Limit Behavior of Maxima in Geometric Words representing Set Partitions

Michael Fuchs

Department of Applied Mathematics

National Chiao Tung University

Hsinchu, 300

Taiwan

Mehri Javanian

Department of Statistics

University of Zanjan

Zanjan

Iran

June 24, 2015

Dedicated to Helmut Prodinger on the occasion of his 60th birthday.

Abstract

We consider geometric words $\omega_1 \cdots \omega_n$ with letters satisfying the restricted growth property

$$\omega_k \leq d + \max\{\omega_0, \ldots, \omega_{k-1}\},$$

where $\omega_0 := 0$ and $d \geq 1$. For $d = 1$ these words are in 1-to-1 correspondence with set partitions and for this case, we show that the number of left-to-right maxima (suitable centered) does not converge to a fixed limit law as n tends to infinity. This becomes wrong for $d \geq 2$, for which we prove that convergence does occur and the limit law is normal. Moreover, we also consider related quantities such as the value of the maximal letter and the number of maximal letters and show again non-convergence to a fixed limit law.

1   Introduction and Results

There exists a vast literature on statistical properties of *geometric words* $\omega_1 \cdots \omega_n$, which are words whose letters are generated by independent, geometrically distributed random variables, i.e.,

$$P(\omega_k = \ell) = pq^{\ell-1}, \qquad (\ell \geq 1),$$

where $0 < p < 1$ is the success probability and, for convenience of notation, we set $q := 1 - p$. For instance, some of the various statistics which have been studied for such words are:

- *left-to-right maxima* (e.g., see Archibald and Knopfmacher [1, 2]; Bai, Hwang and Liang [4]; Brennan, Knopfmacher, Mansour and Wagner [8]; Knopfmacher and Prodinger [19]; Oliver and Prodinger [30]; Prodinger [31, 32, 34, 35, 36]);

- *maximum value* (e.g., see Bruss and O'Cinneide [9]; Eisenberg [10]; Prodinger [36]);

- *number of times the maximum occurs* (e.g., see Kirschenhofer and Prodinger [18]);

- *number of different letters, missing letters and gaps* (e.g., see Archibald and Knopfmacher [3]; Goh and Hitczenko [14]; Louchard and Prodinger [26]; Louchard, Prodinger and Ward [27]),

- *inversions* (e.g., see Prodinger [33]);

- *ascends* and *descends* (e.g., see Brennan [5]; Brennan and Knopfmacher [6, 7]; Knopfmacher and Prodinger [20, 21, 22]; Louchard and Prodinger [25]);

- *runs* (e.g., see Eryilmaz [11]; Grabner, Knopfmacher and Prodinger [15]; Lee and Tsai [23]; Louchard and Prodinger [24]).

Here, we consider geometric words which satisfy the following (generalized) *restricted growth property*

$$\omega_k \leq d + \max\{\omega_0, \ldots, \omega_{k-1}\}, \qquad (1 \leq k \leq n) \tag{1}$$

with $\omega_0 := 0$ and $d \geq 1$.

For $d = 1$ such words are in a bijective correspondence with set partitions: order the blocks of a set partition according to ascending values of the smallest elements from each block; then, define ω_i to be the block which contains the i-th element. It is easy to see that every such word satisfies (1) with $d = 1$, and conversely, every word satisfying (1) with $d = 1$ corresponds to a set partition. This and the fact that they are related to *approximate counting* (see Prodinger [37]) sparked the recent interest in stochastic properties of geometric words satisfying (1) with $d = 1$.

We recall some of the results. The first quantity which was studied was the probability p_n that a geometric word satisfies (1) with $d = 1$, for which exact formulas were obtained by Mansour and Shattuk [28] and Oliver and Prodinger [29]. In addition, the authors of [29] also obtained the following asymptotic result

$$p_n \sim \frac{qQ(p)}{\log\left(1/q\right)Q(1)} n^{\log_{1/q} p} \sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}, \tag{2}$$

where $\chi_k = 2k\pi i/\log(1/q)$ and

$$Q(s) = \prod_{\ell \geq 1} \left(1 - q^\ell s\right).$$

Moreover, in [37], Prodinger considered the number of left-to-right maxima $L_n^{(1)}$ (or equivalently the value of the maximal letter) of a geometric word of length n given (1) with $d = 1$ and derived an asymptotic expansion of the mean

$$\mathbb{E}\left(L_n^{(1)}\right) \sim \log_{1/q} n + \Phi_1^{(1)}\left(\log_{1/q} n\right),$$

where $\Phi_1^{(1)}(z)$ is a one-periodic function. Our first result generalizes this to all higher moments. In particular, this will imply non-convergence to a fixed limit distribution.

Theorem 1. *We have, for all $m \geq 1$,*

$$\mathbb{E}\left(L_n^{(1)} - \log_{1/q} n\right)^m \sim \Phi_m^{(1)}\left(\log_{1/q} n\right),$$

where $\Phi_m^{(1)}(z)$ are one-periodic functions given in (9) below. As a consequence, $L_n^{(1)} - \log_{1/q} n$ does not converge to a fixed limit law.

The interest in this result lies in the fact that for geometric words without (1) it is known that the number of left-to-right maxima (suitable centered and normalized), in fact, does converge to a limit law which is normal; see for example [4]. Thus, if we denote the number of left-to-right maxima of a geometric word of length n given (1) with general $d \geq 1$ by $L_n^{(d)}$, a natural question is whether there is a phase change from non-convergence to a fixed limit law to convergence to a normal limit law as d grows to infinity and if yes, where does the phase change occur? Both questions are answered by our next result for whose formulation we need the following polynomial

$$P(z) = 1 - p \sum_{\ell=1}^{d} q^{\ell-1} z^{\ell}.$$

Moreover, we use ρ to denote its (unique) positive real root.

Theorem 2. *The sequence of random variables $L_n^{(d)}$, suitable centered and normalized, satisfies a central limit theorem:*

$$\frac{L_n^{(d)} + \log_{1/q} n/(\rho P'(\rho))}{\sqrt{\log_{1/q} n}} \xrightarrow{d} N(0, \sigma_d^2),$$

where

$$\sigma_d^2 := \frac{1}{\rho P'(\rho)} + \frac{P''(\rho)}{\rho P'(\rho)^3} + \frac{1}{\rho^2 P'(\rho)^2}.$$

Moreover, $\sigma_d^2 > 0$ if and only if $d \geq 2$.

Note that the above result also holds for $d = 1$, but does not give a meaningful result in this case.

As mentioned above, $L_n^{(1)}$ can also be interpreted as the value of the maximal letter (or also the number of blocks of the corresponding set partition). This, however, becomes wrong for $d \geq 2$, where the value of the maximal letter and the number of left-to-right maxima is different. Thus, we investigate next the value of the maximal letter of a geometric word of length n given (1) with $d \geq 2$ which we denote by $M_n^{(d)}$. Here, we do not have a phase change from non-convergence to convergence to a fixed limit law and, in fact, Theorem 1 can be generalized to all $d \geq 1$ (for the mean this was already proved by Fuchs and Prodinger in [13]).

Theorem 3. *We have, for all $m \geq 1$,*

$$\mathbb{E}\left(M_n^{(d)} - \log_{1/q} n\right)^m \sim \Phi_m^{(d)}\left(\log_{1/q} n\right),$$

where $\Phi_m^{(d)}(z)$ are one-periodic functions given in (11) below. As a consequence, $M_n^{(d)} - \log_{1/q} n$ does not converge to a fixed limit law.

Finally, we will consider the number N_n of times the maximal letter occurs in a geometric word of length n, where for the sake of the simplicity, we only consider the case $d = 1$ (this corresponds to the size of the last block in the corresponding set partition, where the blocks are ordered as above). Again, all the moments exhibit periodic fluctuations preventing convergence to a fixed limit law.

Theorem 4. *We have, for all $m \geq 1$,*

$$\mathbb{E} N_n^m \sim \Psi_m\left(\log_{1/q} n\right),$$

where $\Psi_m(z)$ are one-periodic functions given in (13) below. As a consequence, N_n does not converge to a fixed limit law.

We conclude the introduction with a short sketch of the paper. In the next section, we consider left-to-right maxima and prove Theorem 1 and Theorem 2. The method of proof will be a refinement of the method from [13], which will be recalled in the next section as well. The same approach can be also used to establish Theorem 3, whose proof will be presented in Section 3. Finally, the proof of Theorem 4 is similar, too, and will be briefly sketched in Section 4.

2   Left-to-right Maxima

In this section, we are going to consider the number of left-to-right maxima and prove Theorem 1 and Theorem 2. For the proof, we will refine the method from [13], which relied on the Mellin transform and the theory of analytic depoissonization. The former is a classical tool in analytic combinatorics and the reader is referred to the superb survey by Flajolet, Gourdon and Dumas [12] for background. For the latter, see the survey of Jacquet and Szpankowski [17] (and also Hwang, Fuchs and Zacharovas [16] from which the language used below is borrowed).

First, we use $p_{n,k}$ to denote the probability that a geometric word of length n satisfying (1) has exactly k left-to-right maxima. Moreover, we set $p_n = \sum_{k \geq 0} p_{n,k}$, which is the probability that a geometric word of length n satisfies (1). Note that for $d = 1$, the asymptotics of this probability was given in the introduction. By definition of $L_n^{(d)}$, we have

$$P\left(L_n^{(d)} = k\right) = \frac{p_{n,k}}{p_n}.$$

The crucial observation (already made in [13]) is that $p_{n,k}$ satisfies the following recurrence

$$p_{n+1,k} = \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} p_{j,k-1}, \qquad (n \geq 0, k \geq 1) \tag{3}$$

with initial conditions $p_{n,0} = [\![n = 0]\!]$ and $p_{0,k} = [\![k = 0]\!]$, where $[\![\cdot]\!]$ is the Iverson bracket. This recurrence is easily explained: we first condition on the first letter ℓ, which by definition satisfies $1 \leq \ell \leq d$ (the probability for this is $pq^{\ell-1}$, which is the factor after the first sum); next, we condition on the event that the remaining n letters contain exactly $n - j$ letters which are $\leq \ell$. There are $\binom{n}{j}$ choices of these letters and the probability that they are all $\leq \ell$ is $(1 - q^\ell)^{n-j}$. Moreover, the final j letters are all larger than ℓ (the probability for this is $q^{\ell j}$) and they form again a geometric word satisfying (1) and having one less left-to-right maxima (so, their probability is given by $p_{j,k-1}$).

Our goal is to use this recurrence in order to find an asymptotic expansion of $\sum_{k \geq 0} p_{n,k} e^{kt}$. To this end, we use poissonization

$$\tilde{L}(z,t) := e^{-z} \sum_{n \geq 0} \sum_{k \geq 0} p_{n,k} e^{kt} \frac{z^n}{n!}$$

which means that we replace n by a Poisson random variable of parameter z. Due to concentration of the Poisson distribution, we expect

$$\tilde{L}(n,t) \sim \sum_{k \geq 0} p_{n,k} e^{kt} \tag{4}$$

which is called the *Poisson heuristic* and will be justified below with the theory of analytic depoissonization. The advantage of poissonization is that (3) becomes

$$\tilde{L}(z,t) + \frac{\partial}{\partial z} \tilde{L}(z,t) = pe^t \sum_{\ell=1}^{d} q^{\ell-1} \tilde{L}(q^\ell z, t)$$

4

and this differential-functional equation can be asymptotically solved.

For this, we apply the Mellin transform to the differential-functional equation, where the Mellin transform of a function $\tilde{f}(x)$ is defined as

$$\mathscr{M}[\tilde{f}(x);\omega] = \int_0^\infty \tilde{f}(x)x^{\omega-1}\mathrm{d}x.$$

By properties of the Mellin transform, the equation becomes

$$\mathscr{M}[\tilde{L}(z,t);\omega] - (\omega-1)\mathscr{M}[\tilde{L}(z,t);\omega-1] = (1 - P_t(q^{-\omega}))\mathscr{M}[\tilde{L}(z,t);\omega],$$

where

$$P_t(z) = 1 - pe^t \sum_{\ell=1}^d q^{\ell-1}z^\ell.$$

(Note that $P_0(z)$ equals $P(z)$ from the introduction.) Next, it is advantageous to consider the normalization $\bar{\mathscr{M}}[\tilde{L}(z,t);\omega] = \mathscr{M}[\tilde{L}(z,t);\omega]/\Gamma(\omega)$. This yields

$$\bar{\mathscr{M}}[\tilde{L}(z,t);\omega] = \frac{\bar{\mathscr{M}}[\tilde{L}(z,t);\omega-1]}{P_t(q^{-\omega})}.$$

This recurrence has the general solution

$$\bar{\mathscr{M}}[\tilde{L}(z,t);\omega] = \frac{c(t)}{P_t(q^{-\omega})\Omega_t(q^{-\omega}e^t)},$$

where

$$\Omega_t = \prod_{\ell\geq1} P_t(q^\ell s)$$

and $c(t)$ is a suitable function, which can be obtained by the observation that $\tilde{L}(0,t) = 1$ and applying the direct mapping theorem from [12] (Theorem 3 on page 16) yielding

$$\lim_{\omega\to0}\bar{\mathscr{M}}[\tilde{L}(z,t);\omega] = \frac{\mathscr{M}[\tilde{L}(z,t);\omega]}{\Gamma(\omega)} = \frac{1/\omega + \cdots}{1/\omega + \cdots} = 1.$$

Using this gives that

$$c(t) = P_t(1)\Omega_t(1)$$

and hence

$$\mathscr{M}[\tilde{L}(z,t);\omega] = \Gamma(\omega)\frac{P_t(1)\Omega_t(1)}{P_t(q^{-\omega})\Omega_t(q^{-\omega})}. \tag{5}$$

The reason for using Mellin transform is that there is an inverse formula

$$\tilde{L}(z,t) = \frac{1}{2\pi i}\int_\uparrow \mathscr{M}[\tilde{L}(z,t);\omega]z^{-\omega}\mathrm{d}\omega,$$

where the integral is along a suitable chosen vertical line in the complex plane (here, $\Re(\omega) = \epsilon$ with $\epsilon > 0$ suitable small such that the line $\Re(\omega) = \epsilon$ is entirely contained in the domain of $\mathscr{M}[\tilde{L}(z,t)]$). In order to get now an asymptotic expansion of $\tilde{L}(z,t)$ as $z \to \infty$, we move the line of integration to the right and collect residues. Thus, we have to study the singularity structure of $\mathscr{M}[\tilde{L}(z,t)]$. For this purpose, we first recall a lemma from [13] about the (unique) positive root ρ of $P(z)$.

Lemma 1. *The root ρ of $P(z)$ is a simple root and $\rho > 1$. Moreover, ρ is the only root of $P(z)$ with $|z| \leq \rho$.*

By this lemma and the analyticity of $P_t(z)$ in z and t (recall that $P_0(z) = P(z)$), we obtain that for $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small, $P_t(z)$ has also a root ρ_t with the same properties as ρ in the above lemma (with the only difference that the root might be now complex). Thus, we see from (5), that the singularities of $\mathcal{M}[\tilde{L}(z,t)]$ which are closest to the imaginary axis are simple poles at $\omega = \log_{1/q} \rho_t + \chi_k$ with residues

$$\text{Res}\left(\mathcal{M}[\tilde{L}(z,t); \omega]; \omega = \log_{1/q} \rho_t + \chi_k\right) = \frac{P_t(1)\Omega_t(1)}{\log(1/q)\,\rho_t P_t'(\rho_t)\Omega_t(\rho_t)} \Gamma\left(\log_{1/q} \rho_t + \chi_k\right).$$

Applying now the residue theorem yields, as $z \to \infty$,

$$\tilde{L}(z,t) \sim -\frac{P_t(1)\Omega_t(1)}{\log(1/q)\,\rho_t P_t'(\rho_t)\Omega_t(\rho_t)} z^{-\log_{1/q} \rho_t} \sum_k \Gamma\left(\log_{1/q} \rho_t + \chi_k\right) z^{-\chi_k},$$

where this asymptotic holds uniformly in $|t| \leq \epsilon$ with ϵ sufficiently small (the uniformity can be seen directly or more generally comes from the fact that the denominator of (5) is analytic in both ω and t).

Now, what is left is to justify the Poisson heuristic (4). Here, we use the notion of JS-admissibility from [17] (the name comes from [16]) which ensures that we can depoissonize. The following lemma is crucial.

Lemma 2. *Let $\tilde{f}(z,t)$ and $\tilde{g}(z,t)$ be entire functions in z for $|t| \leq \epsilon$ ($\epsilon > 0$ is constant). Assume that*

$$\tilde{f}(z,t) + \frac{\partial}{\partial z}\tilde{f}(z,t) = pe^t \sum_{\ell=1}^{d} q^{\ell-1}\tilde{f}(q^\ell z, t) + \tilde{g}(z,t).$$

Then,

$$\tilde{g}(z,t) \text{ is uniformly JS-admissible} \quad \Longleftrightarrow \quad \tilde{f}(z,t) \text{ is uniformly JS-admissible},$$

where uniform means here with respect to $|t| \leq \epsilon$.

Proof. Follows with the same method of proof as Proposition 6 in [16] (only minor modifications are needed). ∎

From this result and depoissonization, we obtain that

$$\sum_{k \geq 0} p_{n,k} e^{kt} \sim -\frac{P_t(1)\Omega_t(1)}{\log(1/q)\,\rho_t P_t'(\rho_t)\Omega_t(\rho_t)} n^{-\log_{1/q} \rho_t} \sum_k \Gamma\left(\log_{1/q} \rho_t + \chi_k\right) n^{-\chi_k} \tag{6}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small. Dividing this by p_n (whose asymptotics was derived in [13], or alternatively, is also obtained from the above asymptotics by setting $t = 0$), we obtain the following result.

Proposition 1. *We have,*

$$\mathbb{E}\left(e^{L_n^{(d)} t}\right) \sim \frac{P_t(1)\Omega_t(1)\rho P'(\rho)\Omega_0(\rho)}{q^d \Omega_0(1)\rho_t P_t'(\rho_t)\Omega_t(\rho_t)} n^{-\log_{1/q}(\rho_t/\rho)} \frac{\sum_k \Gamma\left(\log_{1/q} \rho_t + \chi_k\right) n^{-\chi_k}}{\sum_k \Gamma\left(\log_{1/q} \rho + \chi_k\right) n^{-\chi_k}} \tag{7}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small.

Now, we can use this result to prove Theorem 1 and Theorem 2.

Proof of Theorem 1: Left-to-Right Maxima for $d = 1$. Since $d = 1$, (7) (slightly rearranged) becomes

$$\mathbb{E}\left(e^{\left(L_n^{(1)} - \log_{1/q} n\right)t}\right) \sim \frac{(1 - pe^t)\, Q(pe^t)}{qQ(p)} \cdot \frac{\sum_k \Gamma\left(-\log_{1/q} p - t/L + \chi_k\right) n^{-\chi_k}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}} \tag{8}$$

which holds uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small. From this, we obtain the asymptotics of all moments of $L_n^{(1)} - \log_{1/q} n$ by differentiation both sides (which is legitimate because of the uniformity of the above expansions). This gives for the m-th moment

$$\mathbb{E}\left(L_n^{(1)} - \log_{1/q} n\right)^m \sim \Phi_m^{(1)}\left(\log_{1/q} n\right),$$

where

$$\Phi_m^{(1)}(x) = \frac{\mathrm{d}^m}{\mathrm{d}t^m} \frac{(1 - pe^t)\, Q(pe^t)}{qQ(p)} \cdot \left.\frac{\sum_k \Gamma\left(-\log_{1/q} p - t/L + \chi_k\right) e^{-2k\pi i x}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi i x}}\right|_{t=0}. \tag{9}$$

For instance, for $m = 1$, this periodic function becomes

$$\Phi_1^{(1)}(x) = -\alpha_p - \frac{1}{L} \frac{\sum_k \Gamma'\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi i x}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi i x}}$$

with

$$\alpha_p = \sum_{\ell \geq 0} \frac{pq^\ell}{1 - pq^\ell}.$$

This coincides with the result in [37].

As for the non-convergence part, first observe that $\{\log_{1/q} n\}$ is dense in $[0, 1]$. (Here, $\{x\}$ denotes the fractional part of x.) Thus, we can always find subsequences which converge to two different values in $[0, 1]$. Therefore, in order to show that $L_n^{(1)} - \log_{1/q} n$ does not converge weakly to a fixed limit law, we only have to show that for two suitable subsequences, we have that (8) converges to two different functions. This problem, however, reduces to showing that $\Phi_1^{(1)}(x)$ takes on at least two different values. So, assume on the contrary, that $\Phi_1^{(1)}(x)$ is constant. This would imply that

$$\sum_k \Gamma'\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi i x} = c \sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi i x}.$$

These two Fourier series are equal if and only if all the coefficients coincide. This is, however, impossible due to different speed of decay along vertical lines of $\Gamma(x)$ and $\Gamma'(x)$. ∎

Proof of Theorem 2: Left-to-Right Maxima for $d \geq 2$. We again use (7). First, note that from the implicit function theorem, we have that ρ_t is an analytic function in t for small t with Maclaurin series expansion, as $t \to 0$,

$$\rho_t = \rho + \frac{t}{P'(\rho)} - \left(\frac{1}{P'(\rho)} + \frac{P''(\rho)}{P'(\rho)^3}\right) \frac{t^2}{2} + \mathcal{O}(t^3).$$

Thus, as $t \to 0$,

$$-\log_{1/q}\left(\rho_t/\rho\right) = \frac{\mu_d t}{L} + \frac{\sigma_d^2 t^2}{2L} + \mathcal{O}(t^3),$$

where $\mu_d := -1/(\rho P'(\rho))$. Now, we set $t = u/\sqrt{\log_{1/q} n}$ with u fixed. Plugging this into the expansions above and this expansion in turn into (7) yields after some computation

$$\mathbb{E}\left(\exp\left(L_n^{(d)} u/\sqrt{\log_{1/q} n}\right)\right) \sim \exp\left(\mu_d u \sqrt{\log_{1/q} n} + \frac{\sigma_d^2 t^2}{2}\right).$$

Hence,

$$\mathbb{E}\left(\exp\left(\left(L_n^{(d)} - \mu_d \log_{1/q} n\right) u/\sqrt{\log_{1/q} n}\right)\right) \sim e^{\sigma_d^2 t^2/2} \qquad (10)$$

from which the claimed central limit theorem follows.

What is left is to show that $\sigma_d^2 > 0$ for all $d \geq 2$ (note that $\sigma_1^2 = 0$). Since we have an explicit expression of σ_d^2, one might try to show this from this explicit expression. However, we have not been able to do so and leave such a direct proof as an open problem to the reader.

We will use instead a more subtle and (unfortunately) indirect argument. The idea of our proof is to show that $\operatorname{Var}(L_n^{(d)}) \geq c \log n$ for positive $c > 0$ and all n large enough. From this, our claim clearly follows since

$$\operatorname{Var}\left(L_n^{(d)}\right) \sim \mathbb{E}\left(L_n^{(d)} - \mu_d \log_{1/q} n\right)^2 \sim \sigma_d^2 \log_{1/q} n$$

which follows by differentiating (10) which is legitimate since (10) holds uniformly in u with $|u| \leq \epsilon$ with $\epsilon > 0$ suitable small.

In order to show the above lower bound for the variance, we will directly work with recurrences and use some ideas of Schachinger from [38]. We first set

$$S_n(t) := \sum_{k \geq 0} p_{n,k} e^{kt}.$$

Then, from the recurrence for $p_{n,k}$, we obtain that

$$S_{n+1}(t) = \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} e^t S_j(t), \qquad (n \geq 0)$$

with initial condition $S_0(t) = 1$. Next, we shift the mean

$$T_n(t) := \sum_{k \geq 0} p_{n,k} e^{\left(k - \mu_d \log_{1/q}^* n - \Xi_1\left(\log_{1/q}^* n\right)\right)t} = e^{-\mu_d\left(\log_{1/q}^* n\right)t - \Xi_1\left(\log_{1/q}^* n\right)t} S_n(t),$$

where $\log_{1/q}^* n$ is the usual log for $n \geq 1$ and 0 for $n = 0$, and $\Xi_1\left(\log_{1/q} n\right)$, with $\Xi_1(x)$ a one-periodic function, is the second term in the asymptotic expansion of the mean of $L_n^{(d)}$, which can be obtained by differentiation of (7) and setting $t = 0$, i.e,

$$\Xi_1(x) := \frac{d}{dt} \frac{P_t(1)\Omega_t(1)\rho P'(\rho)\Omega(\rho)}{q^d \Omega(1)\rho_t P_t'(\rho_t)\Omega_t(\rho_t)} \cdot \left.\frac{\sum_k \Gamma\left(\log_{1/q} \rho_t + \chi_k\right) e^{-2k\pi ix}}{\sum_k \Gamma\left(\log_{1/q} \rho + \chi_k\right) e^{-2k\pi ix}}\right|_{t=0}.$$

Note that $T_n(t)$ satisfies the recurrence

$$T_{n+1}(t) = \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} e^{\Delta_{n,j} t} T_j(t), \qquad (n \geq 0)$$

8

with initial condition $T_0(t) = e^{-\Xi_1(0)t}$ and

$$\Delta_{n,j} := 1 - \mu_d \log^*_{1/q}(n+1) - \Xi_1\left(\log^*_{1/q}(n+1)\right) + \mu_d \log^*_{1/q} j + \Xi_1\left(\log^*_{1/q} j\right).$$

Differentiating this recurrence twice with respect to t and setting $t = 0$ gives for

$$\nu_n := \sum_{k \geq 0} p_{n,k} \left(k - \mu_d \log^*_{1/q} n - \Xi_1\left(\log^*_{1/q} n\right)\right)^2$$

the following recurrence

$$\nu_{n+1} = \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \nu_j + \rho_n$$

with

$$\rho_n := 2 \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j} m_j + p_n \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j}^2,$$

where

$$m_n := \sum_{k \geq 0} p_{n,k} \left(k - \mu_d \log^*_{1/q} n - \Xi_1\left(\log^*_{1/q} n\right)\right).$$

We need now a series of lemmas.

Lemma 3. *As $n \to \infty$, we have that $m_n = o(n^{-\log_{1/q} \rho})$.*

Proof. Note that

$$m_n = p_n \left(\mathbb{E}\left(L_n^{(d)}\right) - \mu_d \log^*_{1/q} n - \Xi_1\left(\log^*_{1/q} n\right)\right) = o(p_n),$$

where the last equality follows since $\mu_d \log^*_{1/q} + \Xi_1\left(\log^*_{1/q} n\right)$ are the first two terms in the asymptotic expansion of $\mathbb{E}\left(L_n^{(d)}\right)$. The proof is now finished by plugging into this the asymptotics of p_n, which is obtained by setting $t = 0$ in (6). ∎

Lemma 4. *Let $1 \leq \ell \leq d$. Then, uniformly in j with $|j - q^\ell n| \leq n^{2/3}$,*

$$\Delta_{n,j} \sim 1 - \ell \mu_d.$$

Proof. Plugging $j = q^\ell n + \mathcal{O}(n^{2/3})$ into the definition of $\Delta_{j,n}$ and using Taylor series expansion yields the claimed result. ∎

Lemma 5. *For n large enough, we have that $\rho_n \geq cn^{-\log_{1/q} \rho}$ for a suitable $c > 0$.*

Proof. We first consider the second sum in the definition of ρ_n, which we break into two parts

$$\sum_{\ell=1}^{d} pq^{\ell-1} \sum_{|j-q^\ell| \leq n^{2/3}} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j}^2 + \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{|j-q^\ell| > n^{2/3}} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j}^2.$$

The second part is easily shown to be exponentially small by Chernoff's bound. Moreover, for the first part, we can use Lemma 4 which shows that this part is bounded from below by a positive constant for $d \geq 2$ (note that this becomes wrong for $d = 1$ since $\ell = 1$ and $\mu_1 = 1$). Thus, we get

$$p_n \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j}^2 \geq c_0 p_n \geq c_1 n^{-\log_{1/q} \rho}$$

9

with $c_0 \geq c_1 > 0$. Now, by a similar argument applied to the first sum in the definition of ρ_n with Lemma 4 replaced by Lemma 3, we obtain that

$$2 \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} \Delta_{n,j} m_j = o(n^{-\log_{1/q}\rho}).$$

Putting these two estimates together gives the claimed result. ∎

The proof of the positiveness of σ_d^2 for $d \geq 2$ is now completed with the following proposition which was essentially proved by Schachinger in [38].

Proposition 2. *Let $(b_n)_{n\geq 1}$ be a given sequence and assume that a_n is defined by*

$$a_{n+1} = \sum_{\ell=1}^{d} qp^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} a_j + b_n, \qquad (n \geq n_0)$$

with arbitrary initial conditions. If $b_n \geq cn^{-\log_{1/q}\rho}$ for n large enough with a suitable $c > 0$, then $a_n \geq c_0 n^{-\log_{1/q}\rho} \log n$ for n large enough with a suitable $c_0 > 0$.

Proof. This follows with the same method of proof as Lemma 1, part (c) in [38]. ∎

Due to Lemma 5, we can apply the above proposition to ν_n and obtain that

$$\nu_n \geq c_0 n^{-\log_{1/q}\rho} \log n$$

for n large enough with a suitable $c_0 > 0$. But since

$$\frac{\nu_n}{p_n} = \mathbb{E}\left(L_n^{(d)} - \mu_d \log_{1/q}^* n - \Xi_1\left(\log_{1/q}^* n\right)\right)^2 \sim \mathrm{Var}\left(L_n^{(d)}\right),$$

the above bound for ν_n and the asymptotics of p_n imply that $\mathrm{Var}\left(L_n^{(d)}\right) \geq c \log n$ for n large enough with a suitable $c > 0$. From this, it follows that $\sigma_d^2 > 0$ for all $d \geq 2$ as claimed.

3   Maximal Letter

Here, we are going to prove Theorem 3. The method will be similar to the one used in the previous section. Thus, we will only sketch it.

Recurrence.   First, denote by $q_{n,k}$ the probability that a geometric word satisfies (1) with general d and the maximal letter equals to k (note that for $d = 1$, we have that $p_{n,k} = q_{n,k}$). Then, as in the last section

$$q_{n+1,k} = \sum_{\ell=1}^{d} pq^{\ell-1} \sum_{j=0}^{n} \binom{n}{j} \left(1 - q^\ell\right)^{n-j} q^{\ell j} q_{j,k-\ell}, \qquad (n \geq 0, k \geq 1)$$

with initial conditions $q_{n,0} = [\![n = 0]\!]$ and $q_{0,k} = [\![k = 0]\!]$. Note that the only different to the recurrence in the previous section for $p_{n,k}$ is that the last term has second index $k - \ell$ instead of $k - 1$. This is easily explained: after the first letter is fixed to be ℓ, the remaining letters which are larger than ℓ again form a geometric word satisfying (1) but with maximal letter being $k - \ell$.

Moment-generating Function. As in Section 2, we first consider

$$\tilde{M}(z,t) := e^{-z} \sum_{n \geq 0} \sum_{k \geq 0} q_{n,k} e^{kt} \frac{z^n}{n!}.$$

Then, we have

$$\tilde{M}(z,t) + \frac{\partial}{\partial z} \tilde{M}(z,t) = \sum_{\ell=1}^{d} pq^{\ell-1} e^{\ell t} \tilde{M}(q^\ell z, t).$$

Now, by using the Mellin transform and solving the resulting equation in a similar way as in Section 2, we obtain that

$$\mathscr{M}[\tilde{M}(z,t); \omega] = \Gamma(\omega) \frac{P(e^t)\Omega_0(e^t)}{P(e^t q^{-\omega})\Omega_0(e^t q^{-\omega})}.$$

Note that the singularities closest to the imaginary axis are simple poles at $\omega = \log_{1/q} \rho - t/\log(1/q) + \chi_k$ with residues

$$\text{Res}\left(\mathscr{M}[\tilde{M}(z,t); \omega]; \omega = \log_{1/q} \rho - t/\log(1/q) + \chi_k \right)$$

$$= \frac{P(e^t)\Omega_0(e^t)}{\log(1/q)\, \rho P'(\rho)\Omega_0(\rho)} \Gamma\left(\log_{1/q} \rho - t/\log(1/q) + \chi_k \right).$$

Thus, by inverse the Mellin transform and depoissonization,

$$\sum_{k \geq 0} q_{n,k} e^{kt} \sim -\frac{P(e^t)\Omega_0(e^t)}{\log(1/q)\, \rho P'(\rho)\Omega_0(\rho)} n^{-\log_{1/q} \rho + t/\log(1/q)} \sum_k \Gamma\left(\log_{1/q} \rho - t/\log(1/q) + \chi_k \right) n^{-\chi_k}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small. Rearranging and dividing by p_n gives the following result.

Proposition 3. *We have,*

$$\mathbb{E}\left(e^{\left(M_n^{(d)} - \log_{1/q} n \right) t} \right) \sim \frac{P(e^t)\Omega_0(e^t)}{q^d \Omega_0(1)} \cdot \frac{\sum_k \Gamma\left(\log_{1/q} \rho - t/\log(1/q) + \chi_k \right) n^{-\chi_k}}{\sum_k \Gamma\left(\log_{1/q} \rho + \chi_k \right) n^{-\chi_k}}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ sufficiently small.

Moments and Non-convergence to a fixed Limit Law. First, note that differentiating the expression in Proposition 3 once and setting $t = 0$ yields the main result from [13]. Differentiating m times and setting $t = 0$ yields Theorem 3 with the periodic functions

$$\Phi_m^{(d)}(x) = \frac{d^m}{dt^m} \frac{P(e^t)\Omega_0(e^t)}{q^d \Omega_0(1)} \cdot \frac{\sum_k \Gamma\left(\log_{1/q} \rho - t/\log(1/q) + \chi_k \right) e^{-2k\pi ix}}{\sum_k \Gamma\left(\log_{1/q} \rho + \chi_k \right) e^{-2k\pi ix}} \Bigg|_{t=0}. \tag{11}$$

Moreover, the claim about non-convergence to a fixed limit law is proved as in Section 2.

4 Number of Maximal Letters

Here, we will prove Theorem 4. The method is again as in Section 2. Thus, we will only highlight differences.

11

Recurrence. Denote by $r_{n,k}$ the probability that a geometric word of length n satisfies (1) with $d = 1$ and has exactly k occurrences of the maximal letter. Then,

$$r_{n+1} = p \sum_{j=1}^{n} \binom{n}{j} p^{n-j} q^j r_{j,k} + p^{n+1} [\![n+1 = k]\!], \qquad (n \geq 0, k \geq 1)$$

with initial condition $r_{0,k} = 0$ for all $k \geq 1$. The explanation for this recurrence is similar as for the one from Section 2: after fixing the first letter (which now can only be 1 since $d = 1$), either all letters are one, in which case the probability is p^{n+1} if $k = n + 1$ (this is the second term), or there is at least one letter larger than one, in which case the problem can be reduced to considering only the subword with letters all larger than one (this is the first term).

Moment-generating Function. Set

$$\tilde{N}(z,t) := e^{-z} \sum_{n \geq 0} \sum_{k \geq 1} r_{n,k} e^{kt} \frac{z^n}{n!}.$$

Then, we have

$$\tilde{N}(z,t) + \frac{\partial}{\partial z} \tilde{N}(z,t) = p \tilde{N}(qz,t) + pe^t e^{(pe^t-1)z}.$$

The next step is to apply the Mellin transform which gives

$$\mathscr{M}[\tilde{N}(z,t); \omega] - (\omega - 1) \mathscr{M}[\tilde{N}(z,t); \omega - 1] = pq^{-\omega} \mathscr{M}[\tilde{N}(z,t); \omega] + pe^t (1 - pe^t)^{-\omega} \Gamma(\omega).$$

Define $\bar{\mathscr{M}}[\tilde{N}(z,t); \omega] = \mathscr{M}[\tilde{N}(z,t); \omega]/\Gamma(\omega)$. Then, we find

$$\bar{\mathscr{M}}[\tilde{N}(z,t); \omega] = \frac{\bar{\mathscr{M}}[\tilde{N}(z,t); \omega - 1]}{1 - pq^{-\omega}} + \frac{pe^t (1 - pe^t)^{-\omega}}{1 - pq^{-\omega}}.$$

This recurrence is slightly different from the ones encountered before. However, it has again a general solution

$$\bar{\mathscr{M}}[\tilde{N}(z,t); \omega] = \frac{pe^t}{(1 - pq^{-\omega}) Q(pq^{-\omega})} \sum_{\ell \geq 0} (1 - pe^t)^{-\omega + \ell} Q(pq^{-\omega + \ell}) + \frac{c(t)}{(1 - pq^{-\omega}) Q(pq^{-\omega})}.$$

In order to find $c(t)$ observe that $\tilde{N}(0,t) = 0$. Thus, by the direct mapping theorem from [12], we have that $\lim_{\omega \to 0} \bar{\mathscr{M}}[\tilde{N}(z,t); \omega] = 0$. This in turn yields that

$$c(t) = -pe^t \sum_{\ell \geq 0} (1 - pe^t)^{\ell} Q(pq^{\ell}).$$

Plugging this into the expression above gives that

$$\mathscr{M}[\tilde{N}(z,t); \omega] = \frac{\Gamma(\omega) pe^t}{(1 - pq^{-\omega}) Q(pq^{-\omega})} \sum_{\ell \geq 0} \left((1 - pe^t)^{-\omega + \ell} Q(pq^{-\omega + \ell}) - (1 - pe^t)^{\ell} Q(pq^{\ell}) \right).$$

The remaining argument runs now along similar lines as in Section 2. More precisely, after applying the inverse Mellin transform and depoissonization, we obtain that

$$\sum_{k \geq 1} r_{n,k} e^{kt} \sim \frac{pe^t n^{\log_{1/q} p}}{\log(1/q) Q(1)} \sum_{k} \sum_{\ell \geq 0} \Big((1 - pe^t)^{\log_{1/q} p - \chi_k + \ell} Q(q^{\ell})$$

$$- (1 - pe^t)^{\ell} Q(pq^{\ell}) \Big) \Gamma \left(-\log_{1/q} p + \chi_k \right) n^{-\chi_k}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ suitable small. Finally, dividing by p_n gives the following result.

Proposition 4. *We have,*

$$\mathbb{E}(e^{N_n t})$$

$$\sim \frac{pe^t}{qQ(p)} \cdot \frac{\sum_k \sum_{\ell \geq 0} \left((1-pe^t)^{\log_{1/q} p - \chi_k + \ell} Q(q^\ell) - (1-pe^t)^\ell Q(pq^\ell)\right) \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}}$$

uniformly in $|t| \leq \epsilon$ with $\epsilon > 0$ suitable small.

Moments and Non-convergence to a fixed Limit Law. First, note that plugging in $t = 0$ into the result from Proposition 4 must give 1. This gives the following curious identity, for which we will give a direct proof in the appendix.

Corollary 1. *We have,*

$$\sum_{\ell \geq 0} \left(\frac{q^\ell}{p} Q(q^\ell) - q^\ell Q(pq^\ell)\right) = \frac{q}{p} Q(p). \tag{12}$$

Next, observe that the claimed expansions of the moments of Theorem 4 follows from Proposition 4 and differentiation. In particular this yields for the periodic function $\Psi_m(x)$,

$$\frac{\mathrm{d}^m}{\mathrm{d}t^m} \frac{pe^t}{qQ(p)} \cdot \frac{\sum_k \sum_{\ell \geq 0} \left((1-pe^t)^{\log_{1/q} p - \chi_k + \ell} Q(q^\ell) - (1-pe^t)^\ell Q(pq^\ell)\right) \Gamma\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi ix}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) e^{-2k\pi ix}} \Bigg|_{t=0}. \tag{13}$$

For instance, for $m = 1$, this gives

$$1 - \frac{p \log_{1/q} p}{q^2 Q(p)} - \frac{\beta_p}{qQ(p)} + \frac{p}{q^2 Q(p)} \cdot \frac{\sum_k \chi_k \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}}{\sum_k \Gamma\left(-\log_{1/q} p + \chi_k\right) n^{-\chi_k}},$$

where

$$\beta_p = p^2 \sum_{\ell \geq 0} \left(\frac{\ell q^{\ell-1}}{p} Q(q^\ell) - \ell q^{\ell-1} Q(pq^\ell)\right).$$

The non-convergence to a fixed limit law follows from this as in Section 2.

Acknowledgments

5 Appendix: A Direct Proof of (12).

We use the following famous identity of Euler:

$$\sum_{j \geq 0} \frac{(-1)^j q^{\binom{j}{2}}}{(1-q) \cdots (1-q^j)} z^j = \prod_{\ell \geq 0} (1 - q^\ell z). \tag{14}$$

This identity implies that

$$Q(s) = \sum_{j \geq 0} \frac{(-1)^j q^{\binom{j+1}{2}}}{(1-q) \cdots (1-q^j)} s^j.$$

Now, observe that

$$\frac{1}{p} \sum_{\ell \geq 0} q^\ell Q(q^\ell) = \frac{1}{p} \sum_{j \geq 0} \frac{(-1)^j q^{\binom{j+1}{2}}}{(1-q) \cdots (1-q^j)} \sum_{\ell \geq 0} q^{(j+1)\ell}$$

$$= -\frac{1}{p} \sum_{j \geq 0} \frac{(-1)^{j+1} q^{\binom{j+1}{2}}}{(1-q) \cdots (1-q^j)(1-q^{j+1})} = \frac{1}{p},$$

where in the last step, we again used (14). Similarly, one shows that

$$\sum_{\ell \geq 0} q^\ell Q(pq^\ell) = \frac{1 - qQ(p)}{p}.$$

Putting everything together gives the claimed result.

References

[1] M. Archibald and A. Knopfmacher (2007). The average position of the first maximum in a sample of geometric random variables, In 2007 Conference on Analysis of Algorithms, AofA 07, *Discrete Math. Theor. Comput. Sci. Proc.*, 269–278.

[2] M. Archibald and A. Knopfmacher (2009). The average position of the dth maximum in a sample of geometric random variables, *Statist. Probab. Lett.*, **79:7**, 864–872.

[3] M. Archibald and A. Knopfmacher (2014). The largest missing value in a sample of geometric random variables, *Combin. Probab. Comput.*, **23:5**, 670–685.

[4] Z.-D. Bai, H.-K. Hwang, W.-Q. Liang (1998). Normal approximations of the number of records in geometrically distributed random variables, Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997), *Random Structures Algorithms*, **13:3-4**, 319–334.

[5] C. Brennan (2011). Ascents of size less than d in samples of geometric random variables: mean, variance and distribution, *Ars Combin.*, **102:2**, 129–138.

[6] C. Brennan and A. Knopfmacher (2005). The first and last ascents of size d or more in samples of geometric random variables, *Quaest. Math.*, **28:4**, 487–500.

[7] C. Brennan and A. Knopfmacher (2005). The distribution of ascents of size d or more in samples of geometric random variables, In 2005 International Conference on Analysis of Algorithms, *Discrete Math. Theor. Comput. Sci. Proc.*, 343–351.

[8] C. Brennan, A. Knopfmacher, T. Mansour, S. Wagner (2011). Separation of the maxima in samples of geometric random variables, *Appl. Anal. Discrete Math.*, **5:2**, 271–282.

[9] F. T. Bruss and C. A. O'Cinneide (1990). On the maximum and its uniqueness for geometric random samples, *J. Appl. Probab.*, **27:3**, 598–610.

[10] B. Eisenberg (2008). On the expectation of the maximum of IID geometric random variables, *Statist. Probab. Lett.*, **78:2**, 135–143.

[11] S. Eryilmaz (2006). A note on runs of geometrically distributed random variables, *Discrete Math.*, **306:15**, 1765–1770.

[12] P. Flajolet, X. Gourdon, P. Dumas (1995). Mellin transform and asymptotics: Harmonic sums, *Theoret. Comput. Sci.*, **144:1-2**, 3–58.

[13] M. Fuchs and H. Prodinger (2013). Words with a generalized restricted growth property, *Indag. Math.* (special issue in memory of N. G. de Bruijn), **24:4**, 124–133.

[14] W. M. Y. Goh and P. Hitczenko (2007). Gaps in samples of geometric random variables, *Discrete Math.*, **307:22**, 2871–2890.

[15] P. J. Grabner, A. Knopfmacher, H. Prodinger (2003). Combinatorics of geometrically distributed random variables: run statistics, Latin American Theoretical Informatics (Punta del Este, 2000), *Theoret. Comput. Sci.*, **297:1-3**, 261–270.

[16] H.-K. Hwang, M. Fuchs, V. Zacharovas (2010). Asymptotic variance of random symmetric digital search trees, *Discrete Math. Theor. Comput. Sci.* (special issue in honor of Philippe Flajolet), **12:2**, 103–166.

[17] P. Jacquet and W. Szpankowski (1998). Analytical de-Poissonization and its applications, *Theoret. Comput. Sci.*, **201:1-2**, 1–62.

[18] P. Kirschenhofer and H. Prodinger (1998). The number of winners in a discrete geometrically distributed sample, *Ann. Appl. Probab.*, **6**, 687–694.

[19] A. Knopfmacher and H. Prodinger (2001). Combinatorics of geometrically distributed random variables: value and position of the rth left-to-right maximum, *Discrete Math.*, **226:1-3**, 255–267.

[20] A. Knopfmacher and H. Prodinger (2004). The number of descents in samples of geometric random variables, Mathematics and Computer Science. III, 339–350, Trends Math., Birkhäuser, Basel.

[21] A. Knopfmacher and H. Prodinger (2006). The first descent in samples of geometric random variables and permutations, *Discrete Math. Theor. Comput. Sci.*, **8:1**, 215–234.

[22] A. Knopfmacher and H. Prodinger (2007). The last descent in samples of geometric random variables and permutations, *Ars Combin.*, **83**, 365–379.

[23] C.-J. Lee and S.-C. Tsai (2011). Alternating runs of geometrically distributed random variables, *JISE J. Inf. Sci. Eng.*, **27:3**, 1029–1044.

[24] G. Louchard and H. Prodinger (2003). Ascending runs of sequences of geometrically distributed random variables: a probabilistic analysis, *Theoret. Comput. Sci.*, **304:1-3**, 59–86.

[25] G. Louchard and H. Prodinger (2007). A combinatorial and probabilistic study of initial and end heights of descents in samples of geometrically distributed random variables and in permutations, *Discrete Math. Theor. Comput. Sci.*, **9:1**, 137–170.

[26] G. Louchard and H. Prodinger (2008). On gaps and unoccupied urns in sequences of geometrically distributed random variables, *Discrete Math.*, **308:9**, 1538–1562.

[27] G. Louchard, H. Prodinger, M. D. Ward (2005). The number of distinct values of some multiplicity in sequences of geometrically distributed random variables, In 2005 International Conference on Analysis of Algorithms, *Discrete Math. Theor. Comput. Sci. Proc.*, 231–256.

[28] T. Mansour and M. Shattuck (2012). Set partitions as geometric words, *Australas. J. Combin.*, **53**, 31–40.

[29] K. Oliver and H. Prodinger (2011). Words representing set partitions, *Appl. Anal. Discrete Math.*, **5**, 55–59.

[30] K. Oliver and H. Prodinger (2012). Consecutive records in geometrically distributed words, *Afrika Matematika*, **23**, 163–172.

[31] H. Prodinger (1993). Combinatorial problems related to geometrically distributed random variables, Séminaire Lotharingien de Combinatoire (Gerolfingen, 1993), 87–95, Prépubl. Inst. Rech. Math. Av., 1993/34, Univ. Louis Pasteur, Strasbourg.

[32] H. Prodinger (1996). Combinatorics of geometrically distributed random variables: left-to-right maxima, Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993), *Discrete Math.*, **153:1-3**, 253–270.

[33] H. Prodinger (2001). Combinatorics of geometrically distributed random variables: inversions and a parameter of Knuth, *Ann. Comb.*, **5:2**, 241–250.

[34] H. Prodinger (2002). Combinatorics of geometrically distributed random variables: value and position of large left-to-right maxima, *Discrete Math.*, **254:1-3**, 459–471.

[35] H. Prodinger (2006). *d*-records in geometrically distributed random variables, *Discrete Math. Theor. Comput. Sci.*, **8:1**, 273–283.

[36] H. Prodinger (2012). On the *m*-enumeration of maximum and left-to-right maxima in geometrically distributed words, *Pure Math. Appl.*, **23**, 143–154.

[37] H. Prodinger (2012). Set partitions, words, and approximate counting with black holes. *Australas. J. Combin.*, **54**, 303–310.

[38] W. Schachinger (1995). On the variance of a class of inductive valuations of data structures for digital search, *Theoret. Comput. Sci.*, **144**, 251–275.