Pascal's formulas and vector fields

Philippe Chassaing[a] Jules Flin[b] Alexis Zevio[a]

Submitted: Nov 28, 2023; Accepted: Jan 27, 2025; Published: Mar 14, 2025

Abstract

We consider four examples $T = (T(n,k))_{0 \leqslant k \leqslant n}$ of combinatorial triangles (Pascal, Stirling of both types, Euler) : through saddle-point asymptotics, their *Pascal's formulas* define four vector fields, together with their field lines that turn out to be the conjectured limit of sample paths of four well known Markov chains. We prove this asymptotic behaviour in three of the four cases. Our results lead to a new proof of Koršunov's formula for the enumeration of accessible complete deterministic automata, and to the design of an efficient rejection method for the random generation of this class of automata.

Mathematics Subject Classifications: 05A10,05A19,60F17,60J10

1 Introduction

1.1 Pascal's formulas and transition probabilities

Set $S = \{(n,k) \in \mathbb{N}^2,\ 0 \leqslant k \leqslant n\}$, $\mathring{S} = \{(n,k) \in \mathbb{N}^2,\ 0 < k < n\}$, and $S^\star = S \setminus \{(0,0)\}$. Besides Pascal's triangle, many other combinatorial triangles

$$T = (T(n,k))_{(n,k)\in S}$$

of interest satisfy a recursion formula similar to Pascal's formula, i.e. of the following form, for $(n,k) \in S^\star$:

$$T(n,k) = a(n,k)T(n-1,k-1) + b(n,k)T(n-1,k), \tag{1}$$

with the convention that either $(n,k) \in S$ or $T(n,k) = 0$. For instance, relation (1) holds true for the following triangular arrays :

- for Pascal's triangle, if $(a,b)(n,k) = (1,1)$;

[a]Institut Élie Cartan, Université de Lorraine, Nancy, France (chassaingph@gmail.com, zevioalexis@gmail.com).
[b]Télécom Sud-Paris, Évry-Courcouronnes (jules.flin@telecom-sudparis.eu).

- for Stirling numbers of the first kind, if $(a,b)(n,k) = (1, n-1)$;
- for Stirling numbers of the second kind, if $(a,b)(n,k) = (1, k)$;
- for Eulerian numbers, if $(a,b)(n,k) = (n-k, k+1)$.

In view of (1), for $(n,k) \in S^\star$, consider

$$(p_0(n,k), p_1(n,k)) = \left(\frac{b(n,k)T(n-1,k)}{T(n,k)}, \frac{a(n,k)T(n-1,k-1)}{T(n,k)} \right) \tag{2}$$

as some transition probabilities from (n,k) to $(n-1,k)$, resp. to $(n-1,k-1)$.

Definition 1. For each of these four triangular arrays, the transition probabilities

$$(p_\varepsilon(n,k))_{(\varepsilon,(n,k))\in\{0,1\}\times S^\star},$$

together with the initial state (m,ℓ), define a Markov chain $\mathcal{W} = (\mathcal{W}_k)_{0\leqslant k\leqslant m}$.

For the sake of brevity, the Markov chains, as well as their transition probabilities, related to the four combinatorial triangles, are denoted by $(\mathcal{W}^{(Pa)}, p_\varepsilon^{(Pa)})$, $(\mathcal{W}^{(S1)}, p_\varepsilon^{(S1)})$, $(\mathcal{W}^{(S2)}, p_\varepsilon^{(S2)})$, $(\mathcal{W}^{(E)}, p_\varepsilon^{(E)})$, respectively. We shall see later that their terminal state $\mathcal{W}_m$ is always the origin.

1.2 Aim of the article

In this article we shall see that these four Markov chains are closely related to well-studied (to varying degrees) combinatorial stochastic processes (see e.g. [Pit06]) : the simple random walk, the chinese restaurant process, the coupon collector problem and the one-dimensional internal DLA, respectively. In each of these combinatorial stochastic processes, one of the main focuses, if not the main focus, is the study of a random process $X = (X_n)_{n\geqslant 0}$ that fits the following description : $X_0 = 0$, and, for $n \geqslant 0$, $Y_{n+1} = X_{n+1} - X_n$ is a Bernoulli random variable. Consider the time reversal W of X, defined below, and its renormalization w_m.

Definition 2. For a given $m \geqslant 1$, set

$$W_n = (m-n, X_{m-n}) \in S, \quad 0 \leqslant n \leqslant m,$$
$$w_m(t,\omega) = m^{-1} X_{\lfloor mt \rfloor}(\omega) \in S, \quad 0 \leqslant t \leqslant 1,$$

Then $W = (W_n)_{0\leqslant n\leqslant m}$ is called the *time reversal* of X. Note that, by definition, $W_n = (0,0)$ if and only if $n = m$.

For some history about time reversal of Markov processes, see e.g. the notes at the end of [CW05, Ch. 10], or [Kol35]. In this section, we give a simple combinatorial description of W, in terms of the Pascal formula of each triangle. This description is well known for the simple random walk, while it seems largely overlooked, as far as we know, for the other triangles, perhaps because the ratios in relation (2) are easily tractable only in the case of the Pascal's triangle. The proof of this combinatorial description is then given in Section 2. In Section 1.3, we provide simulations of w_m that support the results of Section 1.4. The main results of the article, Theorem 8 and Theorem 9, are given in Section 1.4 : relying on results from analytic combinatorics by [Goo61] and [Ben73], Theorem 8 details the asymptotic behaviour of the transition probabilities of (2). Theorem 9 then states the convergence of w_m to the solutions of some ODEs related to the asymptotic transition probabilities obtained in Theorem 8. The decomposability of the combinatorial structures enumerated by the four triangles has a striking consequence on the structure of the solutions : their sets are invariant by positive homotheties. Finally, in Section 1.5.1, we sketch a simple probabilistic proof of an old result by Koršunov [Kor78] about the enumeration of accessible complete deterministic automata (ACDA) with k letters and n vertices, using Theorem 9 and a nice combinatorial bijection due to Bassino and Nicaud [BN06]. In turn, in Section 1.5.2, Proposition 5 and Theorem 9 provide an efficient rejection method for the random generation of ACDAs. Note that, similarly, Propositions 4 and 7 provide elegant algorithms for the random generation of a permutation with a prescribed number of cycles, or with a prescribed number of descents. Section 3 is devoted to the proof of Theorem 8 and Section 4 is devoted to the proof of Theorem 9. Section 5 contains some trite computations that would have disrupted the thread of arguments in the previous sections.

1.2.1 Simple random walk

Assume that $(Y_i)_{i\geqslant 1}$ is a Bernoulli process, i.e. a sequence of i.i.d. Bernoulli random variables with parameter $p \in (0,1)$, so that X is the simple random walk. Then

Proposition 3. *The distribution of the time reversal of X does not depend on p. More precisely, for any $(m,\ell) \in S^\star$,*

$$((W_n)_{0\leqslant n\leqslant m} \mid W_0 = (m,\ell)) \overset{(d)}{=} \left((\mathcal{W}_n^{(Pa)})_{0\leqslant n\leqslant m} \mid \mathcal{W}_0^{(Pa)} = (m,\ell)\right).$$

This result goes back at least to the introduction of the concept of sufficiency by Fisher around 1920 [Sti73]. We recall its proof in Section 2. In the next cases, the Bernoulli random variables Y_i are not i.i.d. .

1.2.2 Chinese restaurant process

In the Chinese restaurant process with $(0,\theta)$ seating plan, defined at Section 2 (see also, e.g., [Pit06, Ch. 3]), let X_n denote the number of occupied tables after the arrival of the nth customer. The increments $(Y_n)_{n\geqslant 1}$ of the stochastic process X are thus Bernoulli random variables.

Proposition 4. *The distribution of the time reversal of X does not depend on θ. More precisely, for any $(m,\ell) \in S^\star$,*

$$((W_n)_{0\leqslant n\leqslant m} \mid W_0 = (m,\ell)) \overset{(d)}{=} \left((\mathcal{W}_n^{(S1)})_{0\leqslant n\leqslant m} \mid \mathcal{W}_0^{(S1)} = (m,\ell)\right).$$

1.2.3 Coupon collector's problem

Consider the coupon collector's problem with N different items. Let X_n denote the number of different items in the collection after the nth step. Again, the increments $(Y_n)_{n\geqslant 1}$ of the stochastic process X are Bernoulli random variables, and :

Proposition 5. *The distribution of the time reversal of X does not depend on N. More precisely, for any $(m,\ell) \in S^\star$,*

$$((W_n)_{0\leqslant n\leqslant m} \mid W_0 = (m,\ell)) \overset{(d)}{=} \left((\mathcal{W}_n^{(S2)})_{0\leqslant n\leqslant m} \mid \mathcal{W}_0^{(S2)} = (m,\ell)\right).$$

Remark 6. As a consequence, in the three previous cases, given the data $(X_n)_{0\leqslant n\leqslant m}$, X_m (or W_0) are sufficient statistics for the parameters p, θ or N, respectively.

1.2.4 One-dimensional Internal Diffusion Limited Aggregation

Finally, in the one-dimensional Internal Diffusion Limited Aggregation process (or iDLA), let X_n denote the number of particles settled to the right of the origin after the release of the nth particle. Then

Proposition 7. *For any $(m,\ell) \in S^\star$, the distribution of the time reversal of X is given, for any $(m,\ell) \in S^\star$, by*

$$((W_n)_{0\leqslant n\leqslant m} \mid W_0 = (m,\ell)) \overset{(d)}{=} \left((\mathcal{W}_n^{(E)})_{0\leqslant n\leqslant m} \mid \mathcal{W}_0^{(E)} = (m,\ell)\right).$$

Note that, in the case of the Euler triangle, there exists an almost perfect analog to the previous statements, in which some parameter plays the same rôle as p, θ or N, but we were not able to remove some serious bottlenecks in the proofs. More precise definitions, references and proofs related to these four stochastic processes are to be found in Section 2.

1.3 Simulations

The behaviour of $\mathcal{W}^{(Pa)}$ is well understood since forever. Quite recently, [AC19] gave a rather precise analysis of $\mathcal{W}^{(S2)}$, with combinatorial analysis of finite automata as a motivation, see Section 1.5. In this article, we aim to improve and extend some of these results and proofs.

In this section, in order to surmise the behaviour of the four time-reversed Markov chains, we present the result of some simulations. For each case, the figures below show sample paths starting at (m, mt) with $t \in \{0.05, ..., 0.95\}$ and $m = 500$:

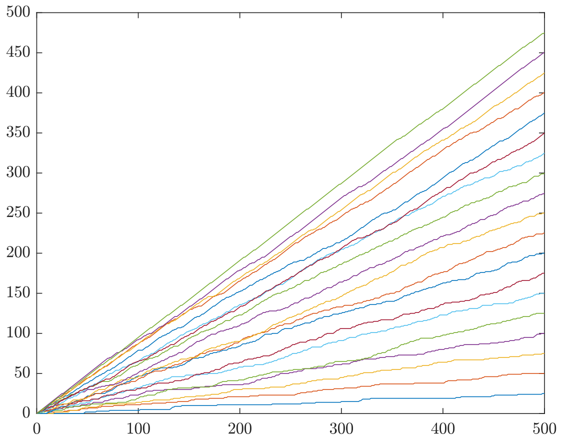

Figure 1: Pascal's triangle.

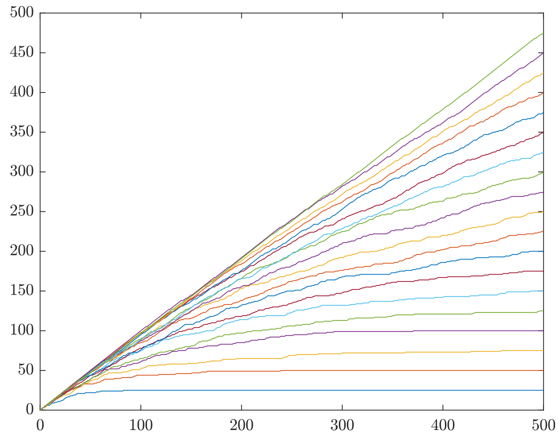

Figure 2: Stirling numbers of the second kind.

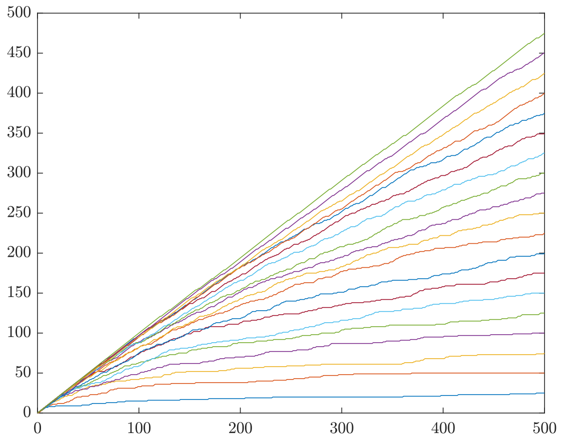

Figure 3: Stirling numbers of the first kind.

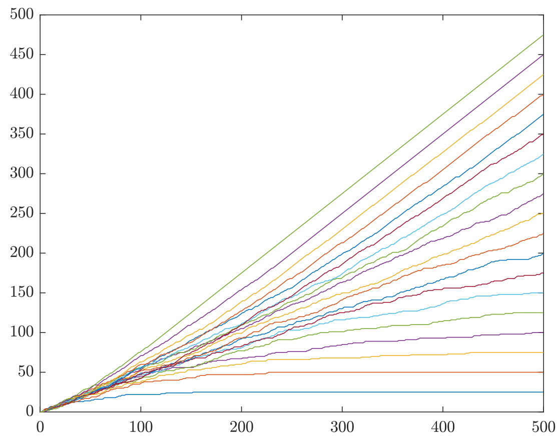

Figure 4: Eulerian numbers.

Now, in order to compare the four combinatorial triangles, we show the average of 100 sample paths for each triangle, for $m = 1000$ and $t \in \{0.05, ..., 0.95\}$:

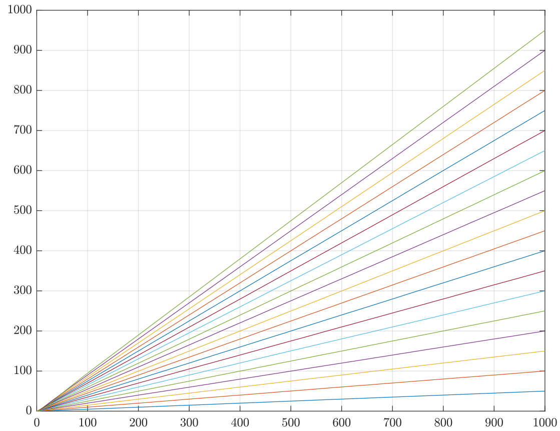

Figure 5: Pascal's triangle.

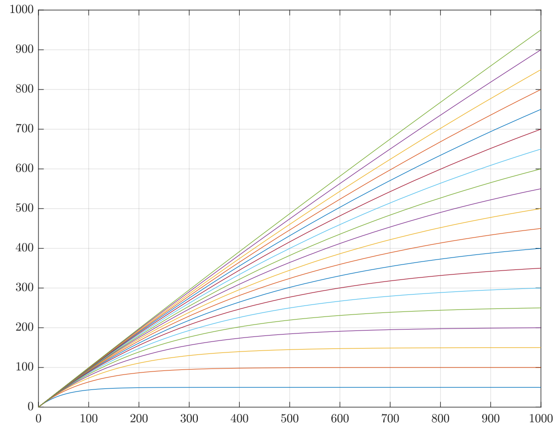

Figure 6: Stirling numbers of the second kind.

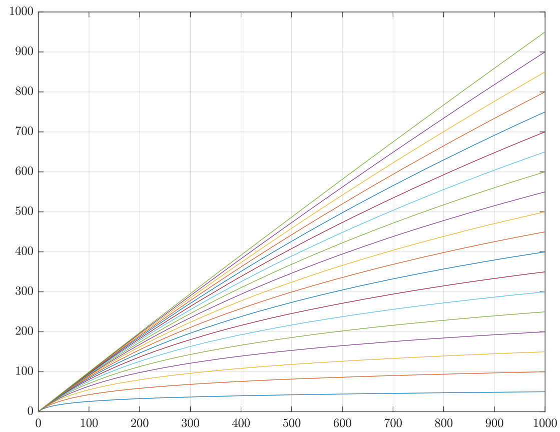

Figure 7: Stirling numbers of the first kind.

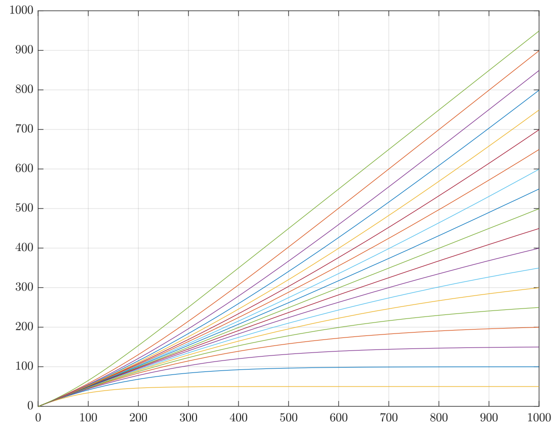

Figure 8: Eulerian numbers.

In the first two cases, the smooth nature of these averaged paths is not unexpected due to old, and more recent, fluid approximation results, see [AC19] or the next sections. This article aims for a global explanation of the asymptotic behaviour of the four Markov chains exhibited by these simulations, see Theorem 8 below.

1.4 Asymptotics of sample paths, and field lines of vector fields

Combinatorial analysis, see [Goo61] or [Ben73], yields that,

Theorem 8. *In each of the four cases, there exists a function $\varphi : (0, +\infty) \to [0, 1]$ such that, for any positive number λ_∞, when $(m, \ell) \to +\infty$ and $\lim m/\ell = 1 + \lambda_\infty$,*

$$\lim p_1(m, \ell) = \varphi(\lambda_\infty).$$

At the end of this section, the function φ is described for each of the four cases.

As a direct consequence of Theorem 8, one expects a fluid approximation of the previous Markov chains by a special family of curves : let $(x, \gamma_\lambda(x))_{0 \leqslant x \leqslant n}$ be the field line going through the point $(1, 1/(1 + \lambda))$ for the vector field $(1, \varphi(-1 + x/y))$, or, equivalently, let γ_λ denote the solution of the ODE

$$y' = \varphi\left(\frac{x - y}{y}\right) \tag{3}$$

that satisfies $y(1) = 1/(1 + \lambda)$. For $(m, \ell) \in \mathring{S}$, set

$$\lambda(m, \ell) = \frac{m - \ell}{\ell},$$

often abridged in λ, and let $\mathbb{P}_{(m,\ell)}$ denote the probability distribution of the Markov chain $\mathcal{W}$ starting from (m, ℓ). For the sake of brevity, the notation $\gamma_{\lambda(m,\ell)}$ is abridged in $\gamma_{m,\ell}$ and denotes the solution of the ODE (3) that satisfies $y(1) = \ell/m$. So far we have a complete proof of the fluid approximation of these Markov chains only in the first three cases :

Theorem 9. *In the first three cases, for any $\eta \in (0, 1/2)$ and any $\lambda_\infty > 0$, when $(m, \ell) \to +\infty$ and $\lim \lambda(m, \ell) = \lambda_\infty$,*

$$\lim \mathbb{P}_{(m,\ell)} \left(\sup_{0 \leqslant t \leqslant 1} \left(|w_m(t) - \gamma_{\lambda_\infty}(t)| \right) \geqslant m^{-\eta} + |\lambda(m, \ell) - \lambda_\infty| \right) = 0.$$

Note that the special form of the ODE (3) entails that the set of field lines is invariant by positive homotheties. Also, note that almost sure convergence of w_m to γ_{λ_∞} with respect to $\|.\|_\infty$ would be a direct consequence of the bounds obtained in the proof of Theorem 9, if the stochastic processes w_m were embedded in the same probability space.

Theorem 9 seems to hold true for Eulerian numbers, according to our simulations (see Section 1.3), but this remains an open question. For Stirling numbers of the first kind, Theorem 9 seems to be new, as far as we know. For Stirling numbers of the second kind, Theorem 9 is a vastly improved version of a result that appeared in [AC19]. In [AC19], the proof relies mainly on Wormald's method [War19], and on uniform bounds for

$$m \ |p_1(m, \ell) - \varphi\left(\lambda(m, \ell)\right)|\,,$$

on domains that approach $\mathring{S}$ as well as possible. These uniform bounds follow from a careful asymptotic analysis of $T(m, \ell)$, that should have some interest in itself. However the proof given in the next pages is much simpler, if only because we obtained an explicit form for the solutions of the four ODEs.

Our choice of four combinatorial triangles may seem arbitrary, and we confess it is : for instance, Bell's triangle or Delannoy's triangle also have Pascal's formulas, but of a slightly different form. We do not know if the approach of this article still produces results for Bell's triangle or Delannoy's triangle, in spite of these slight differences.

1.4.1 Description of the limit rate

- *Pascal's triangle.* Observe that, for all $(m, \ell) \in S^\star$,

$$p_1(m, \ell) = \frac{\ell}{m},$$

so that

$$\varphi_{Pa}(\lambda) = \frac{1}{1+\lambda}.$$

Relation (3) reduces to $y' = y/x$, with the linear functions as solutions, as expected.

- *Stirling numbers of the first kind.* For $\lambda > 0$, set

$$\varphi_{S1}(\lambda) = 1 - \zeta_{S1}(\lambda), \tag{4}$$

where $\zeta_{S1}(\lambda)$is the unique solution, in $(0, 1)$, of

$$\frac{\zeta_{S1}}{(\zeta_{S1} - 1)\ln(1 - \zeta_{S1})} = 1 + \lambda. \tag{5}$$

Relation (3) reduces to

$$y = \frac{xy' \ln(y')}{y' - 1}. \tag{6}$$

Then, for $x \geqslant 0$,

$$\gamma_\lambda(x) = \frac{1 - \zeta_{S1}(\lambda)}{\zeta_{S1}(\lambda)} \ln\left(1 + x\frac{\zeta_{S1}(\lambda)}{1 - \zeta_{S1}(\lambda)}\right). \tag{7}$$

- *Stirling numbers of the second kind.* For $\lambda > 0$, set

$$\varphi_{S2}(\lambda) = e^{-\zeta_{S2}(\lambda)}, \tag{8}$$

where $\zeta_{S2}(\lambda)$ is the unique positive solution of

$$\frac{\zeta_{S2}(\lambda)}{1 - e^{-\zeta_{S2}(\lambda)}} = 1 + \lambda. \tag{9}$$

Relation (3) reduces to

$$y = \frac{x(y' - 1)}{\ln(y')}. \tag{10}$$

Then, for $x \geqslant 0$,

$$\gamma_\lambda(x) = \frac{1 - e^{-x\,\zeta_{S2}(\lambda)}}{\zeta_{S2}(\lambda)}. \tag{11}$$

Incidentally, the solution ζ_{S2} also has an expression in terms of the Lambert function, see [AC19].

- *Eulerian numbers.* For $\lambda > 0$, set

$$\varphi_E(\lambda) = 1 - \frac{\zeta_E(\lambda)}{(1+\lambda)(e^{\zeta_E(\lambda)} - 1)}. \tag{12}$$

where $\zeta_E(\lambda)$ the unique solution, in $\mathbb{R}$, of

$$\frac{1}{1+\lambda} = \frac{1}{1 - e^{-\zeta_E(\lambda)}} - \frac{1}{\zeta_E(\lambda)}. \tag{13}$$

Then, for $\lambda \geqslant 0$ and for $x \geqslant 0$,

$$\gamma_\lambda(x) = x \left(\frac{1}{1 - e^{-x\zeta_E(\lambda)}} - \frac{1}{x\zeta_E(\lambda)} \right) \tag{14}$$

is the solution of

$$y' = \varphi_E\left(\frac{x-y}{y}\right). \tag{15}$$

that satisfies $\gamma_\lambda(1) = 1/(1+\lambda)$.

For the solutions γ_λ of the three last ODEs, basic algorithms failed us, so we had to resort to guessing, with the help of combinatorial and probabilistic arguments detailed in the proof of Theorem 9 (see Section 4).

1.5 Enumeration and random generation of accessible complete deterministic automata

The initial motivation for this article is the study of accessible complete deterministic automata (ACDAs) (see [BN06, AC19] for definitions). ACDAs have nice properties of minimality, as observed in [CP05] : empirically, either, very often, an ACDA is minimal with respect to the regular language $\mathcal{L}$ it recognizes, or it has typically at most one or two additional states, once compared with the minimal automata that recognizes $\mathcal{L}$. Let $a_{k,n}$ denote the number of ACDAs with k letters and n vertices. According to Koršunov [Kor78, Kor86], for any given $k \geqslant 2$,

$$a_{k,n} \sim c_k \left\{ \begin{matrix} kn+1 \\ n \end{matrix} \right\} n!, \tag{16}$$

in which

$$c_k = 1 - k\, e^{-\zeta_{S2}(k-1)}, \tag{17}$$

and ζ_{S2} is defined by (9). In this Section, following [BN06, AC19], we sketch a probabilistic proof of Koršunov's formula, that relies on Theorem 9. The arguments in the proof, namely Proposition 5 and Theorem 9, are also crucial in the design of an efficient rejection method for the random generation of ACDAs, see Section 1.5.2.

1.5.1 Enumeration

According to [BN06], there exists a bijection between the set of ACDA with k letters and n vertices and a subset $\mathcal{A}_{k,n}$ of the set $\Omega_{kn+1,n}$ of surjections from $[\![kn+1]\!]$ to $[\![n]\!]$. Thus (16) states that the ratio $\#\mathcal{A}_{k,n}/\#\Omega_{kn+1,n}$ converges to c_k with n. But an element ω of $\Omega_{kn+1,n}$ can be seen as the history of a coupon collector process such that the collection of n items is complete at step $kn+1$. If, for a random element $\omega \in \Omega_{kn+1,n}$, we set

$$X_m(\omega) = \mathrm{Card}\left(\omega([\![m]\!])\right), \ 1 \leqslant m \leqslant kn+1,$$

and if (Y, W) are defined accordingly, then W has the same distribution as $\mathcal{W}^{(S2)}$ under $\mathcal{P}_{(kn+1,n)}$. As a consequence, in the notations of Section 1.4,

$$\mathbb{P}_{(kn+1,n)}(\mathcal{A}_{k,n}) = \frac{\#\mathcal{A}_{k,n}}{\#\Omega_{kn+1,n}} = \frac{a_{k,n}}{\left\{ {kn+1 \atop n} \right\} n!}, \tag{18}$$

and Koršunov's formula can be rephrased as

$$\lim_n \mathbb{P}_{(kn+1,n)}(\mathcal{A}_{k,n}) = c_k. \tag{19}$$

Now, according to [BN06], $\mathcal{A}_{k,n}$ is the set of elements $\omega \in \Omega_{kn+1,n}$ such that

$$\forall \ell \in [\![0, n-1]\!], \quad X_{\ell k+1}(\omega) \geqslant \ell + 1,$$

or, equivalently,

$$\forall \ell \in [\![0, kn]\!], \quad k\, X_\ell(\omega) \geqslant \ell. \tag{20}$$

Relation (20) is the condition usually required from a breadth first search walk to insure the connexity of the underlying graph : here this condition insures the accessibility of the automata. Note that here $(m, \ell) = (kn+1, n)$, thus :

$$\lambda_\infty = \lim_n \lambda(kn+1, n) = k-1.$$

We shall now sketch the argument, taken from [AC19], which shows, with the help of Theorem 9, that $\Upsilon_n = \overline{\mathcal{A}_{k,n}}$ satisfies

$$\lim_n \mathbb{P}_{(kn+1,n)}(\Upsilon_n) = 1 - c_k = k\, e^{-\zeta_{S2}(k-1)}.$$

If ω is to belong to $\mathcal{A}_{k,n}$, then $w_{kn+1}(\omega)$ is required to stay above the line $y = x/k$. But according to (7), for n large, $w_{kn+1}(\omega)$ is close to the corresponding concave limit field line,

$$\gamma_{k-1}(t) = \frac{1 - e^{-\zeta_{S2}(k-1)t}}{\zeta_{S2}(k-1)},$$

that crosses the line $y = x/k$ only at its endpoints $(0,0)$ and $(1, 1/k)$. Thus, according to Theorem 9, except for an exponentially small probability, the sample path

$\{(\ell, X_\ell), 0 \leqslant \ell \leqslant kn+1\}$ crosses the line $y = x/k$ only close to its endpoints. The probability that such a crossing occurs close to $(0,0)$ is $\mathcal{O}(n^{-2})$, see [AC19, Proposition 2]. As a consequence, $\mathbb{P}_{(kn+1,n)}(\Upsilon_n)$ has the same asymptotic behaviour as the probability that the sample path $\{(\ell, X_\ell), 0 \leqslant \ell \leqslant kn+1\}$ crosses the line $y = x/k$ close to its endpoint $(kn+1, n)$. Close to this endpoint, as a consequence of Theorem 8, the sample path has approximately the same transition probabilities as a simple random walk with step distribution

$$\left(1 - e^{-\zeta_{S2}(k-1)}\right)\delta_0 + e^{-\zeta_{S2}(k-1)}\delta_{-1},$$

thus the sample path and the simple random walk have essentially the same probability of crossing the line close to the endpoint. Note that if the simple random walk crosses the line $y = x/k$ at all, it crosses that line close to the endpoint. Thus $\mathbb{P}_{(kn+1,n)}(\Upsilon_n)$ converges to the probability p_k that the simple random walk crosses the line $y = x/k$ at all. Finally, as follows for instance from the Pollaczek-Khinchine formula (cf. Corollary 6.6 of [Asm03], or cf. Proposition 3 of [AC19]), $p_k = 1 - c_k$. In [AC19], the authors turn these heuristics into a proof of Koršunov's formula.

1.5.2 Random generation

Once a random element ω of $\mathcal{A}_{k,n}$ has been produced, the random ACDA is easily recovered from ω with the help of the bijection described in [BN06]. We propose to produce a random element of $\Omega_{kn+1,n}$, and to reject it if it does not belong to $\mathcal{A}_{k,n}$. This direct generation of a random element ω of $\Omega_{kn+1,n}$, with the help of Proposition 5, saves an average of

$$\frac{n^{kn+1}}{\left\{ {kn+1 \atop n} \right\} \times n!} \simeq \sqrt{\frac{e^{\zeta_{S2}(k-1)}-1-\zeta_{S2}(k-1)}{e^{\zeta_{S2}(k-1)}-1}}\; e^{nJ(\zeta_{S2}(k-1))}$$

steps, compared with the (admittedly clumsy) rejection algorithm in which, at each step, a random mapping ω from $[\![kn+1]\!]$ to $[\![n]\!]$ is produced, and the random mapping is rejected unless it is a surjection. Here J is a positive and decreasing function of k discussed in Section 5.2 of [AC19].

According to Proposition 5, the time reversal W of X has the same distribution as $\mathcal{W}^{(S2)}$, so the first step of our rejection algorithm is the precomputation of the triangle

$$\left(\left\{ {m \atop \ell} \right\}\right)_{0 \leqslant \ell \leqslant m \leqslant kn+1},$$

in order to obtain the transition probabilities of $\mathcal{W}^{(S2)}$. As a very cheap second step, given the transition probabilities, one can proceed to the generation of the sample path, $W(\omega) = (W_t(\omega))_{0 \leqslant t \leqslant kn+1}$, of the Markov chain. At this stage, one can check if ω belongs to $\mathcal{A}_{k,n}$ directly by inspection of $W(\omega)$, with the help of (20), without recovering completely ω itself, and one can stop the generation of the sample path as soon as (20) is violated. Then one obtains the sample path of $X(\omega) = (X_t(\omega))_{0 \leqslant t \leqslant kn+1}$ by time reversal. The cost of rejections caused by (20) is discussed in the next paragraph. In the last step ω is recovered from $X(\omega)$ as follows: $\omega(1)$ is drawn uniformly from $[\![n]\!]$, and, recursively, $\omega(t)$ is drawn uniformly from $\omega([\![t-1]\!])$ if $Y_t(\omega) = 0$ and from $[\![n]\!]\backslash\omega([\![t-1]\!])$ if $Y_t(\omega) = 1$.

Besides the explicit description of the distribution of W given by Proposition 5, time reversal has another benefit : in general, either (20) is violated by ω for some $t \geqslant kn - \mathcal{O}\left(n^{1-\eta}\right)$, or $\omega \in \mathcal{A}_{k,n}$. Actually, due to [AC19], Proposition 2, and due to Theorem 9, the third case, i.e. (20) is violated by $X_t(\omega)$ only for some $t \leqslant kn - \mathcal{O}\left(n^{1-\eta}\right)$, occurs with a probability $\mathcal{O}\left(n^{-2}\right)$. Thus, in most cases, due to time reversal, the rejection takes place in the early stages of the generation of $W(\omega)$, at a cost smaller than $\mathcal{O}\left(n^{1-\eta}\right) = \mathcal{O}\left(n^{1/2+\varepsilon}\right)$, once $\eta = 1/2 - \varepsilon$ is chosen arbitrarily in $(0, 1/2)$. It follows that the average cost of the generation of a sample path of W that satisfies (20) is essentially the cost $\mathcal{O}\left(kn+1\right)$ of the final, successful attempt.

Remark 10. Similarly, Propositions 4 and 7 provide elegant algorithms for the random generation of a permutation of m objects containing exactly ℓ cycles, or having exactly ℓ descents : a first step, as above, is the precomputation of the triangles

$$\left(\begin{bmatrix} t \\ k \end{bmatrix}\right)_{0\leqslant k\leqslant t\leqslant m} \text{ or } \left(\left\langle \begin{matrix} t \\ k \end{matrix} \right\rangle\right)_{0\leqslant k\leqslant t\leqslant m},$$

in order to obtain the transition probabilities of $\mathcal{W}^{(S1)}$ or $\mathcal{W}^{(E)}$. This allows for the cheap generation of a sample path of W that starts from (m, ℓ), and one gets the sample paths of X and Y at once. Then one runs an avatar of the random permutation process of Remark 12 in which :

- if W is a copy of $\mathcal{W}^{(S1)}$ starting from (m, ℓ), for $1 \leqslant t \leqslant m$, the insertion of t in a random list of elements of $[\![t-1]\!]$ takes place at the end if $Y_t(\omega) = 1$, and at any of the remaining $t-1$ positions if $Y_t(\omega) = 0$. One recovers the random permutation ω with ℓ cycles from the final random list of elements of $[\![m]\!]$ using the first fundamental transformation of Foata and Schützenzerger, see [Lot97, Chap. 10.2] ;
- if W is a copy of $\mathcal{W}^{(E)}$ starting from (m, ℓ), for $1 \leqslant t \leqslant m$, the insertion of t in a random list of elements of $[\![t-1]\!]$ takes place at an ascent of this list (or at the beginning of the list) if $Y_t(\omega) = 1$, at a descent or at the end if $Y_t(\omega) = 0$. Then the final random list of elements of $[\![m]\!]$ is a random permutation ω with ℓ descents.

2 Time reversal and Markov property of the reversed process

In this section, we prove Propositions 3, 4, 5 and 7. The notations X_n, Y_n, W_n are defined at Section 1.2.

2.1 Time reversal

As already known at least since Kolmogorov, see [Kol35, (7)], a time-reversed Markov process is still an (eventually inhomogeneous) Markov process. Let us recall the basic facts that we need here : if h_k denotes the probability distribution of X_k and if $X = (X_k)_{k\geqslant 0}$ is an inhomogeneous Markov chain with kernels $(Q_k)_{k\geqslant 0}$, i.e.

$$Q_{k,i,j} = \mathbb{P}\left(X_{k+1} = j \mid X_k = i\right),$$

then

Proposition 11. *$W = (W_n)_{0 \leqslant n \leqslant m}$ is a Markov chain with state space S and with kernel P defined on $S^\star$ by*

$$P_{(n,i),(n-1,j)} = \frac{h_{n-1}(j) Q_{n-1,j,i}}{h_n(i)}.$$

Proof. First, since $(0,0)$ is only reached, eventually, at time m, there is no need to define $P_{(0,0),(.,.)}$. Also, P is a probability kernel due the Chapman-Kolmogorov equations for (h_n) and (Q_n). Then,

$$\mathbb{P}\left((X_k)_{0 \leqslant k \leqslant m} = (x_k)_{0 \leqslant k \leqslant m}\right) = h_0(x_0) \prod_{k=0}^{m-1} Q_{k,x_k,x_{k+1}},$$

thus, provided that $x_m = \ell$,

$$\begin{aligned}\mathbb{P}\left((X_k)_{0 \leqslant k \leqslant m} = (x_k)_{0 \leqslant k \leqslant m} \mid X_m = \ell\right) &= \frac{h_0(x_0)}{h_m(\ell)} \prod_{k=0}^{m-1} Q_{k,x_k,x_{k+1}}, \\ &= \prod_{k=0}^{m-1} P_{(k+1,x_{k+1}),(k,x_k)}.\end{aligned}$$

That is,

$$\mathbb{P}\left((W_k)_{0 \leqslant k \leqslant m} = (m-k, x_{m-k})_{0 \leqslant k \leqslant m} \mid W_0 = (m,\ell)\right) = \prod_{k=0}^{m-1} P_{(k+1,x_{k+1}),(k,x_k)},$$

as expected. □

Except for Eulerian numbers, $h_n(k) = T(n,k)\theta^k / T_n(\theta)$, or $h_n(k) = T(n,k)\theta^{k\downarrow} / T_n(\theta)$, in which $T_n(\theta)$ is a normalizing constant :

$$T_n(\theta) = \sum_{k=0}^{n} T(n,k)\theta^k, \quad \text{or} \quad T_n(\theta) = \sum_{k=0}^{n} T(n,k)\theta^{k\downarrow}. \tag{21}$$

For Eulerian numbers, $h_n(k) = T(n,k)/T_n(1) = T(n,k)/n!$.

Remark 12. Note that Q_n results from a natural growing mechanism with independent steps, that is, a Markovian growth process, obtained as follows :

- by addition of an $n+1$th letter, either $\mathtt{a}$ or $\mathtt{b}$, at the end of a random word of $\{\mathtt{a},\mathtt{b}\}^n$, in order to form an $n+1$-letters long word, for Pascal's triangle,
- by addition of the image of $n+1$ to a random mapping from $[\![n]\!]$ to $[\![N]\!]$, in order to form a random mapping from $[\![n+1]\!]$ to $[\![N]\!]$, for the the second Stirling triangle,

- by random insertion of $n+1$ in order to form a permutation on $[\![n+1]\!]$, starting from a permutation on $[\![n]\!]$, for the 2 other examples.

In each case, the added letter, or integer, is chosen independently of the previous history of the growth process, hence the Markovian character of these growth processes. For the sake of brevity, in this article, we call the last growth process the *random permutation process*.

For these three nonhomogeneous Markov growth processes, there exist well studied functionals that retain the Markov property, and whose one-dimensional distributions are given by the rows of the corresponding combinatorial triangle :

- the sequence of counts of letter `a`, in the sequence of words defined previously, forms one of the most studied Markov chain : the simple random walk, whose one-dimensional distributions h_n are binomial distributions, famously related to Pascal's triangle ;

- the sequence of number of cycles, derived from the random permutation process, is an inhomogeneous Markov chain, related to the chinese restaurant process. Its one-dimensional distributions h_n have a simple expression in terms of the Stirling numbers of the first kind ;

- the sequence of sizes of images, derived from the sequence of random mappings, is a famous inhomogeneous Markov chain, related to the coupon collector problem : it is the sequence of successive sizes of the collection. Its one-dimensional distributions h_n have a simple expression in terms of the Stirling numbers of the second kind ;

- the sequence of the number of descents, also derived from the random permutation process, is an inhomogeneous Markov chain, related to the internal diffusion limited aggregation process. Its one-dimensional distributions h_n have a simple expression in terms of Eulerian numbers.

Chapman-Kolmogorov equations for these Markov chains are derived from Pascal's formulas for corresponding triangles through renormalization : in our settings, Q_n is defined by (a,b) as follows

$$Q_{n,x,y} = c_n(\theta)\ \left(b(n+1,y)\mathbb{1}_{y=x\in[\![n]\!]} + a(n+1,y)\ \theta\ \mathbb{1}_{y=x+1\in[\![n+1]\!]}\right), \tag{22}$$

in which $c_n(\theta)$ denotes a normalizing factor $T_n(\theta)/T_{n+1}(\theta)$, and $\theta = \theta^{x+1}/\theta^x$ should be replaced, in the last factor of (22), with $\theta - x = \theta^{x+1\downarrow}/\theta^{x\downarrow}$ in the case of Stirling numbers of the second kind. For Eulerian numbers, $\theta = 1$. Then, Pascal's formulas appear as special cases of the Chapman-Kolmogorov equation $h_{n-1}Q_{n-1} = h_n$, and relation (2) is just a special case of Proposition 11.

Here, since Y is a sequence of Bernoulli random variables, $Q_{n,x_n,x_{n+1}} \neq 0$ only if $\varepsilon_{n+1} = x_{n+1} - x_n$ belongs to $\{0,1\}$, thus $P_{(n,x),(n-1,y)} \neq 0$ only if $\varepsilon = x - y$ belongs to $\{0,1\}$: in this article, $P_{(n,x),(n-1,x-\varepsilon)}$ is abridged to $p_\varepsilon(n,x)$.

2.2 Simple random walk

Proof of Proposition 3. Here

$$T(n,k) = \binom{n}{k}, \quad \theta = \frac{p}{1-p}, \quad T_n(\theta) = (1+\theta)^n$$
$$h_n(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad c_n(\theta) = \frac{1}{1+\theta} = 1-p.$$

Then, for instance,

$$P_{(n,i),(n-1,i-1)} = \frac{h_{n-1}(i-1)Q_{n-1,i-1,i}}{h_n(i)}$$
$$= \frac{\binom{n-1}{i-1}p^{i-1}(1-p)^{n-i} \times p}{\binom{n}{i}p^i(1-p)^{n-i}} = \frac{\binom{n-1}{i-1}}{\binom{n}{i}} = p_1(n,i)$$

as expected. □

2.3 Chinese restaurant process

Set $\theta \in (0,+\infty)$. The chinese restaurant process, introduced in 1974 by Antoniak in [Ant74], is defined as follows : when entering a metaphoric chinese restaurant, the first customer seats at the first table. For $n > 1$, the nth customer seats at the kth (non-empty) table with probability $\frac{c_{n,k}}{n-1+\theta}$ (where $c_{n,k}$ is the number of customers seated at this table), or at an empty table with probability $\frac{\theta}{n-1+\theta}$. Let X_n denote the number of non-empty tables after the arrival of the nth customer. For exemple, let us sample the first 50 steps of the process, for $\theta = 1$:

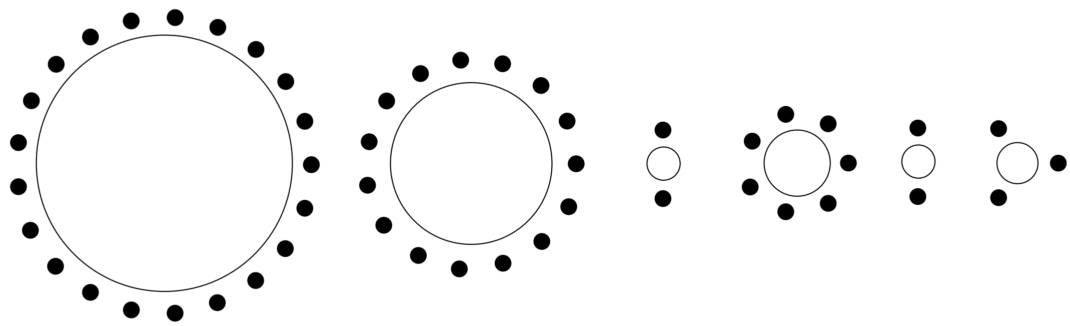

Figure 11: A realization of the chinese restaurant process (here $X_{50} = 6$).

Proof of Proposition 4. In this example, $(Y_i)_{i\geqslant 1}$ is a family of independent Bernoulli random variables with respective parameters $p_i = \theta/(i-1+\theta)$. We have :

$$T(n,k) = \begin{bmatrix} n \\ k \end{bmatrix}, \quad T_n(\theta) = \sum_k T(n,k)\theta^k = (\theta)^{\uparrow n}, \quad c_n(\theta) = \frac{1}{\theta+n}.$$

Thus the probability distribution of X_n is given, for $n \geqslant 1$, by:

$$h_n(\ell) = \mathbb{P}(X_n = \ell) = \frac{\theta^\ell}{(\theta)^{\uparrow n}} \begin{bmatrix} n \\ \ell \end{bmatrix} \mathbb{1}_{1 \leqslant \ell \leqslant n},$$

see [Pit06, Section 3.1.3]. For instance,

$$Q_{n,k,k+1} = \frac{\theta}{n+\theta} = c_n(\theta)\ a(n+1,k)\ \theta.$$

and

$$\begin{aligned} p_1(n,k) = P_{(n,k),(n-1,k-1)} &= \frac{h_{n-1}(k-1)Q_{n-1,k-1,k}}{h_n(k)} \\ &= \frac{\frac{\theta^{k-1}}{(\theta)^{n-1\uparrow}} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} \times \frac{\theta}{n-1+\theta}}{\begin{bmatrix} n \\ k \end{bmatrix} \frac{\theta^k}{(\theta)^{n\uparrow}}} = \frac{\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}}{\begin{bmatrix} n \\ k \end{bmatrix}}, \end{aligned}$$

as expected. □

2.4 Coupon collector's problem

Let us recall the famous problem studied by Gauss and Laplace, among others : a collector wants to complete a collection of N different items (denoted $1, ..., N$). At each step, he receives a coupon chosen uniformly from $[\![1, N]\!]$. The average time to complete the collection is known to be NH_N, where

$$H_N = \sum_{k=1}^{N} \frac{1}{k}$$

is the Nth harmonic number. If X_n denotes the number of different items in the collection after the nth step, then we call the graph of $t \mapsto X_{\lfloor t \rfloor}$ the *completion curve.*

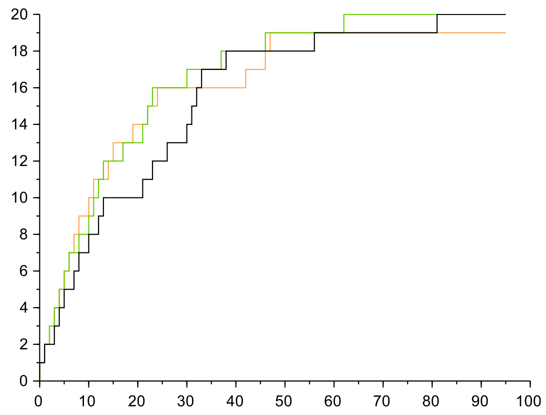

Figure 9: Three completion curves for a $n = 20$ items collection.

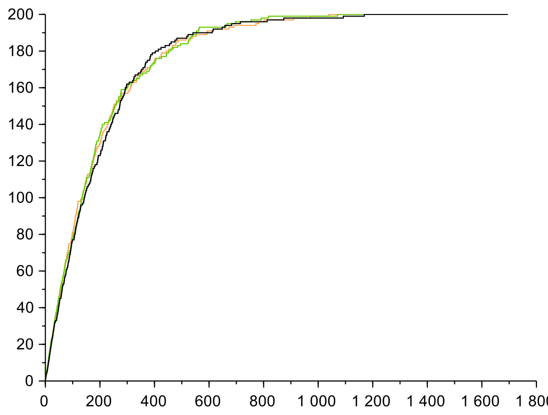

Figure 10: Three completion curves for a $n = 200$ items collection.

Proof of Proposition 5. See [AC19, Proposition 1], in which the proof is given for $m = N$. It fits with the frame given at Section 2.1 as follows : set

$$T(n,k) = \left\{ {n \atop k} \right\}, \quad \theta = N,$$

but consider a variant of T_n. Here :

$$T_n(\theta) = \sum_k T(n,k)\theta^{k\downarrow} = N^n$$

$$h_n(k) = \binom{N}{k} \left\{ {n \atop k} \right\} k! \, \frac{1}{N^n} = \left\{ {n \atop k} \right\} \frac{N^{k\downarrow}}{N^n}, \quad c_n(\theta) = \frac{1}{N}.$$

Then, for instance,

$$Q_{n,k,k+1} = a(n+1,k)\frac{\theta^{k+1\downarrow}}{\theta^{k\downarrow}} c_n(\theta) = \frac{N-k}{N}.$$

and

$$p_1(n,k) = P_{(n,k),(n-1,k-1)} = \frac{h_{n-1}(k-1)Q_{n-1,k-1,k}}{h_n(k)}$$

$$= \frac{\left\{ {n-1 \atop k-1} \right\} \frac{N^{k-1\downarrow}}{N^{n-1}} \times \frac{N-k+1}{N}}{\left\{ {n \atop k} \right\} \frac{N^{k\downarrow}}{N^n}} = \frac{\left\{ {n-1 \atop k-1} \right\}}{\left\{ {n \atop k} \right\}},$$

as expected. □

2.5 One-dimensional Internal Diffusion Limited Aggregation process

Diaconis and Fulton [DF91] introduced the internal Diffusion Limited Aggregation process (iDLA). Lawler, Bramson and Griffeath [LBG92] coined the terminology *iDLA*, and obtained an asymptotic shape behaviour. In the iDLA process, an aggregate of particles on $\mathbb{Z}^d$ is built as follows:

1. the first particle settles at the origin;
2. the next particles perform a symmetric random walk on $\mathbb{Z}^d$, starting from the origin, and settle at the first empty site they encounter.

When $d = 1$, let X_n denote the number of particles settled to the right of the origin after the nth step. Then, according to [Mit20], the process $(X_n)_n$ is an inhomogeneous Markov chain with the same distribution as the sequence of number of descents of the sequence of random permutations defined previously. Both processes have the one-dimensional distribution below

$$\mathbb{P}(X_n = k) = h_n(k) = \frac{\left\langle {n \atop k} \right\rangle}{n!} \mathbb{1}_{(n,k)\in S}.$$

In the case of the one-dimensional iDLA we can stack successive aggregates upon one another to form a space-time diagram. As with Figure 12, the longer it took to visit a cell, the darker we color it.

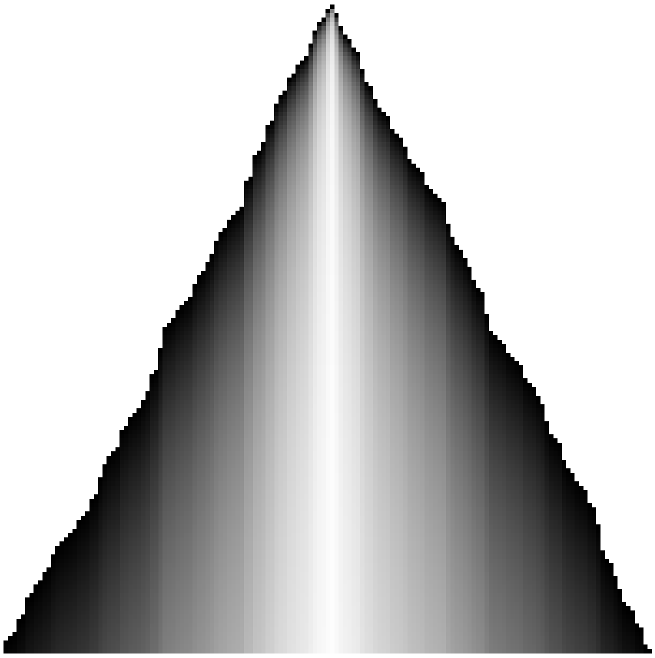

Figure 13: Space-time diagram of a one-dimensional iDLA.

Proof of Proposition 7. In this example, we have :

$$T(n,k) = \left\langle {n \atop k} \right\rangle, \quad T_n(1) = \sum_k T(n,k) = n!, \quad c_n(1) = \frac{1}{n}.$$

Thus the probability distribution of X_n is given, for $n \geqslant 1$, by:

$$h_n(\ell) = \mathbb{P}(X_n = \ell) = \frac{1}{n!} \left\langle {n \atop \ell} \right\rangle \mathbb{1}_{1 \leqslant \ell \leqslant n},$$

see [Pit06, Section 3.1.3]. For instance,

$$Q_{n,k,k+1} = \frac{n-k}{n+1} = c_n(1)\ a(n+1,k).$$

and

$$p_1(n,k) = P_{(n,k),(n-1,k-1)} = \frac{h_{n-1}(k-1) Q_{n-1,k-1,k}}{h_n(k)}$$
$$= \frac{\frac{1}{n-1!} \left\langle {n-1 \atop k-1} \right\rangle \times \frac{n-k}{n}}{\left\langle {n \atop k} \right\rangle \frac{1}{n!}} = \frac{\left\langle {n-1 \atop k-1} \right\rangle (n-k)}{\left\langle {n \atop k} \right\rangle},$$

as expected. □

3 The limit vector field : proof of Theorem 8

One can see the set v of average jumps $v(n,k)$, defined, for $(n,k) \in S$, by

$$\begin{aligned} v(n,k) &= p_0(n,k) \times (-1,0) + p_1(n,k) \times (-1,-1) \\ &= (-1, -p_1(n,k)), \end{aligned}$$

as a kind of discrete vector field v on S, with slope $p_1(n,k)$ at point (n,k). As a consequence, the convergence of the sample paths of the time-reversed Markov chains of Section 1.2 (see Theorem 9) should require a precise asymptotic analysis of

$$p_1(n,k) = \frac{a(n,k)T(n-1,k-1)}{T(n,k)},$$

and thus, of $T(n,k)$.

Consider the generating functions $\mathcal{V}_k$ and $\mathcal{H}_n$ defined by

$$\mathcal{V}_k(z) = \sum_{n=k}^{+\infty} \frac{1}{f_n} T(n,k) z^n, \quad \mathcal{H}_n(w) = \sum_{k=0}^{n} T(n,k) w^k,$$

respectively. Here f_n is either 1 or $n!$: since the two Stirling triangles enumerate sets of labelled objects (cycles and sets, see [FS09, examples II.3 and II.4, page 99]) the factor $n!$ is due to the use of EGFs. Also, since we consider *sets*, not sequences, of k objects, i.e. *unordered collections*, the generating function $\mathcal{V}_k$ contains a factor $1/k!$. For Pascal's triangle and for Euler's triangle, $f_n = 1$. Except for Euler's triangle, $\mathcal{V}_k$ exhibits a factorisation $A \times B^k$ suitable for the saddle-point method for large powers (see [FS09, Ch. VIII.8, p. 585]), while, for Eulerian numbers, $\mathcal{H}_n$ is approximately of the form B^n, allowing the use of large deviations methods.

Due to these factorisations, the limit vector field depends only on the slope

$$\frac{y}{x} = \frac{1}{1+\lambda},$$

and the function φ depends on B alone, in the first three cases through the saddle-point equation

$$\frac{B'(\zeta)}{B(\zeta)} = \frac{1+\lambda}{\zeta}, \tag{23}$$

obtained by optimisation of the function $x \to \frac{B(x)}{x^{1+\lambda}}$ on $(0,+\infty)$, see [FS09, Th. VIII.8, p. 587]. For Eulerian numbers, φ depends on B through the Legendre transformation of $\ln B$, leading to the equation :

$$\frac{B'(\zeta)}{B(\zeta)} = \frac{1}{1+\lambda}. \tag{24}$$

Thus, in the four cases, φ is a simple function of ζ, while ζ is an implicit function of the slope y/x.

Proof of Theorem 8. Except for Eulerian numbers, let $\zeta(\lambda)$ be defined implicitly by (23), i.e. let $\zeta(.)$ be the inverse function of :

$$x \longrightarrow \frac{xB'(x)}{B(x)} - 1.$$

The Eulerian case is similar, but uses large deviations rather than saddle-point methods, and will be handled separately. In the remaining three cases, recall that $a(n,k) = 1$, and set :

$$1 + \lambda = \frac{n}{k}, \quad \zeta = \zeta(\lambda), \quad 1 + \tilde{\lambda} = \frac{n-1}{k-1}, \quad \tilde{\zeta} = \zeta(\tilde{\lambda}).$$

For these three cases, the saddle-point method, see [FS09, Part B, Chap. VIII], leads to

$$T(n,k) \sim \frac{f_n}{f_k} \left(\frac{B(\zeta)}{\zeta^{1+\lambda}}\right)^k g(n,k), \tag{25}$$

in which $g(.,.)$ is some factor such that $g(n,k) \sim g(n-1,k-1)$. The invariance by homothetie of the field lines results from the factorisation $\mathcal{V}_k = A \times B^k$ and from the Cauchy formula, that leads to the key role of the slope $1+\lambda$ in the asymptotic behaviour (25), and is thus a consequence of the decomposability of the underlying combinatorial structures.

The factor f_n/f_k matters only for the two Stirling triangles. As a consequence, for the two Stirling triangles, we have

$$\begin{aligned} p_1(n,k) &\sim \frac{a(n,k)\ k}{n} \ \frac{\tilde{\zeta}^{1+\tilde{\lambda}}}{B\left(\tilde{\zeta}\right)} \left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\tilde{\lambda}}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right)^k \\ &\sim \frac{1}{1+\lambda} \ \frac{\zeta^{1+\lambda}}{B(\zeta)} \ \left(\frac{\tilde{\zeta}^{1+\lambda}}{\tilde{\zeta}^{1+\tilde{\lambda}}}\right)^k \left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right)^k \\ &\sim \frac{1}{1+\lambda} \ \frac{\zeta^{1+\lambda}}{B(\zeta)} \ \zeta^{(\lambda-\tilde{\lambda})k}, \end{aligned}$$

the last step due to

$$\lim_k \ k \ \ln \left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}} \frac{\zeta^{1+\lambda}}{B(\zeta)}\right) = 0. \tag{26}$$

Actually, since ζ is solution of the saddle-point equation, the derivative of

$$x \to \ln \left(\frac{B(x)}{x^{1+\lambda}}\right)$$

vanishes at ζ, thus

$$\ln\left(\frac{B\left(\tilde{\zeta}\right)}{\tilde{\zeta}^{1+\lambda}}\frac{\zeta^{1+\lambda}}{B\left(\zeta\right)}\right)=o\left(\tilde{\zeta}-\zeta\right)=o\left(\lambda-\tilde{\lambda}\right),$$

but

$$\lambda-\tilde{\lambda}=\frac{n}{k}-\frac{n-1}{k-1}\sim\frac{-\lambda}{k},$$

entailing (26). Thus

$$p_1(n,k)\sim\frac{1}{1+\lambda}\;\frac{\zeta}{B\left(\zeta\right)}.$$

Finally, for the two Stirling triangles, the saddle-point equation (23) gives

$$p_1(n,k)\sim\frac{1}{B'\left(\zeta\right)},$$

For Pascal's triangle, $\binom{n}{k}$ enumerates words with n letters, k among them being `a`'s and the $n-k$ others being `b`'s, thus Pascal's triangle enumerates *sequences* (not sets) of *unlabelled* objects[1], for which one usually uses OGFs. As a consequence, $f_n=1$, and, compared with the previous computation, we are rid of the factor $n!/k!$ in $T(n,k)$, and of the factor $k/n=1/(1+\lambda)$ in $p_1(n,k)$, thus we obtain

$$p_1(n,k)\sim\;\frac{\zeta}{B\left(\zeta\right)}.$$

Before we turn to the case of Eulerian numbers, let us derive φ for each of the 3 first cases :

- *Pascal's triangle :*

$$\begin{aligned}\mathcal{V}_{k,Pa}(z)&=\sum_{n\geqslant k}\binom{n}{k}z^n=\;\frac{1}{1-z}\;\left(\frac{z}{1-z}\right)^k,\\ B_{Pa}(z)&=\frac{z}{1-z},\\ p_1(n,k)&\sim\frac{\zeta}{B_{Pa}\left(\zeta\right)}=1-\zeta.\end{aligned}$$

Here (23) can be written

$$\frac{1}{1-\zeta}=1+\lambda,$$

[1] A word with n letters, k among them being `a`'s and the $n-k$ others being `b`'s, can be seen as a sequence of k words of the form $\texttt{b}^m\texttt{a}$ followed by a word of the form $\texttt{b}^m$.

thus $\varphi^{(Pa)}(k/n) = \frac{1}{1+\lambda} = k/n$, that is :

$$p_1(n,k) \sim \frac{k}{n},$$

which is not a surprise, since it is well known that, actually, $p_1(n,k) = \frac{k}{n}$.

- *Stirling numbers of the first kind (unsigned)*

$$\begin{aligned}
\mathcal{V}_{k,S1}(z) &= \sum_{n \geqslant k} \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!} = \frac{1}{k!} \left(-\ln(1-z)\right)^k, \\
B_{S1}(z) &= -\ln(1-z), \\
p_1(n,k) &\sim \frac{1}{B'_{S1}(\zeta_{S1})} = 1 - \zeta_{S1}.
\end{aligned}$$

Here (23) can be written

$$\frac{\zeta_{S1}}{(\zeta_{S1}-1)\ln(1-\zeta_{S1})} = 1 + \lambda, \tag{27}$$

which defines ζ_{S1} as smooth concave function of $\lambda > 0$, with values in $(0,1)$. Thus

$$\varphi_{S1}(k/n) = 1 - \zeta_{S1}\left(\tfrac{n}{k} - 1\right). \tag{28}$$

Note that (27) is the equation to be solved when one wants to tune the parameter ζ of a logarithmic probability distribution in order to obtain the expectation $1+\lambda$.

- *Stirling numbers of the second kind*

Here :

$$\begin{aligned}
\mathcal{V}_{k,S2}(z) &= \sum_{n \geqslant k} \begin{Bmatrix} n \\ k \end{Bmatrix} \frac{z^n}{n!} = \frac{1}{k!} \left(e^z - 1\right)^k, \\
B_{S2}(z) &= e^z - 1, \\
p_1(n,k) &\sim \frac{1}{B'_{S2}(\zeta_{S2})} = e^{-\zeta_{S2}},
\end{aligned}$$

and (23) can be written

$$\frac{\zeta_{S2}}{1 - e^{-\zeta_{S2}}} = 1 + \lambda, \tag{29}$$

see [AC19]. Thus

$$\varphi_{S2}(k/n) = e^{-\zeta_{S2}\left(\frac{n}{k}-1\right)}.$$

Note that, according to Good [Goo61] and others, ζ_{S2} is a smooth concave function of $\lambda > 0$, with positive values. Note also that (29) is the equation to be solved when one wants to tune the parameter ζ of a Poisson random variable *conditioned to be positive* in order to obtain the expectation $1+\lambda$.

For *Eulerian numbers*, though the computation of φ_E has a similar flavour, it presents some notable differences. In order to sum up the asymptotic analysis of Eulerian numbers, set, as done in [Ben73] :

$$t = \frac{k}{n} = \frac{1}{1+\lambda}.$$

In [Ben73, page 97], the main tool is the approximation of $\mathcal{H}_n(e^s)$, the Laplace transform of h_n, by

$$B_E(s)^{n+1} = r(s)^{-n-1} = \left(\frac{e^s - 1}{s}\right)^{n+1}.$$

In other terms, the key point in [Ben73] is that h_n is approximately the distribution of the sum of $n+1$ i.i.d. uniform random variables $(U_k)_{0 \leqslant k \leqslant n}$, each with Laplace transform $B_E(s)$. This is reminiscent of Tanny's representation of Eulerian numbers (cf. [Tan73]) :

$$h_n(k) = \frac{\left\langle {n \atop k} \right\rangle}{n!} = \mathbb{P}\left(\lfloor U_1 + U_2 + \cdots + U_n \rfloor = k\right). \tag{30}$$

Bender obtains the following asymptotic formula for $\left\langle {n \atop k} \right\rangle$ when (n, k) goes to infinity

$$\frac{\left\langle {n \atop k} \right\rangle}{n!} \sim \left(B_E(\zeta_E) e^{-\zeta_E t}\right)^n \; g(n, k)$$

in which $g(.,.)$ is some factor such that $g(n, k) \sim g(n-1, k-1)$, and in which ζ_E is the only real number such that

$$\begin{aligned} \frac{1}{1+\lambda} = t &= \frac{\partial}{\partial \zeta_E} \ln\left(\frac{e^{\zeta_E} - 1}{\zeta_E}\right) \\ &= \frac{e^{\zeta_E}}{e^{\zeta_E} - 1} - \frac{1}{\zeta_E}. \end{aligned} \tag{31}$$

One notices that ζ_E is the derivative of the Legendre-Fenchel transformation of the cumulant-generating function of the uniform distribution, i.e. the unique solution of

$$\frac{\partial}{\partial \zeta_E} \ln\left(B_E(\zeta_E) e^{-\zeta_E t}\right) = \frac{B'_E(\zeta_E)}{B_E(\zeta_E)} - t = 0. \tag{32}$$

As a consequence, for Eulerian numbers, we have

$$\begin{aligned} p_1(n,k) &\sim \frac{a(n,k)}{n}\ \left(\frac{B_E(\tilde{\zeta}_E)^{n-1}e^{-\tilde{\zeta}_E(k-1)}}{B_E(\zeta_E)^n e^{-\zeta_E k}}\right) \\ &\sim \frac{a(n,k)}{n}\ \frac{e^{\tilde{\zeta}_E}}{B_E(\tilde{\zeta}_E)}\left(\frac{B_E(\tilde{\zeta}_E)^{n}e^{-\tilde{\zeta}_E k}}{B_E(\zeta_E)^n e^{-\zeta_E k}}\right) \\ &\sim \frac{(1-t)e^{\zeta_E}}{B_E(\zeta_E)}\left(\frac{B_E(\tilde{\zeta}_E)e^{-\tilde{\zeta}_E t}}{B_E(\zeta_E)e^{-\zeta_E t}}\right)^n \\ &\sim \frac{(1-t)e^{\zeta_E}}{B_E(\zeta_E)} \end{aligned}$$

the last step due to

$$\lim_n\ n\ \ln\left(\frac{B_E(\tilde{\zeta}_E)e^{-\tilde{\zeta}_E t}}{B_E(\zeta_E)e^{-\zeta_E t}}\right) = 0. \tag{33}$$

Actually, since ζ_E is solution of (32), the derivative of

$$x \to \ln\left(B_E(x)e^{-xt}\right)$$

vanishes at ζ_E, thus

$$\ln\left(\frac{B_E(\tilde{\zeta}_E)e^{-\tilde{\zeta}_E t}}{B_E(\zeta_E)e^{-\zeta_E t}}\right) = o\left(\tilde{\zeta}_E - \zeta_E\right) = o\left(t - \tilde{t}\right),$$

but

$$t - \tilde{t} = \frac{k}{n} - \frac{k-1}{n-1} \sim \frac{1-t}{n},$$

entailing (33). Thus

$$p_1(n,k) \sim \frac{(1-t)\zeta_E e^{\zeta_E}}{e^{\zeta_E}-1} = \frac{\lambda\zeta_E}{(1+\lambda)(1-e^{-\zeta_E})} = \varphi_E(\lambda).$$

Note that :

$$\zeta_E(1-t) = -\zeta_E(t), \quad \varphi_E(1-t) = 1 - \varphi_E(t) = \frac{t\zeta_E}{e^{\zeta_E}-1},$$

as expected from the relation $\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n-k-1} \right\rangle$. □

4 Sample path convergence

This section is devoted to the proof of Theorem 9 for the first three triangles. For the sake of completeness, we first give the well known proof of Theorem 9 for Pascal's triangle.

In the case of Stirling numbers of the second kind, a weaker form of Theorem 9 was obtained in [AC19] at the price of a tedious proof using Wormald method and saddle-point asymptotics. For the two Stirling triangles, these alternative, and simpler, proofs are mere adaptations of the proof for Pascal's triangle: we can see the Markov chain $\mathcal{W}$ as a conditioned process with a parameter (p, θ or N) that we can chose arbitrarily. Choosing the parameter equal to its maximum likelihood estimator knowing the terminal point (m, ℓ) of the sample path[2], which suggests the choices

$$p_0 = \frac{\ell}{m}, \quad \theta_0 = m\,\frac{1-\zeta_{S1}(\lambda)}{\zeta_{S1}(\lambda)}, \quad N_0 = \left\lceil \frac{m}{\zeta_{S2}(\lambda)} \right\rceil,$$

makes the probability of the condition $X_m = \ell$ as large as possible, that is

$$\mathbb{P}(X_m = \ell) = \Theta(1/\sqrt{m}). \tag{34}$$

The *unconditional* large deviation probabilities being exponentially small, relation (34) entails that the *conditional* large deviation probabilities are exponentially small too. It follows that the application $t \to m^{-1}\mathbb{E}\left[X_{\lfloor mt \rfloor}\right]$ provides a good approximation for the solutions γ_λ of the ODEs.

For Euler's triangle, we think that Theorem 9 holds true too, but our proof is still incomplete. In the case of Stirling triangles of both kind, we believe that the proofs given in this section are new.

Proof of Theorem 9. Consider two probability distributions for the processes (W, X, Y) defined at section 1.2. Under $\mathbb{P}_{(m,\ell)}$, W is a Markov chain starting from (m, ℓ), distributed as $\mathcal{W}$ (with the same distribution as $\mathcal{W}^{(Pa)}$, $\mathcal{W}^{(S1)}$ or $\mathcal{W}^{(S2)}$, in the three cases we consider), and the processes (X, Y) are distributed accordingly. On the other hand, under $\mathbb{P}_\sigma$, (σ standing for p, θ or N in the three cases we consider) (W, X, Y) has the distribution induced by the simple random walk X with parameter p, the chinese restaurant process with parameter θ, or the coupon collector with parameter N, $Y = (Y_k)_{1 \leqslant k \leqslant m}$ being, in each of these cases, a sequence of Bernoulli random variables. According to Propositions 3, 4 and 5, for any σ, and any set B in the relevant state space,

$$\begin{aligned} \mathbb{P}_{(m,\ell)}\left((W, X, Y) \in B\right) &= \mathbb{P}_\sigma\left(\{(W, X, Y) \in B\} \cap \{X_m = \ell\}\right)/\mathbb{P}_\sigma\left(X_m = \ell\right) \\ &= \mathbb{P}_\sigma\left(\{(W, X, Y) \in B\} \cap \{X_m = \ell\}\right)/\mathbb{P}_\sigma\left(W_0 = (m, \ell)\right). \end{aligned} \tag{35}$$

Let $\|.\|_\infty$ denote the supremum norm over the interval $[0, 1]$, and set

$$\begin{aligned} A_m &= \{\exists n \in [\![0, m]\!] \text{ s.t. } \|W_n - (m-n, \mathbb{E}[X_{m-n}])\|_1 \geqslant m^{1-\eta}/\sqrt{2}\}, \\ &= \{\exists n \in [\![0, m]\!] \text{ s.t. } |X_n - \mathbb{E}[X_n]| \geqslant m^{1-\eta}/\sqrt{2}\}, \\ B_m &= \left\{\|w_m - \gamma_{m,\ell}\|_\infty \geqslant m^{1-\eta}\right\}. \end{aligned}$$

Then, according to (35), for any σ,

$$\mathbb{P}_{(m,\ell)}(A_m) = \frac{\mathbb{P}_\sigma(A_m \cap \{X_m = \ell\})}{\mathbb{P}_\sigma(X_m = \ell)} \leqslant \frac{\mathbb{P}_\sigma(A_m)}{\mathbb{P}_\sigma(X_m = \ell)}. \tag{36}$$

[2]Recall that the terminal point (m, X_m) is a sufficient statistic.

In each of the three cases, $B_m \subset A_m$ for m large enough. On the other hand $\mathbb{P}_\sigma(A_m)$ decays exponentially for any σ : actually, we have, for all $t > 0$ and all $n \in [\![1, m]\!]$,

$$\mathbb{P}_\sigma(|X_n - \mathbb{E}[X_n]| \geqslant t) \leqslant 2 \exp\left(-\frac{2t^2}{n}\right) \leqslant 2 \exp\left(-\frac{2t^2}{m}\right). \tag{37}$$

In particular, for any $\eta \in (0, 1/2)$ and for $t = m^{1-\eta}/\sqrt{2}$,

$$\mathbb{P}_\sigma(|X_n - \mathbb{E}[X_n]| \geqslant m^{1-\eta}/\sqrt{2}) \leqslant 2 \exp\left(-m^{1-2\eta}\right).$$

Thus

$$\mathbb{P}_\sigma\left(A_m\right) \leqslant \sum_{n=1}^{m} \mathbb{P}_\sigma(|X_n - \mathbb{E}[X_n]| \geqslant {\scriptstyle m^{1-\eta}/\sqrt{2}}) \leqslant 2m \exp\left(-m^{1-2\eta}\right). \tag{38}$$

But, for a suitable choice σ_0 of σ,

$$\mathbb{P}_{\sigma_0}(X_m = \ell) = \Theta\left(\frac{1}{\sqrt{m}}\right). \tag{39}$$

Finally, in each case, $\mathbb{P}_{(m,\ell)}\left(B_m\right) = \mathcal{O}(m^{3/2} e^{-m^{1-2\eta}})$ is a direct consequence of (36), (38) and (39), and Theorem 9 follows.

Proof of Theorem 9 : Pascal's triangle. Since Y is a sequence of independent Bernoulli random variables with parameter $\sigma = p \in (0, 1)$, Hoeffding's inequality (37) holds true for all $t > 0$ and all $n \in [\![1, m]\!]$, and $\mathbb{P}_p(A_m)$ decays exponentially for any $p \in (0, 1)$ and any $\eta \in (0, 1/2)$, due to (38). But (36) holds for $\sigma_0 = \tilde{p} = \ell/m$ too. For this choice of $\sigma_0 = \tilde{p}$, using Stirling formula, one finds

$$\begin{aligned}\mathbb{P}_{\tilde{p}}(X_m = \ell) &= \binom{m}{\ell} \left(\frac{\ell}{m}\right)^{\ell} \left(\frac{m-\ell}{m}\right)^{m-\ell} \\ &\sim \frac{1}{\sqrt{2\pi \tilde{p}(1-\tilde{p})}} \, \frac{1}{\sqrt{m}}.\end{aligned}$$

For Pascal's triangle, recall that $\gamma_{m,\ell}(t) = \ell t/m$, thus, for $mt \in [\![0, m]\!]$, we have $\gamma_{m,\ell}(t) = m^{-1} \mathbb{E}_{\tilde{p}}[X_{mt}]$ and, as a consequence,

$$A_m = \left\{ \sup \left\{ \left| w_m(t) - \gamma_{m,\ell}(t) \right|, mt \in [\![0, m]\!] \right\} \geqslant m^{-\eta}/\sqrt{2} \right\}.$$

Now

$$0 \leqslant \sup_{t \in [0,1]} \left\{ \left| w_m(t) - \gamma_{m,\ell}(t) \right| \right\} - \sup_{mt \in [\![0,m]\!]} \left\{ \left| w_m(t) - \gamma_{m,\ell}(t) \right| \right\} \leqslant \frac{\ell}{m^2} \leqslant \frac{1}{m},$$

so that, for m large enough, $B_m \subset A_m$, and $\mathbb{P}_{(m,\ell)}\left(B_m\right) = \mathcal{O}(m^{3/2} e^{-m^{1-2\eta}})$. Since

$$\left\| \gamma_{m,\ell} - \gamma_{\lambda_\infty} \right\|_\infty \leqslant \left| \lambda(m, \ell) - \lambda_\infty \right|,$$

it follows that

$$\mathbb{P}_{(m,\ell)}\left(\|w_m - \gamma_{\lambda_\infty}\|_\infty \geqslant m^{-\eta} + |\lambda(m,\ell) - \lambda_\infty|\right) = \mathcal{O}(m^{3/2} e^{-m^{1-2\eta}}).$$

Proof of Theorem 9 : Stirling numbers of the first kind. Here the parameter is traditionally denoted θ, rather than σ, and under $\mathbb{P}_\theta$, $(Y_k)_{k\geqslant 1}$ is a family of independent Bernoulli random variables with respective parameters

$$p_k = \frac{\theta}{k-1+\theta}.$$

Thus $X = (X_n)$, that can be seen as the number of non-empty tables after the arrival of the nth customer, as explained in Section 2.3, satisfies (36) and (37), due to Hoeffding's inequality. Note that, under $\mathbb{P}_\theta$,

$$\mathbb{E}_\theta[X_n] = \sum_{k=1}^{n} \frac{\theta}{k-1+\theta}.$$

Also, as in (27), recall that for the choice

$$\gamma_{m,\ell}(t) = \frac{1-\zeta}{\zeta} \ln\left(\frac{1-\zeta+t\,\zeta}{1-\zeta}\right).$$

in which

$$\frac{\zeta}{(\zeta-1)\ln(1-\zeta)} = \frac{m}{\ell} = 1+\lambda,$$

we have $\gamma_{m,\ell}(1) = {}^{1}/{}_{1+\lambda} = \ell/m$. But, for the choice $\theta_0 = m(1-\zeta)/\zeta$,

$$\left|\mathbb{E}_{\theta_0}\left[\frac{X_n}{m}\right] - \gamma_{m,\ell}(n/m)\right| \leqslant \frac{n^2}{m^3}\frac{\zeta}{2(1-\zeta)} \leqslant \frac{\zeta}{m(1-\zeta)}. \tag{40}$$

Since $\gamma_{m,\ell}$ is a contraction,

$$0 \leqslant \|w_m - \gamma_{m,\ell}\|_\infty - \sup_{mt\in[\![0,m]\!]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} \leqslant \frac{1}{m},$$

thus $B_m \subset A_m$ provided that

$$\frac{m^{-\eta}}{\sqrt{2}} + \frac{\zeta}{m(1-\zeta)} + \frac{1}{m} \leqslant m^{-\eta},$$

that is, $B_m \subset A_m$ for m large enough. Relation (36) holds true for any $\sigma = \theta > 0$, thus it holds for $\sigma = \theta_0 = (1-\zeta)m/\zeta$ too, but, using relation (13) in [Goo61], one finds

$$\begin{aligned}\mathbb{P}_{\theta_0}(X_m = \ell) &= \frac{\theta_0^\ell}{(\theta_0)^{\uparrow m}} \begin{bmatrix} m \\ \ell \end{bmatrix} \mathbb{1}_{1\leqslant\ell\leqslant m}, \\ &\sim \frac{1}{\sqrt{m}} \sqrt{\frac{\ln(1-\zeta)}{2\pi(1+\lambda)(\zeta+\ln(1-\zeta))}}.\end{aligned} \tag{41}$$

Due to (36), $\mathbb{P}_{(m,\ell)}(A_m) = \mathcal{O}(m^{3/2}e^{-m^{1-2\eta}})$ and decays exponentially for $\eta \in (0, 1/2)$ so that, as expected,

$$\lim_m \mathbb{P}_{(m,\ell)}\left(B_m\right) = 0.$$

Finally, $\|\gamma_\lambda - \gamma_\mu\|_\infty = |\gamma_\lambda(1) - \gamma_\mu(1)| \leqslant |\lambda - \mu|$, thus

$$\mathbb{P}_{(m,\ell)}\left(\|w_m - \gamma_{\lambda_\infty}\|_\infty \geqslant m^{-\eta} + |\lambda(m,\ell) - \lambda_\infty|\right) = \mathcal{O}(m^{3/2}e^{-m^{1-2\eta}}).$$

Proof of Theorem 9 : Stirling numbers of the second kind. Here the parameter is traditionally denoted by N, rather than σ : under $\mathbb{P}_N$, X_n is the number of different coupons that have been collected after n draws with replacement in a collection of N available coupons, and the processes (X, Y, W) are distributed accordingly. Due to [MR95, Ch. 4, Theorem 4.18 and pages 92-95], by Azuma-Hoeffding's inequality, relations (37) and (38) hold true[3]. Relation (36) holds true for any $\sigma = N \geqslant \ell$, thus for $N_0 = \lceil m/\zeta \rceil$ too, but, using relation (3) in [Goo61], one finds that

$$\mathbb{P}_{\lceil m/\zeta \rceil}(X_m = \ell) = \frac{N_0!\, N_0^{-m}}{N_0 - \ell!} \left\{ \begin{matrix} m \\ \ell \end{matrix} \right\}$$

so that

$$\frac{1}{\mathbb{P}_{\lceil m/\zeta \rceil}(X_m = \ell)} = \mathcal{O}(\sqrt{m}). \tag{42}$$

Thus, $\mathbb{P}_{(m,\ell)}(A_m) = \mathcal{O}(m^{3/2}e^{-\zeta m^{1-2\eta}})$ and decays exponentially for $\eta \in (0, 1/2)$. Finally let us check that $B_m \subset A_m$ for m large enough. Note that, under $\mathbb{P}_N$,

$$\mathbb{E}_N[X_n] = N\left(1 - \left(1 - \frac{1}{N}\right)^n\right),$$

so that

$$\left|\mathbb{E}_N[X_n] - N\left(1 - e^{-\frac{n}{N}}\right)\right| \leqslant \frac{n}{2N}. \tag{43}$$

Also, as in (29), set

$$\gamma_{m,\ell}(t) = \left(1 - e^{-\zeta t}\right)/\zeta$$

in which

$$\frac{\zeta}{1 - e^{-\zeta}} = 1 + \lambda = \frac{m}{\ell},$$

and set

$$\frac{m}{\tilde{\zeta}} = \left\lceil \frac{m}{\zeta} \right\rceil = N_0, \quad \text{so that} \quad 0 \leqslant \zeta - \tilde{\zeta} \leqslant \frac{\zeta\tilde{\zeta}}{m}.$$

[3] Note that Motwani and Raghavan provide a slightly sharper bound

Actually, due to (43), for $0 \leqslant n \leqslant m$,

$$\begin{aligned}\left|\frac{\mathbb{E}_{N_0}[X_n]}{m} - \gamma_{m,\ell}(n/m)\right| &\leqslant \frac{n}{2N_0 m} + \left|\frac{N_0}{m}\left(1 - e^{-\frac{n}{N_0}}\right) - \gamma_{m,\ell}(n/m)\right| \\ &\leqslant \frac{n}{2N_0 m} + \left|\frac{1 - e^{-\tilde{\zeta} n/m}}{\tilde{\zeta}} - \frac{1 - e^{-\zeta n/m}}{\zeta}\right| \\ &\leqslant \frac{\tilde{\zeta}}{2m} + \frac{1 + \tilde{\zeta}}{m}, \qquad (44)\end{aligned}$$

Thus $B_m \subset A_m$ for m large enough, i.e. provided that

$$\frac{m^{-\eta}}{\sqrt{2}} + \frac{\tilde{\zeta}}{2m} + \frac{1 + \tilde{\zeta}}{m} \leqslant m^{-\eta}.$$

Finally

$$0 \leqslant \sup_{t \in [0,1]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} - \sup_{mt \in [\![0,m]\!]} \{|w_m(t) - \gamma_{m,\ell}(t)|\} \leqslant \frac{1}{m},$$

so that, for m large enough,

$$\left\{\sup_{t \in [0,1]} |w_m(t) - \gamma_{m,\ell}(t)| \geqslant m^{-\eta}\right\} \subset A_m, \qquad (45)$$

and, as expected,

$$\lim_m \mathbb{P}_{(m,\ell)}\left(\sup_{t \in [0,1]} |w_m(t) - \gamma_{m,\ell}(t)| \geqslant m^{-\eta}\right) = 0.$$

Again $\|\gamma_\lambda - \gamma_\mu\|_\infty = |\gamma_\lambda(1) - \gamma_\mu(1)| \leqslant |\lambda - \mu|$, thus

$$\mathbb{P}_{(m,\ell)}\left(\|w_m - \gamma_{\lambda_\infty}\|_\infty \geqslant m^{-\eta} + |\lambda(m,\ell) - \lambda_\infty|\right) = \mathcal{O}(m^{3/2} e^{-m^{1-2\eta}}).$$

□

Acknowledgements and remarks.

The authors are grateful to Antoine Lejay, Valentin Feray and Rémi Peyre for fruitful discussions, and to Antoine Lejay for reference [Kol35]. A referee made many helpful suggestions for revisions of the manuscript. As concerns Euler's triangle, Theorem 9 is, as far as we know, an open question. However, we believe that a proof similar to those given in Section 4 is within reach : we are aware of a Markov chain X with parameter, such that X_n is a sufficient statistic. This leads us to the solution of ODE (15), see (14), but bottlenecks remain.

References

[AC19] Anis Amri and Philippe Chassaing, *The impatient collector*, working paper or preprint, June 2019.

[Ant74] Charles Edward Antoniak, *Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems*, Ann. Statist. **2** (1974), 1152–1174.

[Asm03] Sören Asmussen, *Applied probability and queues*, second ed., Applications of Mathematics (New York), vol. 51, Springer-Verlag, New York, 2003, Stochastic Modelling and Applied Probability.

[Ben73] Edward A. Bender, *Central and local limit theorems applied to asymptotic enumeration*, J. Combinatorial Theory Ser. A **15** (1973), 91–111.

[BN06] Frédérique Bassino and Cyril Nicaud, *Accessible and deterministic automata: enumeration and Boltzmann samplers*, Fourth colloquium on mathematics and computer science IV. Algorithms, trees, combinatorics and probabilities., Discrete Mathematics & Theoretical Computer Science (DMTCS), 2006, pp. 151–160 (English).

[CP05] Jean-Marc Champarnaud and Thomas Paranthoën, *Random generation of DFAs*, Theor. Comput. Sci. **330** (2005), no. 2, 221–235 (English).

[CW05] Kai Lai Chung and John B. Walsh, *Markov processes, Brownian motion, and time symmetry.*, 2nd ed. ed., Grundlehren Math. Wiss., vol. 249, New York, NY: Springer, 2005 (English).

[DF91] P. Diaconis and W. Fulton, *A growth model, a game, an algebra, Lagrange inversion, and characteristic classes*, vol. 49, 1991, Commutative algebra and algebraic geometry, II (Italian) (Turin, 1990), pp. 95–119 (1993).

[FS09] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.

[Goo61] I. J. Good, *An asymptotic formula for the differences of the powers at zero*, Ann. Math. Statist. **32** (1961), 249–256.

[Kol35] A. Kolmogoroff, *Zur Theorie der Markoffschen Ketten*, Math. Ann. **112** (1935), 155–160 (German).

[Kor78] A. D. Koršunov, *Enumeration of finite automata*, Problemy Kibernet. (1978), no. 34, 5–82, 272.

[Kor86] A. D. Koršunov, *On the number of nonisomorphic strongly connected finite automata*, Elektron. Informationsverarb. Kybernet. **22** (1986), no. 9, 459–462.

[LBG92] Gregory F. Lawler, Maury Bramson, and David Griffeath, *Internal diffusion limited aggregation*, Ann. Probab. **20** (1992), no. 4, 2117–2140 (English).

[Lot97] M. Lothaire, *Combinatorics on words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997.

[Mit20] Kiana Mittelstaedt, *A stochastic approach to Eulerian numbers*, Am. Math. Mon. **127** (2020), no. 7, 618–628 (English).

[MR95] Rajeev Motwani and Prabhakar Raghavan, *Randomized algorithms*, Cambridge: Cambridge Univ. Press, 1995 (English).

[Pit06] J. Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[Sti73] Stephen M. Stigler, *Studies in the history of probability and statistics. XXXII: Laplace, Fisher, and the discovery of the concept of sufficiency*, Biometrika **60** (1973), 439–445 (English).

[Tan73] S. Tanny, *A probabilistic interpretation of Eulerian numbers*, Duke Math. J. **40** (1973), 717–722 (English).

[War19] Lutz Warnke, *On Wormald's differential equation method*, 2019.

5 Appendix : proofs and computations.

Note that in the 4 cases, $\gamma_{m,\ell}$ inherits from X the property of being an increasing contraction. Also, in Theorem 9, the assumption that $(m, \ell) \to +\infty$ and $\lim \lambda(m, \ell) = \lambda_\infty \in (0, +\infty)$ entails that $\zeta = \zeta(\lambda(m, \ell))$ and $\tilde{\zeta}$ are ultimately bounded away from 0 and $+\infty$. For instance, this is needed for relation (45), that holds true when

$$m \geqslant \left(\frac{4 + 3\tilde{\zeta}}{2 - \sqrt{2}} \right)^{1/1-\eta}.$$

Proof of (7).

The relation

$$y' = \varphi_{S1}\left(\frac{x-y}{y}\right)$$

is equivalent to

$$1 - y' = \zeta_{S1}\left(\frac{x-y}{y}\right)$$

or, by (5), to

$$\frac{1-y'}{y'\ \ln(y')} = -\frac{x}{y}. \tag{46}$$

Then, for $x \geqslant 0$, and the choice

$$y = a\ln\left(1+\frac{x}{a}\right), \quad y' = \frac{a}{a+x}$$

relation (46) holds, for it can be written

$$\frac{x}{a\ln(a/a+x)} = -\frac{x}{y}.$$

Then a is chosen in such a way that

$$\frac{1}{1+\lambda} \quad = \gamma_\lambda(1) = a\ln\left(1+\tfrac{1}{a}\right),$$

that is, $a = \frac{1-\zeta_{S1}(\lambda)}{\zeta_{S1}(\lambda)}$, leading to

$$\frac{1}{1+\lambda} \quad = \gamma_\lambda(1) = \tfrac{1-\zeta_{S1}(\lambda)}{\zeta_{S1}(\lambda)}\ln\left(1+\tfrac{\zeta_{S1}(\lambda)}{1-\zeta_{S1}(\lambda)}\right) = \tfrac{\zeta_{S1}(\lambda)-1}{\zeta_{S1}(\lambda)}\ln\left(1-\zeta_{S1}(\lambda)\right).$$

Proof of (11)**.**

The relation

$$y' = \varphi_{S2}\left(\frac{x-y}{y}\right)$$

is equivalent to

$$-\ln(y') = \zeta_{S2}\left(\frac{x-y}{y}\right)$$

or, by (9), to

$$\frac{-\ln(y')}{1-y'} = \frac{x}{y}. \tag{47}$$

Then, for $x \geqslant 0$, and the choice

$$y = \frac{1-e^{-ax}}{a}, \quad y' = e^{-ax},$$

relation (47) holds, for it can be written

$$\frac{ax}{1-e^{-ax}} = \frac{x}{y}.$$

Then a is chosen in such a way that

$$\frac{1}{1+\lambda} \quad = \gamma_\lambda(1) = \tfrac{1-e^a}{a},$$

that is, $a = \zeta_{S2}(\lambda)$.

Proof of (40)**.**

This is a classic bound for Riemann sums.

Proof of (41)**.**

According to [Goo61, Section4],

$$\begin{bmatrix} m \\ \ell \end{bmatrix} \sim \frac{m!}{\ell!}\,\frac{(\ln(1-\zeta))^{\ell}}{\zeta^m\sqrt{2\pi\kappa\ell}},$$

in which ζ is defined by (5), and

$$\kappa = -(1+\lambda)^2\,\frac{\zeta+\ln(1-\zeta)}{\zeta}.$$

Set $\alpha = (1-\zeta)/\zeta$ and $\zeta = 1/(1+\alpha)$, so that $\theta_0 = \alpha m$. For $1 \leqslant \ell \leqslant m$,

$$\begin{aligned}\mathbb{P}_{\theta_0}(X_m = \ell) &= \frac{\theta_0^{\ell}}{(\theta_0)^{\uparrow m}}\begin{bmatrix} m \\ \ell \end{bmatrix} \\ &\sim \frac{m!}{\ell!}\,\frac{(-\alpha m\ln(1-\zeta))^{\ell}}{\zeta^m\sqrt{2\pi\kappa\ell}}\,\frac{\Gamma(m(\alpha+1))}{\Gamma(\alpha m)}\end{aligned}$$

and, applying Stirling formula four times,

$$\begin{aligned}\mathbb{P}_{\theta_0}(X_m = \ell) &\sim \left(\frac{-m\alpha e\ln(1-\zeta)}{\ell}\right)^{\ell}\left(\frac{\alpha^{\alpha}}{\zeta(\alpha+1)^{1+\alpha}}\right)^{m}\sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}} \\ &\sim \left(\frac{-(1+\lambda)e\ln(1-\zeta)\alpha^{1+\alpha(1+\lambda)}}{\zeta^{1+\lambda}(\alpha+1)^{(1+\alpha)(1+\lambda)}}\right)^{\ell}\sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}} \\ &\sim \left(\frac{-e(1+\lambda)\,(1-\zeta)^{(1+\alpha)(1+\lambda)}\ln(1-\zeta)}{\alpha^{\lambda}\zeta^{1+\lambda}}\right)^{\ell}\sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}},\end{aligned}$$

for $\alpha/(1+\alpha) = 1-\zeta$. Using $(1+\alpha)(1+\lambda) = 1/((\zeta-1)\ln(1-\zeta))$, one obtains

$$\begin{aligned}\mathbb{P}_{\theta_0}(X_m = \ell) &\sim \left(\frac{-e^{\zeta/(\zeta-1)}(1+\lambda)\,\ln(1-\zeta)}{\alpha^{\lambda}\zeta^{1+\lambda}}\right)^{\ell}\sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}} \\ &\sim \left(-\alpha(1+\lambda)\,\ln(1-\zeta)\right)^{\ell}\sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}} \\ &\sim \sqrt{\frac{1+\alpha}{2\pi\alpha\kappa m}} \\ &\sim \frac{1}{\sqrt{m}}\sqrt{\frac{\ln(1-\zeta)}{2\pi(1+\lambda)(\zeta+\ln(1-\zeta))}}.\end{aligned}$$

Proof of (42)**.**

Recall that, according to [Goo61, relation (3)],

$$\left\{ {m \atop \ell} \right\} \sim \frac{1}{2\pi} \, \frac{m!}{\ell!} \, \left(\frac{e^{\zeta} - 1}{\zeta^{1+\lambda}} \right)^{\ell} \, \sqrt{\frac{\pi}{v\ell}}, \tag{48}$$

in which ζ is defined by (9), and

$$v = \frac{(\lambda+1)(\zeta-\lambda)}{2}.$$

Using (48), we prove that, for the choice $N = \lceil m/\zeta \rceil$,

$$\mathbb{P}_{\lceil m/\zeta \rceil}(X_m = \ell) = \frac{N! \, N^{-m}}{N-\ell!} \left\{ {m \atop \ell} \right\}$$

satisfies (42) :

$$\frac{1}{\mathbb{P}_{\lceil m/\zeta \rceil}(X_m = \ell)} = \mathcal{O}(\sqrt{m}).$$

First, set $\varepsilon_m = \lceil {}^{m}/_{\zeta} \rceil - {}^{m}/_{\zeta} \in [0,1)$. Then, with the help of

$$\frac{1}{\zeta} - \frac{1}{1+\lambda} = \frac{e^{-\zeta}}{\zeta},$$

using Stirling formula twice,

$$\begin{aligned}
\frac{N-\ell! \, N^m}{N!} &\sim e^{\ell} \, \frac{(N-\ell)^{N-\ell}}{N^{N-m}} \sqrt{\frac{N-\ell}{N}} \\
&\sim e^{\frac{m}{1+\lambda}} \, \frac{m^{m(\frac{1}{\zeta}-\frac{1}{1+\lambda})} (\frac{1}{\zeta} - \frac{1}{1+\lambda})^{m(\frac{1}{\zeta}-\frac{1}{1+\lambda})}}{m^{m(\frac{1}{\zeta}-1)} \, \zeta^{m(1-\frac{1}{\zeta})}} \sqrt{1 - \frac{\zeta}{1+\lambda}} \; \psi(m,\lambda), \\
&\sim \frac{\left(e \left(\zeta e^{\zeta} \right)^{\frac{1}{1-e^{\zeta}}} \right)^{\frac{m}{1+\lambda}}}{\sqrt{\zeta e^{\zeta}}} \, \frac{m^{m(1-\frac{1}{1+\lambda})} \sqrt{\zeta}}{\zeta^{m(1-\frac{1}{\zeta})}} \; \psi(m,\lambda),
\end{aligned}$$

in which $\psi(m,\lambda)$ is defined by

$$\begin{aligned}
\psi(m,\lambda) &= e^{-\zeta\varepsilon_m} \left(1 + \frac{\zeta\varepsilon_m}{m} \right)^{m(1-\frac{1}{\zeta})-\varepsilon_m-1/2} \left(1 + \frac{\zeta\varepsilon_m}{me^{-\zeta}} \right)^{\frac{me^{-\zeta}}{\zeta}+\varepsilon_m+1/2} \\
&\sim e^{-\zeta\varepsilon_m} \left(1 + \frac{\zeta\varepsilon_m}{m} \right)^{m(1-\frac{1}{\zeta})} \left(1 + \frac{\zeta\varepsilon_m}{me^{-\zeta}} \right)^{\frac{me^{-\zeta}}{\zeta}},
\end{aligned}$$

so that $\psi(m,\lambda)$ is bounded away from 0 and $+\infty$, since $\varepsilon_m \in [0,1)$. Finally, using Stirling formula twice again,

$$\begin{aligned}
\frac{N-\ell!\, N^m}{N!\,\left\{{m \atop \ell}\right\}\psi(m,\lambda)} &\sim \frac{2\left(e^{1+\lambda}\left(\zeta e^{\zeta}\right)^{\frac{1}{1-e^{\zeta}}}\ \frac{\zeta^{1+\lambda}}{(1+\lambda)(e^{\zeta}-1)}\right)^{\frac{m}{1+\lambda}}}{\sqrt{(1+\lambda)e^{\zeta}}}\ \frac{\sqrt{\pi v\ell}}{\zeta^{m(1-\frac{1}{\zeta})}} \\
&\sim 2\left(e^{1+\lambda}\left(\zeta e^{\zeta}\right)^{\frac{1}{1-e^{\zeta}}}\ \zeta^{\lambda} e^{-\zeta}\ \zeta^{\frac{(1-\zeta)(1+\lambda)}{\zeta}}\right)^{\frac{m}{1+\lambda}} \sqrt{\frac{\pi v\ell}{(1+\lambda)e^{\zeta}}} \\
&\sim 2\left(e^{0}\zeta^{0}\right)^{\frac{m}{1+\lambda}} \sqrt{\frac{\pi v\ell}{(1+\lambda)e^{\zeta}}} \\
&\sim \frac{2\sqrt{\pi v e^{-\zeta}}}{1+\lambda}\ \sqrt{m},
\end{aligned}$$

so that (42) is satisfied.

Proof of (43)**.**

$$\begin{aligned}
\left|N\left(1-\left(1-\frac{1}{N}\right)^{n}\right)-N\left(1-e^{-\frac{n}{N}}\right)\right| &= N\left|\left(1-\frac{1}{N}\right)^{n}-e^{-\frac{n}{N}}\right| \\
&\leqslant Nn\left|\left(1-\frac{1}{N}\right)-e^{-\frac{1}{N}}\right| \\
&\leqslant Nn\times\frac{1}{2N^2}=\frac{n}{2N},
\end{aligned}$$

the last inequality due to relation :

$$\forall u\geqslant 0,\quad \left|e^{-u}-1+u\right|\leqslant\frac{u^2}{2}.$$

Proof of (44)**.**

This amounts to prove that, for $0\leqslant x\leqslant 1$,

$$A(x)=\left|\frac{1-e^{-\tilde{\zeta}x}}{\tilde{\zeta}}-\frac{1-e^{-\zeta x}}{\zeta}\right|\leqslant\frac{1+\tilde{\zeta}}{m},$$

But

$$0\leqslant\frac{m}{\tilde{\zeta}}-\frac{m}{\zeta}\leqslant 1,$$

thus

$$
\begin{aligned}
A(x) &\leqslant \left|\frac{1-e^{-\tilde{\zeta}x}}{\tilde{\zeta}} - \frac{1-e^{-\tilde{\zeta}x}}{\zeta}\right| + \left|\frac{1-e^{-\tilde{\zeta}x}}{\zeta} - \frac{1-e^{-\zeta x}}{\zeta}\right| \\
&\leqslant \frac{1}{m} + \frac{1}{\zeta}\left|e^{-\tilde{\zeta}x} - e^{-\zeta x}\right| \\
&\leqslant \frac{1}{m} + \frac{1}{\zeta}\left|\tilde{\zeta} - \zeta\right| \leqslant \frac{1+\tilde{\zeta}}{m}.
\end{aligned}
$$