrm{Var}(L_m)} + O(1)) \tag{38}$$

Assume that the $|O(1)| \leq B$

$$P(L_m < E(L_m) + (y + \frac{A}{\sqrt{\mathrm{Var}(L_m)}})\sqrt{\mathrm{Var}(L_m)}) \leq P(M_n > \ell^{-1} + y\frac{s_n}{\ell'(\ell^{-1}(n))}) \tag{39}$$

since for all y' $\lim_{m\to\infty} P(L_m < E(L_m) + y'\sqrt{\mathrm{Var}(L_m)}) = \Phi(y')$ and therefore

$$\limsup_{m\to\infty} P(L_m < E(L_m) + (y + \frac{A}{\sqrt{\mathrm{Var}(L_m)}}) = \Phi(y) \ . \tag{40}$$

similarly we prove that

$$\liminf_{m\to\infty} P(L_m < E(L_m) + (y - \frac{A}{\sqrt{\mathrm{Var}(L_m)}}) = \Phi(y) \ . \tag{41}$$

Therefore $\lim_{m\to\infty} P(M_n > \ell^{-1}(n) + y\frac{s_n}{\ell'(\ell^{-1}(n))}) = \Phi(y)$. Similarly we prove $\lim_{m\to\infty} P(M_n < \ell^{-1}(n) - y\frac{s_n}{\ell'(\ell^{-1}(n))})$. This prove two things: First this prove that $(L_m - \ell^{-1}(n))\frac{\ell'(\ell(^{-1}(n))}{s_n}$ tends to the normal distribution in probability.

Second, since by the moderate deviation result the $(L_m - \ell^{-1}(n))\frac{\ell'(\ell(^{-1}(n))}{s_n}$ has bounded moments then, by virtue of the dominated convergence:

$$\lim_{n\to\infty} \mathbf{E}((L_m - \ell^{-1}(n))\frac{\ell'(\ell(^{-1}(n))}{s_n})) = 0 \quad (42)$$

$$\lim_{n\to\infty} \mathbf{E}(\left((L_m - \ell^{-1}(n))\frac{\ell'(\ell(^{-1}(n))}{s_n})\right)^2) = 1 \quad (43)$$

or in other words $\mathrm{Var}(M_n) \sim \frac{v(\ell^{-1}(n))}{(\ell'(\ell^{-1}(n)))^2}$ which proves the variance estimate. □

2.2 path length in Digital Search Tree

We denote L_m the path length of the resulting tree after m insertion. We use the notation $L_m(u) = E(u^{L_m})$. In the following $\mathbf{k}$ is a tupple in $\mathbb{N}^{|\mathcal{A}|}$ and k_a for $a \in \mathcal{A}$ is the component of $\mathbf{k}$ for symbol a. We take the analysis in [2]Due to the fact that the source of sequence X is memoriless, the phrases are independent. Therefore the DST behaves as if we were inserting m independent strings and the following recursion holds:

$$P_{m+1}(u) = 1 + u^m \sum_{\mathbf{k}\in\mathbb{N}^{|\mathcal{A}|}} \binom{m}{\mathbf{k}} \prod_{a\in\mathcal{A}} p_a^{k_a} P_{k_a} \ . \quad (44)$$

where the binomial $\binom{m}{\mathbf{k}} = \frac{m!}{\prod_{a\in\mathcal{A}} k_a!}$ when $\sum_{a\in\mathcal{A}} k_a = m$ and 0, otherwise. Here we introduce the exponential generating function $L(z,u) = \sum_m \frac{z^m}{m!} L_m(u)$ to get the functional equation:

$$\frac{\partial}{\partial z} L(z,u) = \prod_{a\in\mathcal{A}} L(p_a u z, u) \ . \quad (45)$$

In [**?**] there is a generalization to random sequences built on a Markovian source. But here we will limit ourselves to memoriless sources. It is clear by construction that $L(z,1) = e^z$, since $L_m(1) = 1$ for all integer m. Via the cumulant formula we know that for each integer m we for t complex sufficient small, for which $\log(L_m(e^t)$ exists:

$$\log(P_n(e^t)) = t\mathbf{E}(L_n) + \frac{t^2}{2}\mathrm{Var}(L_n) + O(t^3) \ . \quad (46)$$

Notice that the term in $O(t^3)$ is not expected to be uniform in m.

We remark in passing that $\mathbf{E}(L_m) = L'_m(1)$ and $\mathrm{Var}(L_m) = L''_m(1) + L'_m(1) - (L'_m(1))^2$.

3 Moment analysis of path length in Digital Search Tree

In [2] we prove the following theorem:

Theorem 3. *When $m \to \infty$ we have the following estimate*

$$\mathbf{E}(L_m) = \ell(m) + O(1) \tag{47}$$
$$\mathrm{Var}(L_m) = v(m) \tag{48}$$

We don't change the proofs of [2].

4 Distribution analysis of path length in Digital Search Tree

4.1 Results headlines

Our aim is to prove that the limiting distribution of the path length is normal when $m \to \infty$. Namely the following theorem:

Theorem 4. *For all $\delta > 0$ and for all $\delta' < \delta$ there exists $\varepsilon > 0$ such that for all t such $|t| \leq \varepsilon$: $\log L_m(e^{tm^{-\delta}})$ exists and*

$$\log L_m(e^{tm^{-\delta}}) = O(m) \tag{49}$$
$$\log L_m(e^{tm^{-\delta}}) = \frac{t}{m^\delta}\mathbf{E}(L_m) + \frac{t^2}{2m^{2\delta}}\mathrm{Var}(L_m) + t^3 O(m^{1-3\delta'}) \ . \tag{50}$$

To this end we use depoissonization argument [4] on the generating function $L(m,u)$ with $u = \exp(\frac{t}{m^\delta})$. We delay the proof of the theorem in next section dedicated to dePoissonization. Therefore we can now prove our main result:

Theorem 5. *The random variable $\frac{L_m - \mathbf{E}(L_m)}{\sqrt{\mathrm{Var} L_m}}$ tends to a normal distribution of mean 0 and variance 1 in probability and in moment. More precisely, for all $\delta > 0$ we have the following. For any given real number x:*

$$\lim_{m\to\infty} P(L_m > \mathbf{E}(L_m) + x\sqrt{\mathrm{Var}(L_m)}) = \phi(x) \ . \tag{51}$$

and for all integer k:

$$\mathbf{E}\left((\frac{L_m - \mathbf{E}(L_m)}{\sqrt{\mathrm{Var} L_m}})^k \right) = \mu_k + O(m^{-\frac{1}{2}+\delta}) \ . \tag{52}$$

where $\mu_k = 0$ when k is odd and $\mu_k = \frac{k!}{2^{k/2}(\frac{k}{2})!}$.

Proof. The first point is a consequence of Levy Theorem. Indeed let t a complex number. Using theorem 4, with some $\delta > \frac{1}{2}$ and $\delta' > \delta$ such that $1 - 3\delta' < 0$. We have $\frac{t}{\sqrt{\mathrm{Var}(L_m)}}$ which will drop below $\varepsilon m^{-\delta}$ since $\frac{m^\delta}{\sqrt{\mathrm{Var}(L_m)}} \to 0$ when $m \to \infty$. therefore $\log L_m(e^{t/\sqrt{\mathrm{Var}(L_m)}})$ exists and tends to $\frac{t^2}{2}$. According to continuity Levy theorem this implies that the distribution density tends to the distribution density of $\exp(\frac{t^2}{2})$ which is the normal distribution. Notice that the application of Levy theorem does not provide insight in the convergence rate of density functions.

The second point comes from the fact that the kth moments of $\frac{L_m - \mathbf{E}(L_m)}{\sqrt{\mathrm{Var}(L_m)}}$ is equal to $k!$ the kth derivative of $L_m(\exp(\frac{t}{\sqrt{\mathrm{Var}(L_m)}}))e^{-t/\sqrt{\mathrm{Var}(L_m)}}$ at $t = 0$. Since for t in any complex compact set and for any small $\delta > 0$, the latter function has evaluation equal to $\exp(\frac{t^2}{2})(1 + O(m^{-\frac{1}{2}+\delta})$, then by application of Cauchy formula on a circle of radius R encircling the origin

$$\begin{aligned}
\mathbf{E}\left((\frac{L_m - \mathbf{E}(L_m)}{\sqrt{\mathrm{Var} L_m}})^k \right) &= \frac{1}{2i\pi} \oint \frac{dt}{t^{k+1}} L_m(\exp(\frac{t}{\sqrt{\mathrm{Var}(L_m)}}))e^{-t/\sqrt{\mathrm{Var}(L_m)}} \\
&= \frac{1}{2i\pi} \oint \frac{dt}{t^{k+1}} \exp(\frac{t^2}{2})(1 + O(m^{-\frac{1}{2}+\delta}) \\
&= \mu_k + O(R^{-k} \exp(\frac{R^2}{2})m^{-\frac{1}{2}+\delta}) \ .
\end{aligned}$$

□

We also have deviation results:

Theorem 6. • **Large deviation** *Let $\delta > \frac{1}{2}$, there exists $\varepsilon > 0$ there exists $B > 0$ and $\beta > 0$ such that for all numbers $x \geq 0$:*

$$P(|L_m - \mathbf{E}(L_m)| > xm^\delta) \leq B \exp(-\beta m^\varepsilon x) \tag{53}$$

• **Moderate deviation** *Let $\delta < \frac{1}{6}$ and $A > 0$, there exists $B > 0$ such that for all integer m and for all non-negative real number $x < An^\delta$:*

$$P(|L_m - \mathbf{E}(L_m)| \geq x\sqrt{\mathrm{Var}(L_m)}) \leq Be^{-\frac{x^2}{2}} \ . \tag{54}$$

Proof. **Large deviation:** This is an application of Chernov bounds. We take t as being a non negative real number. We have the identity:

$$P(L_m > E(L_m) + xm^\delta) = P(e^{tL_m} > e^{(E(L_m)+xm^\delta)t}) \ . \tag{55}$$

using Tchebychev inequality:

$$\begin{aligned}
P(e^{tL_m} > e^{(E(L_m)+xm^\delta)t}) &\leq \frac{\mathbf{E}(e^{tL_m})}{e^{(E(L_m)+xm^\delta)t})} &\quad (56) \\
&= L_m(e^t)\exp(-t\mathbf{E}(L_m) - xm^\delta t) \ . &\quad (57)
\end{aligned}$$

Here we take $\delta' = \frac{\delta+1/2}{2}$, and $\varepsilon = \delta' - \frac{1}{2}$. Let $t = \beta m^{-\delta'}$ such that the estimate $\log L_m(e^t) = tE(L_m) + O(t^2 m^{1+\varepsilon})$ is valid. Therefore we have

$$\log L_m(e^{\beta m^{-\delta'}}) - \frac{\beta}{m^{\delta'}} = O(m^{-\varepsilon}) \ , \tag{58}$$

which tends to zero. We achieve the proof with taking $B > 0$ such that

$$B \geq \lim \sup_{m\to\infty} \exp(\log L_m(e^{\beta m^{-\delta'}})) - \frac{\beta}{m^{\delta'}} \ , \tag{59}$$

and noticing that $tm^\delta x = \beta m^\varepsilon x$.

The other side is similar except that we consider:

$$
\begin{aligned}
P(L_m < E(L_m) - xm^\delta) &= P(e^{-tL_m} > e^{-(E(L_m)-xm^\delta)t}) && (60)\\
&\leq L_m(e^{-t})\exp(t\mathbf{E}(L_m) - xm^\delta t) \ , && (61)
\end{aligned}
$$

and the proof follows the same lines.

Moderate deviation: We use the same proof as in large deviation except that we take $t = \frac{x}{\sqrt{\mathrm{Var}(L_m)}}$. We notice that $t = O(m^{-\delta'})$ for some $\delta' > \frac{1}{3}$. Thanks to theorem 4 we have

$$
\log L_m(e^t) = \mathbf{E}(L_m)t + \frac{t^2}{2}\mathrm{Var}(L_m) + o(1) \tag{62}
$$

We now develop for the Chernov estimate $P(L_m > \mathbf{E}(L_m) + x\sqrt{\mathrm{Var}(L_m)})$, the other estimate on $P(L_m < \mathbf{E}(L_m) - x\sqrt{\mathrm{Var}(L_m)})$ goes similar way. Therefore

$$
\begin{aligned}
P(L_m < \mathbf{E}(L_m) + x\sqrt{\mathrm{Var}(L_m)}) &\leq \exp(\log L_m(e^t) - t\mathbf{E}(L_m) + xt\sqrt{\mathrm{Var}(L_m)})\\
&= \exp(\frac{t^2}{2}\mathrm{Var}(L_m) - xt\sqrt{\mathrm{Var}(L_m)} + o(1))\\
&= (1+o(1))\exp(-\frac{x^2}{2}) \ .
\end{aligned}
$$

□

4.2 Streamlined technical proofs

4.2.1 DePoissonization tool

We recall that $X(z) = \frac{\partial}{\partial u}L(z,1)$ and $V(z) = \frac{\partial^2}{\partial u^2}L(z,1) + \frac{\partial}{\partial u}L(z,1) - (\frac{\partial}{\partial u}L(z,1))^2$. Our aim is to prove the theorem 4. To this end we will use the diagonal exponential depoissonization tool. Let θ a non-negative number smaller tan $\frac{\pi}{2}$. We define $\mathcal{C}(\theta)$ as the complex cone around the positive real axis defined by $\mathcal{C} = \{z, |\arg(z)| \leq \theta\}$. We will use of the following theorem from [4] in theorem 8 about diagonal exponential depoissonization tool:

Theorem 7. *Let assume a sequence u_k of complex number, a number $\theta \in]0, \frac{\pi}{2}[$. For al $\varepsilon > 0$ there exist $c > 1$, $\alpha < 1$, $A > 0$ and $B > 0$ such that:*

$$
\begin{aligned}
z \in \mathcal{C}(\theta) \ , \quad |z| \in [\frac{n}{c}, cn] \quad &\Rightarrow \quad |\log(L(z,u_m)| \leq B|z| && (63)\\
z \notin \mathcal{C}(\theta) \ , \quad |z| = n \quad &\Rightarrow \quad |L(z,u_m)| \leq Ae^{\alpha n} && (64)
\end{aligned}
$$

Then

$$
L_m(u_m) = L(m,u_m)\exp(-m - \frac{m}{2}\left(\frac{\partial}{\partial z}\log(L(m,u_m)) - 1\right)^2)(1 + o(m^{-\frac{1}{2}+\varepsilon})) \ . \tag{65}
$$

We prove in the next section the following result:

Theorem 8. *Let $\delta \in]0,1[$. There exists a numbers $\theta \in]0, \frac{\pi}{2}[$, $\alpha < 1$, $A > 0$, $B > 0$ and $\varepsilon > 0$ such that for all complex t such $|t| \leq \varepsilon$:*

$$z \in \mathcal{C}(\theta) \quad \Rightarrow \quad |\log(L(z, e^{t|z|^{-\delta}}))| \leq B|z| \tag{66}$$

$$z \notin \mathcal{C}(\theta) \quad \Rightarrow \quad |L(z, e^{t|z|^{-\delta}})| \leq Ae^{\alpha|z|} \tag{67}$$

The proof is delayed to the next section. In order to achieve the proof of theorem 4 we will make use of the following lemmas:

Lemma 1. *Let δ be an arbitrary non negative number. There exist $\varepsilon > 0$ such that when $|t| \leq \varepsilon$ and $z \in \mathcal{C}(\theta)$ the following estimate holds*

$$X(z) = O(|z|^{1+\delta}) \tag{68}$$

$$V(z) = O(|z|^{1+2\delta}) \tag{69}$$

$$\log L(z, e^{t|z|^{-\delta}}) = z + X(z)\frac{t}{|z|^{\delta}} + V(z)\frac{t^2}{2|z|^{2\delta}} + O(t^3|z|^{1+3\delta}) \ . \tag{70}$$

Proof. We have $\log L(z,1) = z$. We notice that $X(z)$ and $V(z)$ are respectively the first and second derivative of $L(z, e^t)$ with respect to t at $t = 0$. Via Cauchy formula we have

$$X(z) = \frac{1}{2i\pi}\oint \log L(z, e^t)\frac{dt}{t^2} \tag{71}$$

$$V(z) = \frac{2}{2i\pi}\oint \log L(z, e^t)\frac{dt}{t^3} \tag{72}$$

where the integral is made on the circle of center 0 and radius $\varepsilon|z|^{-\delta}$. On this integral path the estimate $|\log L(z, e^t)| \leq B|z|$ and therefore we have

$$|X(z)| \leq \frac{B}{\varepsilon}|z|^{1+\delta} \tag{73}$$

$$|V(z)| \leq \frac{2B}{\varepsilon^2}|z|^{1+2\delta} \tag{74}$$

which proves the two first points. For the third point we need to estimate $R(z,t) = \log L(z, e^t) - z - X(z)t - V(z)\frac{t^2}{2}$ and we use the Cauchy formula [4]

$$R(z,t) = \frac{2t^3}{2i\pi}\oint \log L(z, e^{t'})\frac{dt'}{(t')^3(t'-t)} \tag{75}$$

the integral being done on the circle of center 0 and radius $\varepsilon|z|^{-\delta}$. If we restrict $|t| \leq \frac{\varepsilon}{2}|z|^{-\delta}$ to have $|t - t'| \geq \frac{\varepsilon}{2}|z|^{-\delta}$, then we get the estimate $|R(z,t)| \leq \frac{2B}{\varepsilon^3}|t|^3|z|^{1+3\delta}$. □

Let $D(z,t) = \frac{\partial}{\partial z}\log L(z, e^t)$. We have the second technical lemma:

Lemma 2. *Let $\delta > 0$. There exists $\varepsilon > 0$ and $B' > 0$ such that for all t such $|t| < \varepsilon$ have the estimate $|D(m, tm^{-\delta)})| \leq B'$, and the following estimate holds*

$$D(m, tm^{-\delta}) = 1 + X'(m)\frac{t}{m^{\delta}} + O(t^2m^{2\delta}) \tag{76}$$

and $X'(m) = O(m^{\delta})$.

Proof. The main point here is the bound $D(m, tm^{-\delta})) = O(1)$. In order to get it, we again use the Cauchy formula:

$$D(m,t) = \frac{1}{2i\pi} \oint \log L(z, e^t) \frac{dz}{(z-m)^2} \tag{77}$$

where the integration loop encircles m and is included in the cone $\mathcal{C}(\theta)$ and $|t| \leq \varepsilon m^{-\delta}$ so that the inequality $|\log L(z, e^t)| \leq B$ holds. To this end one can take the circle of center 0 and radius $m \sin(\theta)$, and therefore

$$|D(m,t)| \leq B \frac{m}{\sin^2(\theta)} \ . \tag{78}$$

from there we can take a similar path as in the previous lemma proof, noticing that $D(m, 0) = 1$:

$$X'(m) = \frac{\partial}{\partial t} D(m,0) \quad = \quad \frac{1}{2i\pi} \oint D(m,t') \frac{dt'}{(t')^2} \tag{79}$$

$$D(m,t) - 1 - X'(m)t \quad = \quad \frac{t}{2i\pi} \oint D(m,t') \frac{dt'}{(t')^2(t'-t)} \tag{80}$$

the integral loop being the circle of center 0 and radius $\varepsilon m^{-\delta}$ □

Theorem 9. *For all $\delta > \frac{1}{2}$, and let $A > 0$ be an arbitrary number. For all arbitrarily small $\delta' < \delta$, and for all t complex such that $|t| \leq A$ we have*

$$\begin{aligned} \log L_m(e^{tm^{-\delta}}) \quad &= \quad X(m)\frac{t}{m^\delta} + \left(V(m) - m(X'(m))^2\right)\frac{t^2}{2m^{2\delta}} \\ &\quad + O(t^2 m^{1-2\delta+2\delta'}) + O(t^3 m^{1-3\delta+3\delta'}) \ . \end{aligned} \tag{81}$$

Proof. We take $\delta > \frac{1}{2}$ and $\delta' < \delta$ arbitrary small, Applying lemma 1 we get that

$$L(m, e^{tm^{-\delta}}) = m + X(m)\frac{t}{m^\delta} + V(m)\frac{t^2}{2m^{2\delta}} + O(t^3 m^{1-3\delta+3\delta'}) \ . \tag{82}$$

Now applying lemma 2 we got

$$D(m, e^{tm^{-\delta}}) = 1 + X'(m)\frac{t}{m^\delta} + O(t^2 m^{1-2\delta+2\delta'}) \ . \tag{83}$$

Therefore, using $X'(m) = O(m^{\delta'})$:

$$\left(D(m, e^{tm^{-\delta}}) - 1\right)^2 = (X'(m))^2 \frac{t^2}{m^{2\delta}} + O(t^3 m^{1-3\delta+3\delta'}) \ . \tag{84}$$

Putting everything together, with result of theorem 7, gives the expected estimate on $\log L_m(e^{tm^{-\delta}})$. □

Corollary 3. *By simple identification $\delta = \frac{1}{2}$ in $L_m(e^{t/\sqrt{m}})$ expansion in last theorem, we recover the following estimate valid for any $\delta' > 0$*

$$\mathbf{E}(L_m) \quad = \quad X(m) + o(1) = O(m^{1+\delta'}) \tag{85}$$

$$\mathrm{Var}(L_m) \quad = \quad V(m) - m(X'(m))^2 + o(1) = O(m^{1+2\delta'}) \ . \tag{86}$$

Proof of theorem 4. From theorem 7 and lemma 2, for any $\delta' > 0$, we basically get the estimate $\log L_m(e^{tm^{-\delta'}}) = O(m)$. Denoting $R_m(t) = \log L_m(e^t) - \mathbf{E}(L_m)t - \mathrm{Var}(L_m)\frac{t^2}{2}$ we have via Cauchy

$$R_m(t) = \frac{2t^3}{2i\pi}\oint \log L_m(e^{t'})\frac{dt'}{(t')^3(t'-t)} \ . \tag{87}$$

Assuming $|t| < \varepsilon m^{-\delta'}$ we get $R_m(t) = t^3 O(\frac{m^{1+3\delta'}}{\varepsilon^3})$ and therefore for $\delta > \delta'$ we have

$$R_m(e^{tm^{-\delta}}) = t^3 o(m^{1-3\delta+3\delta'}) \ . \tag{88}$$

□

4.2.2 Growth analysis of the exponential generating function

The key of our analysis is the following theorem

Theorem 10. *There exists a complext neighborhood $\mathcal{U}(0)$ of 0 and $B > 0$ such that for all $t \in \mathcal{U}(0)$ and for all $z \in \mathcal{C}$: $\log(L(z, u(z,t))$ exists and*

$$\log(L(z, u(z,t)) \leq B|z| \ . \tag{89}$$

To this end we introduce the following function $f(z,u)$ that we call the *kernel* function, defined by

$$f(z,u) = \frac{L(z,u)}{\frac{\partial}{\partial z}L(z,u)} = \frac{L(z,u)}{\prod_{a\in\mathcal{A}} L(p_a uz, u)} \tag{90}$$

Notice that formally $\frac{1}{f(z,u)} = \frac{\partial}{\partial z}\log(L(z,u)$. Indeed if we show that the kernel function is well defined and is never zero in a convex set that contains the real positive line then we will prove that $\log(L(z,u))$ exists since

$$\log(L(z,u)) = \int_0^z \frac{dx}{f(x,u)} \tag{91}$$

If we can prove that the estimate $f(x,u) = \Omega(1)$ then we get:

$$\log(L(z,u)) = \int_0^z \frac{dx}{f(x,u)} = O(z) \tag{92}$$

The estimate on the kernel function is therefore the key of our analysis. The Kernel function satisfies the following differential equation:

$$\frac{\partial}{\partial z}f(z,u) = 1 - f(z,u)\sum_{a\in\mathcal{A}} \frac{p_a u}{f(p_a uz, u)} \ , \tag{93}$$

and the results can be generalized as well.

We first start with a trivial lemma, whose proof is left to the reader.

Lemma 3. *Let for (x,ε) real positive tuple a function $h(x,\varepsilon)$ which is defined on an open set which contains all tuples $(x,0)$ with $x \geq 0$. Assume that function $h(x,\varepsilon)$ is real positive and continuously differentiable. If $\forall x \geq 0$: $\frac{\partial}{\partial x}h(x,0) < 1$, then for all compact set $\mathcal{K}_x$ there exist a compact neighborhood of $\mathcal{U}(0)$ of 0 $(x_0,t) \in \mathcal{K}_x \times \mathcal{U}(0)$, the sequence defined for k integer*

$$x_{k+1} = h(x_k, \varepsilon) \tag{94}$$

converges to a bounded fixed point when $k \to \infty$.

We denote $\frac{1}{f(z,u)} = 1 + a(z,u)$. We now prove the following more difficult lemma for t real:

Lemma 4. *Let δ' be a real number such that $\delta < \delta' < 1$. For all number $a > 0$ there exists a real number $\varepsilon > 0$ suc f(h that for all real t such that $|t| < \varepsilon$ and for all $z \in \mathcal{C}(\theta)$:*

$$|a(z, u(z,t))| \leq a\frac{|t|}{|z|^{\delta'}} \ . \tag{95}$$

Proof. We choose a real compact neighborhood $\mathcal{U}(1)$ of 1 and we let

$$\rho = \min_{u \in \mathcal{U}(1), a \in \mathcal{A}} \{\frac{1}{p_a|u|}\} \ . \tag{96}$$

We assume that $\mathcal{U}(1)$ is small enough that ρ is greater than 1. Let R be a real number. We denote $\mathcal{C}_0(\theta)$ the subset of $\mathcal{C}(\theta)$ made of complex number of modulus smaller than or equal to R. By extension let k be an integer, and we denote $\mathcal{C}_k(\theta)$ the subset of $\mathcal{C}(\theta)$ made of complex numbers of modulus smaller than or equal to $R\rho^k$. By construction if $z \in \mathcal{C}_k(\theta)$ for $k > 0$ and $u \in \mathcal{U}(1)$ then all $p_a uz$ for $a \in \mathcal{A}$ belong to $\mathcal{C}_{k-1}(\theta)$.

It is difficult to keep $u(z,t)$ as really dependent on z without unneeded complication in the differential equation of $f(u, u(z,t))$. Instead we will make $u(z,t)$ independent of z in each portion $\mathcal{C}_k(\theta)$. We introduce $u_k(t) = e^{t\nu^k}$ for $\nu = \rho^{-\delta}$, and we fix $\mu = \rho^{-\delta'} > \nu$.

In the following we will denote $f_k(z) = f(z, u_k(t))$, and $u_k = u_k(t)$, omitting variable t when no confusion is possible. The kernel function satisfies the differential equation:

$$f'_k(z) = 1 - f_k(z) \sum_{a\in\mathcal{A}} \frac{p_a u_k}{f(p_a u_k z, u_k)} \ . \tag{97}$$

Let $a_k(z,t) = a(z, u_k(t))$. Since $L(z,1) = e^z$ for all z and therefore $f(z,1) = 1$. Since $\frac{\partial}{\partial u} f(z,u)$ is well defined and continuous, we can restrict the neighborhood $\mathcal{U}(1)$ such that $f(z,u)$ is non zero and therefore $a(z,u)$ is well defined for $z \in \mathcal{C}_0(\theta)$ and $u \in \mathcal{U}(1)$. Let a_0 be a non negative numbers such that

$$\forall u \in \mathcal{U}(1) \ , \quad \forall z \in \mathcal{C}_0(\theta) \ : \quad |a_0(z,t)| \leq a_0|t| \tag{98}$$

Now we fix ε such that $a_0\varepsilon < 1$. We fix our aim is to prove that there exists a number $\epsilon > 0$ such that there exists an increasing sequence of non negative numbers a_k such that for all $z \in \mathcal{C}_k$: and for all t such that $|t| \leq \epsilon$:

$$|a_k(z,t)| \leq a_k|t|\mu^k \tag{99}$$

and $\limsup_{k\to\infty} a_k < \infty$.

For this purpose we will fix a recursion property that state that the previous property is true up to some integer $k-1$. We will prove that the property is also true for k. But to conclude the proof for all k, we have to be careful that we must take the same ϵ for all k.

Let $z \in \mathcal{C}_k(\theta)$. We denote

$$g_k(z) = \sum_{a\in\mathcal{A}} \frac{p_a u_k}{f(p_a u_k z, u_k)} \ . \tag{100}$$

Thus equation 97 rewrites $f'_k(z) = 1 - g_k(z)f_k(z)$ and the differential equation can be solved by

$$f_k(z) = 1 + \int_0^z (1 - g_k(x)) \exp(G_k(x) - G_k(z))dx \ , \tag{101}$$

where function $G_k(z)$ is a primitive of function $g_k(z)$.

We now will give some bounds on $g_k(z)$ when $z \in \mathcal{C}_k(\theta)$ and $|t| < \varepsilon$. For all $a \in \mathcal{A}$ $p_a u_k z \in \mathcal{C}_{k-1}(\theta)$. We have $u_k(t) = u_{k-1}(\nu t)$ and we can use the recursion since $|\nu t| < \varepsilon$. In particular we have

$$\begin{aligned} g_k(z) &= \sum_{a\in\mathcal{A}} p_a u_k (1 + a_{k-1}(p_a u_k z, \nu t)) && (102)\\ &= 1 + b_k(z,t) && (103)\end{aligned}$$

with

$$b_k(z,t) = \sum_{a\in\mathcal{A}} p_a(u_k - 1 + u_k a_{k-1}(p_a u_k z, \nu t) \ , \tag{104}$$

Since both $|a_{k-1}(pu_k, \nu t)|$ and $|a_{k-1}(qu_k, \nu t)|$are smaller than $a_{k-1}\nu\mu^{k-1}|t|$ and since $|u_k - 1| \le \beta\nu^k|t|$ for some β close to 1, we have $|b(z,t)| \le b_k|t|$ with

$$b_k = (a_{k-1}\nu\mu^{k-1} + \beta\nu^k)(1 + \beta\nu^k\varepsilon) \ . \tag{105}$$

Thus pluging in equation (101) we got

$$\begin{aligned} |f_k(z) - 1| &\le \int_0^z |b_k(x,t)| \exp(\Re(G_k(x) - G_k(z)))dx && (106)\\ &\le \int_0^1 b_k|t||z| \exp(\Re(G_k(zy) - G_k(z)))dy && (107)\\ &\le \frac{b_k|t|}{\cos(\theta) - b_k|t|} && (108)\end{aligned}$$

Since we consider

$$\begin{aligned} \Re(G_k(yz) - G_k(z)) &= -\Re(z)(1-y) + \int_y^1 \Re(zb_k(zx,t))dx \\ &\le -\cos(\theta)|z| + b_k|z| \end{aligned}$$

And therefore

$$\left|\frac{1}{f_k(z)} - 1\right| \le \frac{\frac{b_k|t|}{\cos(\theta)-b_k|t|}}{1 - \frac{b_k|t|}{\cos(\theta)-b_k|t|}} = \frac{b_k|t|}{\cos(\theta) - 2b_k|t|} \ . \tag{109}$$

Therefore we get an evaluation for a_k which is

$$a_k \le (a_{k-1}\frac{\nu}{\mu} + \beta\frac{\nu^k}{\mu^k})(1 + \beta\nu^k\varepsilon)\frac{1}{\cos(\theta) - b_k\varepsilon} \ . \tag{110}$$

Clearly if θ is chosen such $\frac{\nu}{\mu\cos(\theta)} < 1$, we can identify $a_k \le h(a_{k-1}, \varepsilon)$ with $h(x, \varepsilon)$ as a function described in lemma 3. Moreover $h(x, \varepsilon)$ is increasing. With lemma 3 we can make ε small enough such that $\limsup_{k\to\infty} a_k < \infty$. □

We extend this lemma to a complex neighborhood of 0:

Lemma 5. *For all number $a > 0$ there exists $\varepsilon > 0$, $\theta \in]0, \frac{\pi}{2}[$ such for complex t such that $|t| < \varepsilon$ and for all $z \in \mathcal{C}$:*

$$|a(z, u(z,t))| \leq a\frac{|t|}{|z|^{\delta'}} \ . \tag{111}$$

Proof. The proof is essentially the same as the previous lemma excepted that we have to extend the cone $\mathcal{C}$ and a larger set $\mathcal{C}'(\theta)$ defined by $\{z, |\arg(z)| \leq \theta + \phi|z|^{\delta-1}\}$ so that if $z \in \mathcal{C}'(\theta)$ then for all $a \in \mathcal{A}$ $p_a u(z,t)z$ belongs to $\mathcal{C}'(\theta)$ (taking into account the small rotation of angle $\Im(\frac{t}{|z|^\delta})$ that would have made the two points getting out $\mathcal{C}(\theta)$ and missed the recursion hypothesis. □

We now check the growth of function $L(z, a, u(z,t))$ when z is outside the cone $\mathcal{C}(\theta)$.

Theorem 11. *Let $\theta \in]0, \frac{\pi}{2}[$, there exist numbers $A > 0$, $\alpha < 1$ and $\varepsilon > 0$ such that for all complex t such $|t| \leq \varepsilon$:*

$$z \notin \mathcal{C}(\theta) \Rightarrow |L(z, u(z,t))| \leq Ae^{\alpha|z|} \ . \tag{112}$$

Proof. We proceed as with previous proofs. We first prove for t real. We take a neighborhood $\mathcal{U}(1)$ and we define ρ as in the same condition as in the proof of lemma 4. We define $\bar{\mathcal{C}}(\theta)$ as the complementary of $\mathcal{C}(\theta)$ in the complex plan. We also denote

$$\lambda = \min_{u\in\mathcal{U}(1), a\in\mathcal{A}} \{p_a|u|\} \tag{113}$$

We take $R > 0$ and we define $\bar{\mathcal{C}}_0(\theta)$ and $\bar{\mathcal{C}}_k(\theta)$ for $k > 0$ integer as subsets of $\bar{\mathcal{C}}(\theta)$:

$$\begin{aligned} \bar{\mathcal{C}}_0(\theta) &= \{z \in \bar{\mathcal{C}}(\theta), |z| \leq \lambda R\} \\ \bar{\mathcal{C}}_k(\theta) &= \{z \in \bar{\mathcal{C}}(\theta), \lambda R < |z| \leq \rho^k R\} \end{aligned}$$

With this definition, if $u \in \mathcal{U}(1)$ when z is in $\bar{\mathcal{C}}_k(\theta) - \bar{\mathcal{C}}_{k-1}(\theta)$ then both puz and quz are in $\bar{\mathcal{C}}_{k-1}(\theta)$.

Since $L(z, 1) = e^z$ and that if $\alpha > \cos(\theta)$ then $|L(z, 1)| \leq e^{\alpha|z|}$. There exist $A_0 > 0$ and ε such that for all t such $|t| \leq \varepsilon$ and for all $z \in \bar{\mathcal{C}}_0(\theta)$: $|L(z, e^t)| \leq A_0 e^{\alpha|z|}$. We also tune ε so that $\alpha \prod_k u_k(\varepsilon) < 1$.

We can do the same analysis with $z \in \bar{\mathcal{C}}_1(\theta)$, but since $|L(z, 1)|$ is strictly smaller than $e^{\alpha|z|}$ for all $z \in \bar{\mathcal{C}}_1(\theta)$, we can find $A_1 < 1$ and $\varepsilon > 0$ such that for all t such $|t| \leq \varepsilon$ and for all $z \in \bar{\mathcal{C}}_1(\theta)$: $|L(z, e^t)| \leq A_1 e^{\alpha|z|}$. In fact since $\min_{z\in\bar{\mathcal{C}}_1(\theta)}\{\frac{|e^z|}{e^{\alpha|z|}}\} \to 0$ when $R \to \infty$ we can make A_1 as small as we want.

We define $\alpha_k = \alpha \prod_{i=0}^{i=k} u_k(\varepsilon)$. We will prove by induction that there exists an increasing sequence $A_k < 1$ such that for t such $|t| \leq \varepsilon$:

$$z \in \bar{\mathcal{C}}_k(\theta) \Rightarrow |L(z, u_k(t))| \leq A_k e^{\alpha_k|z|} \ . \tag{114}$$

Our plan is to prove this property by induction. Assume is true until some $k-1$ and we will prove it is true for k. Assume $z \in \bar{\mathcal{C}}_k(\theta) - \bar{\mathcal{C}}_{k-1}(\theta)$. We can make use of the differential equation:

$$L(z, u_k) = L(\frac{z}{\rho}, u_k) + \int_{z/\rho}^{z} \prod_{a\in\mathcal{A}} L(p_a u_k x, u_k)dx \ . \tag{115}$$

Via obvious inequality manipulation we get

$$|L(z,u_k)| \leq |L(\frac{z}{\rho},u_k)| + |z|\int_{1/\rho}^{1}\prod_{a\in\mathcal{A}}|L(p_a u_k zy,u_k)|dy \ . \tag{116}$$

Using the hypothesis of recursion: $|L(\frac{z}{\rho},u_k)| \leq A_{k-1}e^{\alpha_{k-1}|z|/\rho}$, for all $a \in \mathcal{A}$: $|L(p_a u_k yz,u_k)| \leq A_{k-1}e^{\alpha_{k-1}p_a|u_k||z|y} \leq A_{k-1}e^{\alpha_k p|z|y}$ (we have $\alpha_{k-1}|u| \leq \alpha_{k-1}e^{\varepsilon} = \alpha_k$):

$$|L(z,u_k) \leq A_{k-1}e^{\alpha_{k-1}|z|/\rho} + \frac{A_{k-1}^2}{\alpha_k}\left(e^{\alpha_k|z|} - e^{\alpha_k|z|/\rho}\right) \ . \tag{117}$$

This gives an estimate

$$A_k \leq \frac{A_{k-1}^2}{\alpha_k} + A_{k-1}e^{-\rho^{k-2}(\rho-1)\alpha_k R} \ . \tag{118}$$

Clearly the term in $e^{-\rho^{k-2}(\rho-1)\alpha_k R}$ can be made as small as we want by increasing R. If we choose A_1 such that $\frac{A_1}{\alpha_1} + e^{-\rho^{k-2}(\rho-1)\alpha_k R} < 1$ for all k then we get $A_k \leq A_{k-1}$ and the theorem is proven for t real.

Second, we need to expand our proof to the case where t is complex and $|t| \leq \varepsilon$. To this end we use a similar trick as in the proof of lemma 5. We expand $\bar{\mathcal{C}}(\theta)$ to $\bar{\mathcal{C}}'(\theta)$ as the set $\{z, |\arg(z)| \geq \theta + \phi R^{\delta-1} - \phi|z|^{\delta-1}\}$ when $|z| > R\rho$, in order to keep $p_a u_k z$ in $\bar{\mathcal{C}}'(\theta)$ for all $a \in \mathcal{A}$ when $z \in \bar{\mathcal{C}}_k'(\theta) - \bar{\mathcal{C}}_{k-1}'(\theta)$ in absorbing the tiny rotation that the factor u_k implies when t is complex. Of course one must choose ϕ such that $\theta + \phi R^{\delta-1} < \frac{\pi}{2}$ and to tune ε small in consequence. □

References

[1] J Ziv, A Lempel "Compression of individual sequences via variable-rate coding" IEEE Transactions on Information Theory, 1978

[2] P Jacquet, W Szpankowski "Asymptotic behavior of the Lempel-Ziv parsing scheme and digital search trees", Theoretical Computer Science, 1995 - Elsevier

[3] G Louchard, W Szpankowski "On the average redundancy rate of the Lempel-Ziv code" IEEE Transactions on Information Theory, 1997.

[4] P Jacquet, W Szpankowski "Analytical depoissonization and its applications", Theoretical Computer Science, 1998 - Elsevier

[5] P Jacquet, W Szpankowski, J Tang "Average profile of the Lempel-Ziv parsing scheme for a Markovian source", Algorithmica, 2001 - Springer

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

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY 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 - 38334 Montbonnot Saint-Ismier (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