



**HAL**  
open science

## Occupancies within zeta urns revisited

Thierry Huillet

► **To cite this version:**

Thierry Huillet. Occupancies within zeta urns revisited. The International Mathematical Forum, 2008, 3 (no 26), pp.1269 - 1293. hal-00096631

**HAL Id: hal-00096631**

**<https://hal.science/hal-00096631>**

Submitted on 19 Sep 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Occupancies within zeta urns revisited

Thierry HUILLET

Laboratoire de Physique Théorique et Modélisation,  
CNRS-UMR 8089 et Université de Cergy-Pontoise,  
2 Avenue Adolphe Chauvin, 95302, Cergy-Pontoise, FRANCE  
E-mail: Thierry.Huillet@u-cergy.fr

July 4, 2006

## Abstract

Equilibrium statistical properties of occupancy distributions within zeta urns are revisited and developed.

**Keywords:** Bose-Einstein statistics, balls in boxes, zeta-urns, occupancy distributions, sampling, equilibrium statistical mechanics.

AMS Classification 2000: 60C05, 60E05, 05Axx, 82Bxx, 60-02.

## 1 Introduction and outline of main results

Ten years ago, motivated by simplicial quantum gravity and the statistics of branched polymers, a model nicknamed the zeta-urn model was introduced (see [1], [2] and [3]). This model consists in a symmetric balls in boxes model with many interesting features both at and out of equilibrium (see [8]). The purpose of this work is to further investigate the statistical equilibrium properties of the zeta urn model. It is first underlined that the model is in the Bose-Einstein statistics class where undistinguishable particles are allocated within distinguishable boxes, the energy required to place particles within each box being independent of box label. As a result, it may be seen as a random allocation scheme obtained while conditioning on its sum a vector of infinitely divisible zeta-distributed discrete random variables. This class of models has recently received considerable attention (see [11] and [4], for instance).

One of the specificities of this model is that it presents a phase transition between a fluid and a condensed phase at finite inverse-temperature  $\beta > 1$ . This is because box energy is sub-linear (actually logarithmic). The critical properties are governed by the Riemann zeta function  $\zeta(\beta)$ . We recall this property and investigate some further statistical consequences in some details. Then, we

focus on the canonical partition function when the number of boxes and particles are fixed. Next, the box occupancies are investigated in the thermodynamic limit. Fixing the number of boxes, we give the asymptotic behaviour of box occupancies when the number of particles becomes large, for different regimes in parameter space  $\beta$ . The total energy of the configurations is next studied under the same limiting conditions. We use singularity analysis of poly-logarithmic functions to do so (see [6]). In the sequel, the number of states (boxes) with prescribed number of particles is studied in the canonical ensemble. In particular, since the number of occupied states deserves interest, we shall study it in some detail. Several other statistical issues of interest are briefly discussed.

The last Section is devoted to the occupancy distributions in the grand-canonical ensemble after the number of particles was suitably randomized. In a specific low temperature - large number of boxes asymptotical regime, it is shown that occupancies are governed by uniques or singletons as in Fermi-Dirac statistics.

## 2 Urn models and occupancies

To fix the background, we start with generalities on thermalized urn models before concentrating on the remarkable zeta-urn sub-class.

### 2.1 Generalities on thermalized urn models

Consider an urn model with  $n$  distinguishable boxes within which  $k$  particles are to be allocated ‘at random’. Suppose first the energy required to put  $k_m$  particles within box number  $m$ ,  $m = 1, \dots, n$ , is  $e_{k_m, m} > 0$ . Two cases arise:

1/  $e_{k_m, m}$  depends explicitly on box label  $m$ . A famous example is  $e_{k_m, m} = k_m \epsilon_m$  where  $\epsilon_m$  is the energy required to put a *single* particle within box number  $m$ ,  $m = 1, \dots, n$ . Typically,  $\epsilon_m = m^\alpha$ , for some  $\alpha > 1$ . Note that  $e_{k_m, m}$  is an increasing sequence in both arguments  $(k_m, m)$ . In this case, energy is box dependent (**BDE**).

2/  $e_{k_m, m}$  does not depend on  $m$ , hence  $e_{k_m, m} = e_{k_m}$  where  $e_{k_m}$  is simply assumed to increase with  $k_m$ . In this case, energy is box independent (**BIE**).

Occupancy distributions which we shall consider are Gibbs distributions which can be obtained while maximizing occupancies distribution entropy under the constraint that the average total energy  $\langle h \rangle := \langle h_{k, n} \rangle$  of the  $k$ -particle system configurations within  $n$  boxes is fixed. In this setup, as usual, a parameter  $\beta$  (the inverse of temperature) pops in; it is the Legendre conjugate to the average energy  $\langle h \rangle$ . Let  $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ . We shall be interested into the law of

$$\mathbf{K}_{k, n} := (K_{k, n}(1), \dots, K_{k, n}(n)) \in \mathbb{N}_0^n$$

which is an integral-valued random vector which counts the occupancy numbers within the  $n$  different boxes in a  $k$ -system of particles. Depending now on whether particles to be allocated are distinguishable or not, two additional cases arise; finally, we are left with four cases:

- Assume first particles to be allocated within labelled boxes are distinguishable (Maxwell-Boltzmann statistics).  
- With  $\sum_{m=1}^n k_m = k$  and  $\mathbf{k}_n := (k_1, \dots, k_n) \in \mathbb{N}_0^n$ , if energy is box dependent,  $\mathbf{K}_{k,n}$  follows the **BDE-DP** (box-dependent energy, distinguishable particles) distribution if:

$$\mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \frac{1}{Z_{k,n}(\beta)} \prod_{m=1}^n \frac{\sigma_{k_m,m}^{-\beta}}{k_m!},$$

where partition function

$$Z_{k,n}(\beta) = [z^k] \prod_{m=1}^n Q_{\beta,m}(z) \text{ and } Q_{\beta,m}(z) = \sum_{k_m \in \mathbb{N}_0} \frac{z^{k_m}}{k_m!} e^{-\beta e_{k_m,m}}$$

is a product of ‘exponential’ generating functions [In the latter formula,  $[z^k] f(z)$  stands for the  $z^k$ -coefficient in the series expansion of function  $f(z)$ ]. Here,  $\sigma_{k_m,m}^{-\beta} := e^{-\beta e_{k_m,m}}$  are the usual Boltzmann weights. In addition,  $\beta$  and  $\langle h \rangle := \langle h_{k,n} \rangle$  are Legendre conjugates, related as usual through  $-\partial_\beta \log Z_{k,n}(\beta) = \langle h \rangle$ .

- when energy is box independent,  $\mathbf{K}_{k,n}$  follows the **BIE-DP** (box-independent energy, distinguishable particles) distribution if:

$$\mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \frac{1}{Z_{k,n}(\beta)} \prod_{m=1}^n \frac{\sigma_{k_m}^{-\beta}}{k_m!},$$

where, with  $\sigma_k := \exp e_k$

$$Z_{k,n}(\beta) = [z^k] Q_\beta(z)^n \text{ and } Q_\beta(z) = \sum_{k \in \mathbb{N}_0} \frac{z^k}{k!} e^{-\beta e_k}.$$

In this case, the distribution of  $\mathbf{K}_{k,n}$  is exchangeable or symmetric (as a result of its invariance under permutation of the entries).

- Assume now particles are undistinguishable (Bose-Einstein statistics).  
- If energy is box-dependent,  $\mathbf{K}_{k,n}$  follows the **BDE-UP** distribution if:

$$\mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \frac{1}{Z_{k,n}(\beta)} \prod_{m=1}^n \sigma_{k_m,m}^{-\beta},$$

where partition function

$$Z_{k,n}(\beta) = [z^k] \prod_{m=1}^n P_{\beta,m}(z) \text{ with } P_{\beta,m}(z) = \sum_{k_m \in \mathbb{N}_0} z^{k_m} e^{-\beta e_{k_m,m}}$$

is now the product of ‘ordinary’ generating functions.

- if energy is box-independent,  $\mathbf{K}_{k,n}$  follows the **BIE-UP** distribution if:

$$\mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \frac{1}{Z_{k,n}(\beta)} \prod_{m=1}^n \sigma_{k_m}^{-\beta},$$

where

$$Z_{k,n}(\beta) = [z^k] P_{\beta}(z)^n \text{ and } P_{\beta}(z) = \sum_{k \in \mathbb{N}_0} z^k e^{-\beta e_k}.$$

## 2.2 Zeta-urns: statistical properties

In this manuscript, we shall limit ourselves to the **BIE-UP** distribution. We shall therefore suppose that the energy required to put  $k_m$  particles within box number  $m$ ,  $m = 1, \dots, n$ , is  $e_{k_m} > 0$  where  $e_{k_m}$  is an increasing function with  $k_m$ , independently of box label  $m$ . Assume  $e_0 := 0$ . We shall also further specify and assume that

$$\frac{e_k}{k} \xrightarrow{k \nearrow \infty} 0$$

meaning that energy is sub-linear. For example,  $e_k = k^\alpha$  with  $\alpha \in (0, 1)$  and  $e_k = \log(1 + k)$  would do. Interest into such specific allocation models is because they are likely to present a phase transition phenomenon at all temperatures in the first case and when temperature is small enough in the second case. We shall in fact further restrict ourselves to the **BIE-UP** model with  $e_k = \log(1 + k)$  or equivalently  $\sigma_k = 1 + k$ , which is the zeta urn model (see [1] and [8]).

### • Occupancy distribution statistics

With  $\sum_{m=1}^n k_m = k$  and  $\mathbf{k}_n := (k_1, \dots, k_n) \in \mathbb{N}_0^n$ , the zeta occupancy numbers  $\mathbf{K}_{k,n}$  within the  $n$  different boxes in a  $k$ -system of particles follows the exchangeable distribution:

$$(2.1) \quad \mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \frac{1}{Z_{k,n}(\beta)} \prod_{m=1}^n \sigma_{k_m}^{-\beta},$$

with  $\mathbf{E}(K_{k,n}(m)) = k/n$  and  $\sigma_k = 1 + k$ . Clearly,

$$(2.2) \quad Z_{k,n}(\beta) = [z^k] D_{\beta}(z)^n$$

where  $D_{\beta}(z) := 1 + \sum_{k \geq 1} \sigma_k^{-\beta} z^k$  is a thermalized ordinary generating function (here, a Dirichlet series). Because all the information on the model is enclosed

in the two-parameters function  $D_\beta(z)$ , let us first study it before proceeding with the evaluation of  $Z_{k,n}(\beta)$ .

• **Some properties of the function  $D_\beta(z)$**

First, the (real) definition domain of  $D_\beta(z)$  is  $\beta > 0$  and  $z \in [0, 1)$  or  $\beta > 1$  and  $z = 1$ . Incidentally, with  $\Gamma(\cdot)$  the Euler gamma function,  $D_\beta(z)$  has alternative Bose-Einstein integral representation:

$$D_\beta(z) = \frac{1}{\Gamma(\beta)} \int_0^\infty \frac{t^{\beta-1} e^{-t}}{1 - ze^{-t}} dt,$$

expressing the poly-logarithmic (multiplicative) Dirichlet series

$$zD_\beta(z) := \sum_{k \geq 1} k^{-\beta} z^k$$

as a Mellin transform of its additive counterpart  $\sum_{k \geq 1} e^{-\beta kt} z^k = ze^{-t} / (1 - ze^{-t})$ ,  $t > 0$ . This straightforward number theoretic representation is well-known, especially for  $D_\beta(1) =: \zeta(\beta)$ , the Riemann zeta function. Next, when  $\beta > 0$  and  $z \in [0, 1)$

$$D_{\beta+1}(z) = \frac{1}{z} \int_0^z D_\beta(z') dz'.$$

From the statistical point of view, this suggests to consider the positive random variable  $T_{\beta,z}$  with generalized Planck density (see [12]), namely

$$f_{T_{\beta,z}}(t) = \frac{1}{\Gamma(\beta) \cdot D_\beta(z)} \frac{t^{\beta-1} e^{-t}}{1 - ze^{-t}}, \quad t > 0$$

where  $\beta > 0$  and  $z \in [0, 1)$  or  $\beta > 1$  and  $z = 1$ . This family of probability densities includes the classical (standard) Planck density ( $z = 1$ ) and the gamma( $\beta$ ) densities ( $z = 0$ ). Actually, it is a discrete scale-mixture of gamma( $\beta$ ) distributions, since

$$f_{T_{\beta,z}}(t) = \sum_{k \geq 0} \frac{z^k (1+k)^{-\beta}}{D_\beta(z)} \frac{(1+k)^\beta}{\Gamma(\beta)} t^{\beta-1} e^{-(1+k)t}.$$

In other words,  $T_{\beta,z} \stackrel{d}{=} S_{\beta,z} \cdot T_{\beta,0}$  where  $S_{\beta,z} := (1 + K_{\beta,z})^{-1}$  is the random scale change. In the latter expression,  $K_{\beta,z}$  is such that  $\mathbf{P}(K_{\beta,z} = k) = \frac{z^k (1+k)^{-\beta}}{D_\beta(z)}$ ,  $k \in \mathbb{N}_0$ ; it is a (say) discrete zeta( $\beta, z$ )–distributed random variable, independent of gamma( $\beta$ )–distributed  $T_{\beta,0}$  with shape parameter  $\beta$ . When  $\beta > 1$ ,  $z \rightarrow D_\beta(z)$  is absolutely monotone on  $(0, 1)$  in the sense that order  $l$  derivatives  $D_\beta^{(l)}(z) \geq 0$  for all  $l \geq 0$  and  $z \in (0, 1)$ . Indeed,  $D_\beta(z)/D_\beta(1)$  is the generating function of the integral valued random variable  $K_{\beta,1}$ . Raising  $D_\beta(z)$  to

the power  $n$  also gives an absolutely monotone  $z$ -function on  $(0, 1)$  (because, up to a constant, it is the generating function of the sum of  $n$  independent and identically distributed (iid) copies of  $K_{\beta,1}$ ).

We finally note from the expression of  $f_{T_{\beta,z}}(t)$  that for  $\beta > 0$  and  $z \in [0, 1)$

$$\mathbf{E} \left( T_{1,z}^{\beta-1} \right) = \Gamma(\beta) \frac{D_\beta(z)}{D_\beta(1)}$$

where  $\Gamma(\beta) = \mathbf{E} \left( T_{1,0}^{\beta-1} \right)$ . Thus  $\tilde{D}_\beta(z) := \frac{D_\beta(z)}{D_\beta(1)}$  also interprets as the moment function of  $S_z := (1 + K_{1,z})^{-1}$ , namely:  $\tilde{D}_\beta(z) = \mathbf{E} \left( S_z^{\beta-1} \right)$  recalling that

$$T_{1,z} \stackrel{d}{=} T_{1,0} \cdot S_z$$

where  $T_{1,0}$  has exponential distribution and is independent the lattice scale random variable  $S_z$  supported by  $(1+k)^{-1}$ ,  $k \in \mathbb{N}_0$ . As a scale mixture of exponentially distributed random variables, the random variable  $T_{1,z}$  is infinitely divisible (see [13]).

This shows that, with  $z \in [0, 1)$ , the function  $\beta \rightarrow \tilde{D}_\beta(z)$  is completely monotone as a function of  $\beta$  (in the sense that  $(-1)^n \partial_\beta^n \tilde{D}_\beta(z) \geq 0$  for all  $\beta > 1$ ). Indeed, by Bernstein theorem,  $\tilde{D}_\beta(z)$  is the Laplace-Stieltjes transform of  $-\log S_z > 0$ . Raising  $\tilde{D}_\beta(z)$  to the power  $n$  also gives rise to a completely monotone  $\beta$ -function on  $(1, \infty)$  (because it is the moment function of the product of  $n$  iid copies of  $S_z$ ).

Since energy is sub-linear, the convergence radius of the series  $D_\beta(z)$  is  $z_c = 1$ . Function  $D_\beta(z)$  is absolutely monotone on  $(0, z_c)$ . It increases with  $z$  and its  $l$ -th derivative at point 1, say  $D_\beta^{(l)}(1)$ , is finite if and only if  $\beta > l + 1$ . In particular,  $D_\beta(1)$  is finite if and only  $\beta > 1$  and both  $D_\beta(1)$  and  $D'_\beta(1)$  are finite if and only  $\beta > 2$ .

#### • Canonical partition function

Let us now turn back our attention to  $Z_{k,n}(\beta)$ . Clearly,  $Z_{k,n}(\beta)$  fulfills the recurrence:

$$Z_{k,n+1}(\beta) = \sum_{l=0}^k Z_{k-l,n}(\beta) \sigma_l^{-\beta}, \quad n \geq 1$$

with boundary conditions  $Z_{k,1}(\beta) = (1+k)^{-\beta}$ ,  $k \geq 0$  and  $Z_{0,n}(\beta) = 1$ ,  $n \geq 1$ .

Let  $\mathbb{N} := \{1, 2, \dots\}$ . By Faa di Bruno formula for potentials (integral powers), with  $\{n\}_l := n(n-1)\dots(n-l+1)$ , a closed-form solution is

$$Z_{k,n}(\beta) = \frac{1}{k!} \sum_{l=1}^k \{n\}_l B_{k,l}(\bullet! \sigma_\bullet^{-\beta})$$

where, with  $x_\bullet := (x_1, x_2, \dots)$ ,  $B_{k,l}(x_\bullet)$  is a Bell polynomials in the variables  $x_\bullet$  (see [5]):

$$B_{k,l}(x_\bullet) := \frac{k!}{l!} \sum_{k_m \in \mathbb{N}: \sum_{m=1}^l k_m = k} \prod_{i=1}^l \frac{x_{k_m}}{k_m!}, \quad l = 1, \dots, k.$$

In other words, after some simplifications

$$Z_{k,n}(\beta) = \sum_{p=1}^{k \wedge n} \binom{n}{p} \sum_{k_m \in \mathbb{N}: \sum_{m=1}^p k_m = k} \prod_{m=1}^p (1 + k_m)^{-\beta}$$

is the closed-form expression of  $Z_{k,n}(\beta)$ . Note that, with

$$\sigma \in \text{Span} \left\{ \prod_1^n (1 + k_m), \text{ when } k_1, \dots, k_n \in \mathbb{N}_0^n \text{ and } k_1 + \dots + k_n = k \right\}$$

we also have

$$Z_{k,n}(\beta) = \sum_{\sigma} |C_{k,n}(\log \sigma)| \sigma^{-\beta}$$

where  $|C_{k,n}(h)| = \#C_{k,n}(h)$  and

$$C_{k,n}(h) = \left\{ k_1, \dots, k_n \in \mathbb{N}_0^n: k_1 + \dots + k_n = k \text{ and } \log \prod_{m=1}^n (1 + k_m) = h \right\}.$$

• **The thermodynamic limit and evidence of a phase transition**

Let  $\rho > 0$  and assume  $k = \kappa_n := \lfloor n\rho \rfloor$ , so that  $\rho$  interprets as the box density of particles. When  $n \nearrow \infty$ , we further get the thermodynamic limit. Observing that

$$Z_{k,n}(\beta) = \frac{1}{2i\pi} \oint z^{-(k+1)} D_\beta(z)^n dz,$$

a saddle point estimate gives

$$\begin{aligned} -\frac{1}{n} \log Z_{\lfloor n\rho \rfloor, n}(\beta) &\underset{n \nearrow \infty}{\sim} -\frac{1}{n} \log \left[ z_\beta(\rho)^{-\kappa_n} D_\beta(z_\beta(\rho))^n \right] \\ &\underset{n \nearrow \infty}{\sim} \rho \log z_\beta(\rho) - \log D_\beta(z_\beta(\rho)) =: \beta F_\beta(\rho) \end{aligned}$$

where saddle point  $z_\beta(\rho)$  is defined implicitly by

$$z_\beta(\rho) \Phi'_\beta(z_\beta(\rho)) = \rho \text{ and } \Phi_\beta(z) := \log D_\beta(z).$$

The function  $F_\beta(\rho)$  is the free energy per box in the thermodynamic limit.

The range of the function  $z \Phi'_\beta(z)$  when  $z \in (0, 1)$  is  $(0, \rho_c)$  where  $\rho_c = \infty$  if  $\beta \leq 2$  and  $\rho_c = \Phi'_\beta(1) = \zeta(\beta - 1) / \zeta(\beta) - 1 < \infty$  if  $\beta > 2$ . Since  $z \in (0, 1) \rightarrow$

$z\Phi'_\beta(z)$  is monotone increasing,  $z_\beta(\rho)$  is uniquely defined for each  $\rho \in (0, \rho_c)$ . By the Lagrange-inversion formula, when  $\rho \in (0, \rho_c)$ :

$$z_\beta(\rho) = 1 + \sum_{l \geq 1} \frac{\rho^l}{l} h_l(\beta) \quad \text{with } h_l(\beta) := [z^{l-1}] \Phi'_\beta(z)^{-l}.$$

For each  $\beta > 0$ , the free energy function  $\rho \rightarrow F_\beta(\rho)$  is a convex function of  $\rho > 0$ . Further  $F_\beta(0) = 0$ ,  $F'_\beta(0) = -\infty$  and  $F_\beta(\rho)$  decreases with  $\rho$ . When  $\beta > 1$ ,  $F_\beta(\rho_c) = -\frac{1}{\beta} \log \zeta(\beta) \in (-\infty, 0)$  and  $F'_\beta(\rho_c) = 0$ . Next,  $\rho_c = \infty$  if  $\beta \in (1, 2]$  whereas, when  $\beta > 2$ ,  $\rho_c < \infty$  and  $F_\beta(\rho) = F_\beta(\rho_c)$  in the range  $\rho \in [\rho_c, \infty)$ . In any case, when  $\beta > 1$ ,  $F_\beta(\rho)$  is bounded below by  $F_\beta(\rho_c)$  whereas it is unbounded below when  $\beta \in (0, 1)$ . When  $\beta > 2$ , the critical density  $\rho_c < \infty$  separates a fluid phase ( $\rho < \rho_c$ ) from a condensed phase ( $\rho > \rho_c$ ). Clearly, the critical properties of this phase transition model are dictated by the Riemann zeta function.

The partition function behaviour shows that, in the thermodynamic limit,  $\mathbf{K}_{\kappa_n, n}$  is asymptotically iid in the sense that

$$-\frac{1}{n} \log \mathbf{P}(\mathbf{K}_{\kappa_n, n} = \mathbf{k}_n) \underset{n \nearrow \infty}{\sim} -\frac{1}{n} \log \left[ \prod_{m=1}^n \frac{\sigma_{k_m}^{-\beta} z_\beta(\rho)^{k_m}}{D_\beta(z_\beta(\rho))} \right].$$

• **Additional aspects of the occupancy distribution**

We now draw various immediate conclusions from the occupancy distribution expressions.

(i) The conditional probability to occupy state  $m$  is

$$\pi_m := \mathbf{P}(K_{k, n}(m) > 0) = 1 - \frac{Z_{k, n-1}(\beta)}{Z_{k, n}(\beta)}.$$

(ii) If  $k \leq n$ , with  $\mathbf{1}_k := (0, 1, 0, \dots, 1, 0, 1)$  a vector with  $k$  “1” in any of the  $\binom{n}{k}$  possible positions

$$\mathbf{P}(\mathbf{K}_{k, n} = \mathbf{1}_k) = \binom{n}{k} \frac{\sigma_1^{-\beta k}}{Z_{k, n}(\beta)}$$

is the probability that the  $k$  particles will occupy any  $k$  distinct boxes.

(iii) (Bose-Einstein): When  $\beta$  tends to 0, the joint law of  $\mathbf{K}_{k, n}$  looks uniform

$$\mathbf{P}(\mathbf{K}_{k, n} = \mathbf{k}_n) = \frac{1}{\binom{n+k-1}{k}}$$

and the 1–dimensional distribution reads

$$\mathbf{P}(K_{k,n}(1) = l) = \frac{\binom{n+k-l-2}{k-l}}{\binom{n+k-1}{k}}, \quad l = 0, \dots, k.$$

(iv) Occupancies  $\mathbf{K}_{k,n}$  have exchangeable distribution. We observe from Eqs. (2.1, 2.2) that

$$\mathbf{E} \left( \prod_{m=1}^n u_m^{K_{k,n}(m)} \right) = \frac{[z^k] \prod_{m=1}^n D_\beta(u_m z)}{[z^k] D_\beta(z)^n}.$$

Putting  $u_2 = \dots = u_n = 1$ , the one-dimensional distribution of  $K_{k,n}(1)$  can easily be checked from the convolution formula to be

$$(2.3) \quad \mathbf{P}(K_{k,n}(1) = l) = \frac{\sigma_l^{-\beta} Z_{k-l, n-1}(\beta)}{Z_{k,n}(\beta)}, \quad l = 0, \dots, k.$$

(v) Using the saddle-point analysis of  $Z_{k,n}(\beta)$ , in the thermodynamic limit, we immediately have

$$K_{\kappa_n, n}(1) \xrightarrow[n \nearrow \infty]{} \mathcal{K}_\rho$$

where  $\mathcal{K}_\rho$  has  $\text{zeta}(\beta, z_\beta(\rho))$  distribution, namely:

$$\mathbf{P}(\mathcal{K}_\rho = l) = \frac{\sigma_l^{-\beta} \cdot z_\beta(\rho)^l}{D_\beta(z_\beta(\rho))}, \quad l \in \mathbb{N}_0.$$

Assuming  $\beta > 1$ , when  $\rho \nearrow \rho_c$ ,  $z_\beta(\rho) \nearrow 1$  and  $\mathcal{K}_{\rho_c}$  has  $\text{zeta}(\beta)$  – distribution with power-law tails:

$$\mathbf{P}(\mathcal{K}_{\rho_c} = l) = \frac{\sigma_l^{-\beta}}{\zeta(\beta)}, \quad l \in \mathbb{N}_0.$$

The critical random variable  $\mathcal{K}_{\rho_c}$  has finite moments of order strictly less than  $\beta - 1$ . Its distribution is  $\text{zeta}(\beta)$  law, also known as Zipf law (see [10]).

(vi) Let us now investigate *correlations*.

The joint *falling factorial moments* of  $(K_{k,n}(m))_{m=1}^r$  are also available in closed form. Let us also consider the falling factorial moments of  $\mathbf{K}_{k,n}$ . Fix  $\mathbf{l}_n := (l_1, \dots, l_n) \in \mathbb{N}_0^n$ , summing to  $l$  with  $l \leq k$ . Expressing the joint generating function  $\mathbf{E} \left( \prod_{m=1}^n u_m^{K_{k,n}(m)} \right)$  in terms of  $v_m = u_m - 1$ , we have

$$\mathbf{E} \left[ \prod_{m=1}^n \{K_{k,n}(m)\}_{l_m} \right] = \frac{\prod_{m=1}^n l_m! [z^k] \prod_{m=1}^n [v_m^{l_m}] D_\beta(z(v_m + 1))}{[z^k] D_\beta(z)^n}.$$

Let  $\{k\}_l := k(k-1)\dots(k-l+1)$  stand for the  $l$ -falling factorials of  $k$ . Since  $l_m! [v_m^{l_m}] D_\beta(z(v_m+1)) = \sum_{\mathbf{k}_m \geq l_m} \{k_m\}_{l_m} \sigma_{k_m}^{-\beta} \cdot z^{k_m}$ , with  $\mathbf{k}_m$  summing to  $k$ , we get

$$(2.4) \quad \mathbf{E} \left[ \prod_{m=1}^n \{K_{k,n}(m)\}_{l_m} \right] = \frac{\sum_{\mathbf{k}_n \geq \mathbf{l}_n} \prod_{m=1}^n \{k_m\}_{l_m} \sigma_{k_m}^{-\beta}}{Z_{k,n}(\beta)} = \frac{[z^k] \prod_{m=1}^n D_\beta^{(l_m)}(z)}{Z_{k,n}(\beta)}.$$

Particularizing  $l_m = 1, m = 1, \dots, r, l_m = 0, m = r+1, \dots, n$ , we get the joint moments

$$\mathbf{E} \left( \prod_{m=1}^r K_{k,n}(m) \right) = \frac{[z^k] \left( D'_\beta(z)^r D_\beta(z)^{n-r} \right)}{Z_{k,n}(\beta)}.$$

If in particular  $l_1 = 2, l_m = 0, m = 1, \dots, n$

$$\mathbf{E} (\{K_{k,n}(1)\}_2) = \frac{[z^k] \left( D''_\beta(z) D_\beta(z)^{n-1} \right)}{Z_{k,n}(\beta)}$$

in such a way that

$$\sigma^2(K_{k,n}(1)) = \frac{[z^k] \left( D''_\beta(z) D_\beta(z)^{n-1} \right)}{Z_{k,n}(\beta)} + \frac{k}{n} \left( 1 - \frac{k}{n} \right).$$

Squaring  $\sum_{m=1}^n K_{k,n}(m) = k$ , averaging and using exchangeability, for each  $m_1, m_2 \in [n], m_1 \neq m_2$ , we get the covariance

$$\text{Cov}(K_{k,n}(m_1), K_{k,n}(m_2)) = -\frac{\sigma^2(K_{k,n}(1))}{n-1}.$$

(vii) Fixed number of boxes and large number of particles: We now prove the following convergence in distribution.

**Lemma 1** Assume  $n(1-\beta) \notin \{-1, -2, \dots\}$ .

(i) Let  $(S_1, \dots, S_n) := \mathbf{S}_n \stackrel{d}{\sim} D_n(1-\beta)$ , the Dirichlet distribution on the simplex with parameter  $1-\beta$ . If  $\beta \in (0, 1)$ , it holds that

$$(2.5) \quad \mathbf{K}_{k,n}/k \xrightarrow{d} \mathbf{S}_n \text{ as } k \nearrow \infty.$$

In particular,

$$K_{k,n}(1)/k \xrightarrow{d} S_n \stackrel{d}{\sim} \text{beta}(1-\beta; (n-1)(1-\beta)) \text{ as } k \nearrow \infty.$$

(ii) If  $\beta > 1$ ,

$$K_{k,n}(1) \xrightarrow{d} K_n \stackrel{d}{\sim} \frac{1}{n} \delta_\infty + \left( 1 - \frac{1}{n} \right) \text{zeta}(\beta), \text{ as } k \nearrow \infty,$$

where a  $\mathbb{N}_0$ -valued random variable has the zeta( $\beta$ ) distribution if its probability generating function is  $D_\beta(u)/D_\beta(1)$ .

**Proof:** Suppose  $n(1-\beta) \notin \{-1, -2, \dots\}$  (If this were not the case, the singularity would be logarithmic and would deserve a special treatment which we skip).

(i) When  $\beta \in (0, 1)$ , to the leading algebraic order, we have

$$D_\beta(z)^n \underset{z \nearrow z_c=1}{\sim} \Gamma(1-\beta)^n \cdot (1-z)^{-n(1-\beta)}$$

where  $D_\beta(z) := \sum_{k \geq 0} \sigma_k^{-\beta} z^k$ . [The exact asymptotic equivalent of  $D_\beta(z)$  is

$$D_\beta(z) \underset{z \nearrow z_c=1}{\sim} \Gamma(1-\beta) (-\log z)^{-(1-\beta)} + \sum_{l \geq 0} \frac{(-1)^l}{l!} \zeta(\beta-l) z^l. ]$$

It follows from standard singularity analysis (see [6] for generalities on transfer theorems and [7] for the specific poly-log functions) that

$$Z_{k,n}(\beta) := [z^k] D_\beta(z)^n \underset{k \nearrow \infty}{\sim} \frac{\Gamma(1-\beta)^n}{\Gamma(n(1-\beta))} k^{n(1-\beta)-1}$$

Next, we have

$$\mathbf{P}(K_{k,n}(1) = l) = \frac{Z_{l,1}(\beta) Z_{k-l,n-1}(\beta)}{Z_{k,n}(\beta)}, l = 0, \dots, k.$$

Thus, with  $s \in (0, 1)$ , as  $k \nearrow \infty$

$$k \mathbf{P}\left(\frac{K_{k,n}(1)}{k} = s\right) \sim \frac{\Gamma(n(1-\beta))}{\Gamma(1-\beta) \Gamma((n-1)(1-\beta))} s^{(1-\beta)-1} (1-s)^{(n-1)(1-\beta)-1}$$

where one recognizes the density of a Beta( $1-\beta, (n-1)(1-\beta)$ ) distributed random variable  $S_n$ . This shows that

$$\frac{K_{k,n}(1)}{k} \underset{k \nearrow \infty}{\overset{d}{\sim}} S_n \overset{d}{\sim} \text{Beta}(1-\beta, (n-1)(1-\beta)).$$

This limiting distribution is the marginal of a Dirichlet distribution  $D_n(1-\beta)$  with parameters  $n$  and  $(1-\beta)$  towards which  $\mathbf{K}_{k,n}/k$ , more generally, converges weakly as  $k \nearrow \infty$ . We recall that the density on the simplex of  $\mathbf{S}_n \overset{d}{\sim} D_n(1-\beta)$  is exchangeable with

$$f_{\mathbf{S}_n}(s_1, \dots, s_n) = \frac{\Gamma(n(1-\beta))}{\Gamma(1-\beta)^n} \prod_{m=1}^n s_m^{(1-\beta)-1} \delta_{(\sum_1^n s_m=1)}.$$

(ii) If  $\beta > 1$ ,  $D_\beta(z) \underset{z \nearrow z_c=1}{\sim} D_\beta(1) + \Gamma(1-\beta) \cdot (1-z)^{\beta-1}$ . Thus,

$$D_\beta(z)^n \underset{z \nearrow z_c=1}{\sim} D_\beta(1)^n + n\Gamma(1-\beta) \zeta(\beta)^{n-1} (1-z)^{-(1-\beta)}.$$

By singularity analysis,

$$Z_{k,n}(\beta) = [z^k] D_\beta(z)^n \underset{k \nearrow \infty}{\sim} n\zeta(\beta)^{n-1} k^{-\beta}.$$

We have

$$\mathbf{E}u^{K_{k,n}(1)} = \frac{[z^k] \left( D_\beta(zu) D_\beta(z)^{n-1} \right)}{Z_{k,n}(\beta)} \underset{k \nearrow \infty}{\sim} \left( 1 - \frac{1}{n} \right) \frac{D_\beta(u)}{D_\beta(1)}$$

because the dominant singularity of  $D_\beta(zu) D_\beta(z)^{n-1}$  is at  $z = 1$  so that the analysis of the numerator reduces to  $D_\beta(u) [z^k] D_\beta(z)^{n-1} = D_\beta(u) Z_{k,n-1}(\beta)$ .

This shows that, as  $k$  goes to  $\infty$ , with probability  $1/n$ ,  $K_{k,n}(1) = \infty$  whereas with probability  $1 - 1/n$ ,  $K_{k,n}(1)$  converges weakly to a zeta( $\beta$ ) distributed  $\mathbb{N}_0$ -valued random variable whose generating function is  $\frac{D_\beta(u)}{D_\beta(1)}$ ,  $u \in [0, 1]$ .  $\square$

• **Energy of the configurations**

A random variable of interest: the *total energy of the occupancy configurations*. It is the random variable

$$\mathcal{H}_{k,n} := \sum_{m=1}^n e_{K_{k,n}(m)}.$$

Clearly, it is characterized by its Laplace-Stieltjes transform (LST)

$$\mathbf{E} \left( e^{\lambda \mathcal{H}_{k,n}} \right) = \frac{Z_{k,n}(\beta + \lambda)}{Z_{k,n}(\beta)}; \quad \lambda \geq 0.$$

With  $h \in \mathcal{S}_{k,n} := \text{Span}\{\sum_{m=1}^n e_{k_m}, \text{ when } k_1, \dots, k_n \in \mathbb{N}_0^n \text{ and } k_1 + \dots + k_n = k\}$ , this is also

$$\begin{aligned} \mathbf{P}[\mathcal{H}_{k,n} = h] &= \sum_{k_1 + \dots + k_n = k} \mathbf{1} \left( \sum_{m=1}^n e_{K_{k,n}(m)} = h \right) \frac{\prod_{m=1}^n \sigma_{k_m}^{-\beta}}{Z_{k,n}(\beta)} \\ &= \frac{e^{-\beta h} \cdot |C_{k,n}(h)|}{Z_{k,n}(\beta)}, \end{aligned}$$

where  $|C_{k,n}(h)| = \#C_{k,n}(h)$  and

$$C_{k,n}(h) = \left\{ k_1, \dots, k_n \in \mathbb{N}_0^n: k_1 + \dots + k_n = k \text{ and } \sum_{m=1}^n e_{k_m} = h \right\}.$$

Recalling  $e_{k_m} = \log(1 + k_m)$ , we also have  $h \in \mathcal{S}_{k,n}$  if and only if  $\sigma := e^h$  and

$$\sigma \in \text{Span} \left\{ \prod_1^n (1 + k_m), \text{ when } k_1, \dots, k_n \in \mathbb{N}_0^n \text{ and } k_1 + \dots + k_n = k \right\}.$$

Next,  $C_{k,n}(\log \sigma) := \{k_1, \dots, k_n \in \mathbb{N}_0^n: k_1 + \dots + k_n = k \text{ and } \prod_1^n (1 + k_m) = \sigma\}$  where  $\sigma$  is an integer belonging to the above set. There is no known explicit expression of the number of both additive and multiplicative integer ‘compositions’, namely of  $|C_{k,n}(\log \sigma)|$ .

Note however that, as a probability,  $\frac{e^{-\beta h} \cdot |C_{k,n}(h)|}{Z_{k,n}(\beta)} \leq 1$  so that

$$\log |C_{k,n}(h)| \leq \beta h + \log Z_{k,n}(\beta).$$

Finally, the joint law of the occupancies conditionally given  $\mathcal{H}_{k,n} = h$  is also of interest in the micro-canonical ensemble: With  $\mathbf{k}_n \in C_{k,n}(h)$  and  $h \in \mathcal{S}_{k,n}$ , we have

$$\mathbf{P}[\mathbf{K}_{k,n} = \mathbf{k}_n \mid \mathcal{H}_{k,n} = h] = \frac{1}{|C_{k,n}(h)|}$$

and the distribution is uniform over the set  $C_{k,n}(h)$ .

*Remark:* As is well-known, in a neighborhood of  $\beta = 1$ ,  $\zeta(\beta) = (\beta - 1)^{-1} + c + o(1)$  where  $c$  is Euler’s constant. Thus  $\zeta(\beta)^n \underset{\beta \searrow 1}{\sim} (\beta - 1)^{-n}$ . Let  $\mathcal{S}_n(\sigma') := \text{Span}\{k_1 + \dots + k_n, \text{ when } k_1, \dots, k_n \in \mathbb{N}_0^n \text{ and } \prod_1^n (1 + k_m) = \sigma'\}$ . It follows from singularity analysis that

$$\sum_{\sigma'=1}^{\sigma} \sum_{k \in \mathcal{S}_n(\sigma')} |C_{k,n}(\log \sigma')| \underset{\sigma \nearrow \infty}{\sim} \frac{\sigma^{n-1}}{\Gamma(n)}$$

Note that  $\sum_{k \in \mathcal{S}_n(\sigma')} |C_{k,n}(\log \sigma')| = \#\{k_1, \dots, k_n \in \mathbb{N}_0^n: \prod_1^n (1 + k_m) = \sigma'\}$  is the number of multiplicative compositions of  $\sigma'$  with  $n$  summands each possibly taking the value 0.  $\diamond$

We now supply limit laws for total configurational energy when  $k \nearrow \infty$  in the different regimes for  $\beta$ .

**Proposition 2** (i) *Let  $\beta \in (0, 1)$ . Let  $\eta_0 > 0$  be a random variable with  $\text{gamma}(n(1 - \beta))$  distribution. With  $\eta_0$  independent of  $\mathcal{H}_{k,n}$*

$$\mathcal{H}_{k,n} - n \log \left( \frac{k}{\eta_0} \right) \xrightarrow{d} \sum_{m=1}^n \eta_m \text{ as } k \nearrow \infty$$

where  $(\eta_m; m = 1, \dots, n)$  are  $\text{exp-gamma}(1 - \beta)$  real-valued iid random variables.

(ii) When  $\beta > 1$  is not integer, we get

$$\mathcal{H}_{k,n} - \log k \xrightarrow{d} \sum_{m=1}^{n-1} \delta_m \text{ as } k \nearrow \infty$$

where  $(\delta_m; m = 1, \dots, n)$  are iid infinitely divisible random variables with common lattice exp-zeta( $\beta$ ) distribution

$$\mathbf{P}(\delta_1 = \log l) = \frac{l^{-\beta}}{\zeta(\beta)}, \quad l \in \mathbb{N} := \{1, 2, \dots\}.$$

**Proof:**

(i)  $\beta \in (0, 1)$ : In this case,

$$Z_{k,n}(\beta) = [z^k] D_\beta(z)^n \underset{k \nearrow \infty}{\sim} \frac{\Gamma(1-\beta)^n}{\Gamma(n(1-\beta))} k^{n(1-\beta)-1}$$

and so

$$\frac{Z_{k,n}(\beta + \lambda)}{Z_{k,n}(\beta)} \underset{k \nearrow \infty}{\sim} k^{-n\lambda} \left( \frac{\Gamma(1-\beta-\lambda)}{\Gamma(1-\beta)} \right)^n \frac{\Gamma(n(1-\beta))}{\Gamma(n(1-\beta-\lambda))}.$$

From this, the law of  $\eta_0$  can be read: it is given by  $\mathbf{E}\eta_0^{-\lambda} = \frac{\Gamma(n(1-\beta)-\lambda)}{\Gamma(n(1-\beta))}$ ,  $\lambda < n(1-\beta)$ . Also, the ones of  $(\eta_m; m = 1, \dots, n)$  are seen to be given by  $\mathbf{E}e^{-\lambda\eta_1} = \frac{\Gamma(1-\beta-\lambda)}{\Gamma(1-\beta)}$  for  $\lambda < 1-\beta$ . Thus  $\eta_0$  and  $e^{-\eta_1}$  are gamma distributed with the announced parameters and Part (i) follows up.

(ii)  $\beta > 1$ : Recalling  $Z_{k,n}(\beta) = [z^k] D_\beta(z)^n \underset{k \nearrow \infty}{\sim} n\zeta(\beta)^{n-1} k^{-\beta}$ , we indeed have

$$\mathbf{E}e^{-\lambda(\mathcal{H}_{k,n} - \log k)} = k^\lambda \frac{Z_{k,n}(\beta + \lambda)}{Z_{k,n}(\beta)} \underset{k \nearrow \infty}{\sim} \left( \frac{\zeta(\beta + \lambda)}{\zeta(\beta)} \right)^{n-1}$$

which is a product LST of the  $\delta_m$ s, with  $\mathbf{E}(e^{-\lambda\delta_1}) = \zeta(\beta + \lambda) / \zeta(\beta)$ . It is known (see [9]) that  $\delta_1$  is an infinitely divisible compound Poisson distribution, namely:

$$\delta_1 \stackrel{d}{=} \sum_{q=1}^{\mathcal{P}} \varepsilon_q$$

with  $(\varepsilon_q; q \geq 1)$  an iid sequence, independent of  $\mathcal{P} \stackrel{d}{\sim} \text{Poisson}(\log \zeta(\beta))$  and common law:

$$\mathbf{P}(\varepsilon_1 = \log i) = \frac{\phi_i(\beta)}{\log \zeta(\beta)}, \quad i = 2, 3, \dots$$

In the latter expression,  $\phi_i(\beta) = \frac{\Lambda(i) \cdot i^{-\beta}}{\log i}$ , where  $\Lambda(i) = \log p \cdot \mathbf{1}(i = p^l)$  (for some prime  $p \geq 2$  and some integer  $l \geq 1$ ) is the von Mangoldt function. This

follows from taking the logarithm of Euler product representation of Riemann zeta function, stating that, for  $\beta > 1$

$$\zeta(\beta) = \prod_{p \geq 2} (1 - p^{-\beta})^{-1}$$

where the infinite product runs over all prime numbers  $p$ . Stated differently, the jumps law also reads

$$\mathbf{P}(\varepsilon_1 \in dx) = \sum_{p \geq 2} \sum_{l \geq 1} \frac{p^{-l\beta}}{l} \delta_{l \log p}(dx). \quad \square$$

• **Occupancy distribution as a random allocation scheme**

Let  $z \in (0, 1)$ . Let  $(\xi_{z,m}; m \geq 1)$  be an iid sequence on  $\mathbb{N}_0 := \{0, 1, \dots\}$ , with discrete zeta( $\beta, z$ ) distribution, namely

$$(2.6) \quad \mathbf{P}(\xi_{z,1} = k) =: p(k) = \frac{\sigma_k^{-\beta} z^k}{D_\beta(z)}, \quad k \in \mathbb{N}_0.$$

The generating function of  $\xi_{z,1}$  is

$$(2.7) \quad \mathbf{E}(u^{\xi_{z,1}}) = \frac{D_\beta(zu)}{D_\beta(z)}, \quad 0 \leq u < 1/z.$$

The random variable  $\xi_{z,1}$  has mean  $\rho = z \frac{D'_\beta(z)}{D_\beta(z)}$  and finite variance  $\sigma^2$ .

Let  $P(k) := \sum_{k_1=0}^k p(k_1)$  and  $\bar{P}(k) := 1 - P(k)$  the tail distribution of  $\xi_{z,1}$ . Then,  $\frac{\bar{P}(k)}{\bar{P}(k-1)} \leq z \in (0, 1)$  and

$$\begin{aligned} p(k_1) &\underset{k_1 \nearrow \infty}{\sim} \frac{k_1^{-\beta} z^{k_1}}{D_\beta(z)} \\ \bar{P}(k) &\underset{k \nearrow \infty}{\sim} \frac{k^{-\beta} z^k}{D_\beta(z)}. \end{aligned}$$

The distribution of  $\xi_{z,1}$  can be obtained as follows: suppose we randomize the success probability  $p$  of a geometrically distributed random variable  $N$  by:  $p \rightarrow P := ze^{-X}$  where  $z \in (0, 1]$  and  $X \stackrel{d}{\sim} \text{gamma}(\beta)$ , independent of  $N$ . We assume  $\beta > 0$  if  $z \in (0, 1)$  and  $\beta > 1$  if  $z = 1$ . Then, with  $k \in \mathbb{N}_0$

$$\mathbf{P}(N \geq k) = \mathbf{E}(P^k) = z^k \mathbf{E}(e^{-kX}) = (1+k)^{-\beta} z^k.$$

$N$  is a log-gamma-geometric mixture with mean  $\mathbf{E}(N) = \sum_{k \geq 0} (1+k)^{-\beta} z^k = D_\beta(z)$ .

Next,  $\xi_{z,1}$  is the size-biased of  $N$  for which  $\mathbf{P}(\xi_{z,1} = k) = \mathbf{P}(N \geq k) / \mathbf{E}(N)$ , also interpreting as the limiting residual lifetime in a discrete renewal process

generated by  $N$ . Because of this representation, the sequence  $\mathbf{P}(\xi_{z,1} = k)$  is completely monotone and so  $\xi_{z,1}$  is infinitely divisible, meaning that it is a compound Poisson random variable, with

$$\xi_z = \sum_{q=1}^{\mathcal{P}_z} \varepsilon_{z,q}$$

where  $\mathcal{P}_z$  is a Poisson random variable with intensity  $\Phi_\beta(z) := \log D_\beta(z)$ , independent of the iid sequence of jumps  $(\varepsilon_{z,q}; q \geq 1)$  with common generating function

$$(2.8) \quad \mathbf{E}(u^{\varepsilon_{z,1}}) = \frac{\log D_\beta(zu)}{\log D_\beta(z)}.$$

Therefore, with  $\varphi_k(\beta) := [z^k] (\log D_\beta(z)) > 0$

$$\mathbf{P}(\varepsilon_{z,1} = k) = \frac{z^k \varphi_k(\beta)}{\log D_\beta(z)}, \quad k \in \mathbb{N}.$$

Let  $\zeta_{z,n} := \sum_{m=1}^n \xi_{z,m}$ ,  $n \geq 1$ , be the partial sum sequence of  $(\xi_{z,m}; m \geq 1)$  with  $\zeta_{z,0} := 0$ . Then, for any  $\rho > 0$ , one can easily check that

$$\mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) = \mathbf{P}(\xi_{z,1} = k_1, \dots, \xi_{z,n} = k_n \mid \zeta_{z,n} = k).$$

The zeta-urn distribution is in the class of random allocation schemes as the ones obtained by conditioning a random walk by its terminal value (see [11] and [4]). Further, the involved random variables  $(\xi_{z,m}; m = 1, \dots, n)$  are compound Poisson.

• **Order statistics:** Let

$$\mathbf{K}_{k,(n)} := (K_{k,(n)}(1), \dots, K_{k,(n)}(n))$$

be the ordered version of  $\mathbf{K}_{k,n}$ , with  $K_{k,(n)}(1) \geq \dots \geq K_{k,(n)}(n)$ . Due to the random allocation scheme representation of  $\mathbf{K}_{k,n}$ , it follows that

$$\mathbf{P}(K_{k,(n)}(m) \leq r) = \sum_{l=0}^{m-1} \binom{n}{l} \bar{P}(r)^l P(r)^{n-l} \frac{\mathbf{P}(\zeta_{z,l} + \zeta_{z,n-l} = k \mid \xi_{z,1}, \dots, \xi_{z,n} \leq r)}{\mathbf{P}(\zeta_{z,n} = k)}.$$

In particular, if  $m = 1$ , with  $r > \frac{k}{n}$

$$\mathbf{P}(K_{k,(n)}(1) \leq r) = (1 - \bar{P}(r))^n \frac{\mathbf{P}(\zeta_{z,n} = k \mid \xi_{z,1}, \dots, \xi_{z,n} \leq r)}{\mathbf{P}(\zeta_{z,n} = k)}$$

whereas, for  $m = n$ , with  $r < \frac{k}{n}$

$$\mathbf{P}(K_{k,(n)}(n) \leq r) = 1 - \bar{P}(r)^n \frac{\mathbf{P}(\zeta_{z,n} = k \mid \xi_{z,1}, \dots, \xi_{z,n} > r)}{\mathbf{P}(\zeta_{z,n} = k)}.$$

Let  $\kappa_n = \lfloor n\rho \rfloor$  in such a way that  $\kappa_n/n \rightarrow_{n \nearrow \infty} \rho$  (the asymptotic occupancy density). As  $n \nearrow \infty$ , the occupancy  $K_{\kappa_n, (n)}(1)$  of the box with largest amount of particles satisfies

$$\begin{aligned} \mathbf{P}(K_{\kappa_n, (n)}(1) = r_n - 1) &\rightarrow e^{-1} \\ \mathbf{P}(K_{\kappa_n, (n)}(1) = r_n) &\rightarrow 1 - e^{-1} \end{aligned}$$

where  $r_n$  is the sequence fulfilling  $n\bar{P}(r_n) \rightarrow 1$ , which is of order

$$(2.9) \quad r_n \sim -\log_{z_\beta(\rho)} \left( \frac{n}{D_\beta(z_\beta(\rho))} \left[ -\log_{z_\beta(\rho)} \left( \frac{n}{D_\beta(z_\beta(\rho))} \right) \right]^{-\beta} \right).$$

The support of  $K_{\kappa_n, (n)}(1)$  law consists in the two points  $r_n - 1$  and  $r_n$ , where  $r_n$  slowly moves to infinity as indicated. The discreteness of the distributions involved prevents the maximum from converging properly and instead forces this oscillatory behavior.

The smallest term  $K_{\kappa_n, (n)}(n)$  tends to 0 with probability 1 as  $n \nearrow \infty$ .

- **Sampling without replacement from zeta urn**

Let  $(\xi_{z, m}; m \geq 1)$  be an iid sequence of  $\text{zeta}(\beta, z)$  distributed random variables on  $\mathbb{N}_0$ , with mean  $\rho > 0$ . Let  $\zeta_{z, n} := \sum_{m=1}^n \xi_{z, n}$ ,  $n \geq 1$ . As noted above:

$$\mathbf{K}_{k, n} \stackrel{d}{=} (\xi_{z, 1}, \dots, \xi_{z, n} \mid \zeta_{z, n} = k) .$$

Assume the number of particles  $k$  is larger than  $n$ . We would like to extract a random sub-sample of size  $n$ , without replacement, from  $\mathbf{K}_{k, n}$ .

Let  $\mathbf{K}_n := (K_n(m), m = 1, \dots, n)$  be the number of occurrences of energy state  $m$  in this random size- $n$  sub-sample, with  $\sum_{m=1}^n K_n(m) = n$ . With  $(k_1, \dots, k_n) \in \mathbb{N}_0^n$  satisfying  $\sum_{m=1}^n k_m = n$ , the sampling without replacement strategy yields:

$$\begin{aligned} \mathbf{P}(K_n(1) = k_1, \dots, K_n(n) = k_n) &= \frac{1}{\{k\}_n} \frac{n!}{\prod_{m=1}^n k_m!} \mathbf{E} \left( \prod_{m=1}^n \{K_{k, n}(m)\}_{k_m} \right) \\ &= \frac{1}{\binom{k}{n}} \frac{[z^k] \prod_{m=1}^n D_\beta^{(k_m)}(z) / k_m!}{Z_{k, n}(\beta)}, \end{aligned}$$

where, in the second step, we used the expression of the falling factorial moments of  $\mathbf{K}_{k, n}$  displayed in Eq. (2.4). In the sub-sampling without replacement strategy, a knowledge of these moments is essential.

## 2.3 Frequency of frequencies: canonical approach

- **The number of non-empty states**

Let now  $P_{k,n} := \sum_{m=1}^n \mathbf{I}(K_{k,n}(m) > 0)$  count the number of non empty cells. With  $p \leq n \wedge k$ , using exchangeability of  $\mathbf{K}_{k,n}$ , with  $\mathbf{k}_p := (k_1, \dots, k_p) \in \mathbb{N}^p$  summing to  $k$

$$\mathbf{P}(K_{k,n}(1) = k_1, \dots, K_{k,n}(p) = k_p; P_{k,n} = p) = \binom{n}{p} \frac{\prod_{q=1}^p \sigma_{k_q}^{-\beta}}{Z_{k,n}(\beta)}$$

is the probability that  $p$  cells only are occupied with occupancy numbers  $\mathbf{k}_p$ . Thus

$$(2.10) \quad \begin{aligned} \mathbf{P}(P_{k,n} = p) &= \binom{n}{p} \frac{1}{Z_{k,n}(\beta)} \sum_{\mathbf{k}_p \geq 1: k_1 + \dots + k_p = k} \prod_{q=1}^p \sigma_{k_q}^{-\beta} \\ &= \binom{n}{p} \frac{[z^k] (D_\beta(z) - 1)^p}{[z^k] D_\beta(z)^n} \end{aligned}$$

and

$$\mathbf{P}(K_{k,n}(1) = k_1, \dots, K_{k,n}(p) = k_p \mid P_{k,n} = p) = \frac{\prod_{q=1}^p \sigma_{k_q}^{-\beta}}{[z^k] (D_\beta(z) - 1)^p}.$$

• **The number of states with prescribed number of particles**

This suggests to look at the frequency of frequencies distribution problem. For  $i = 0, \dots, k$ , let now

$$(2.11) \quad A_{k,n}(i) = \sum_{m=1}^n \mathbf{I}(K_{k,n}(m) = i)$$

count the number of cells visited  $i$  times by the  $k$ -sample, with  $A_{k,n}(0) = n - P_{k,n}$ , the number of empty cells. Let  $(a_0, a_1, \dots, a_k)$  be non-negative integers satisfying  $\sum_{i=0}^n a_i = n$  and  $\sum_{i=1}^n i a_i = k$ . Then

$$\mathbf{P}(A_{k,n}(0) = a_0, A_{k,n}(1) = a_1, \dots, A_{k,n}(k) = a_k) = \frac{n!}{Z_{k,n}(\beta)} \prod_{i=0}^k \frac{\sigma_i^{-\beta a_i}}{a_i!}.$$

Note from this that, with  $\sum_{i=1}^n i a_i = k$  and  $\sum_1^k a_i \leq n$ , the normalization condition gives

$$(2.12) \quad \sum_{a_1, \dots, a_k} \frac{1}{(n - \sum_1^k a_i)!} \prod_{i=1}^k \frac{\sigma_i^{-\beta a_i}}{a_i!} = \frac{Z_{k,n}(\beta)}{n!}.$$

From this, we get

**Proposition 3** *If  $p = n - a_0$ , the joint distribution of  $(A_{k,n}(1), \dots, A_{k,n}(k))$  and  $P_{k,n}$  reads*

$$(2.13) \quad \mathbf{P}(A_{k,n}(1) = a_1, \dots, A_{k,n}(k) = a_k; P_{k,n} = p) = \frac{\{n\}_p}{Z_{k,n}(\beta)} \prod_{i=1}^k \frac{\sigma_i^{-\beta a_i}}{a_i!}.$$

so that

$$\begin{aligned} \mathbf{P}(A_{k,n}(1) = a_1, \dots, A_{k,n}(k) = a_k \mid P_{k,n} = p) \\ = \frac{p!}{[z^k] (D_\beta(z) - 1)^p} \prod_{i=1}^k \frac{\sigma_i^{-\beta a_i}}{a_i!} \end{aligned}$$

Let us now compute the falling factorial moments of  $A_{k,n}(i)$ ,  $i = 1, \dots, k$ .

**Proposition 4** *Let  $r_i$ ,  $i = 1, \dots, k$  be non-negative integers satisfying  $\sum_1^k r_i = r \leq n$  and  $\sum_1^k i r_i = \kappa \leq k$ . We have*

$$(2.14) \quad \mathbf{E} \left[ \prod_{i=1}^k \{A_{k,n}(i)\}_{r_i} \right] = \{n\}_r \frac{Z_{k-\kappa, n-r}(\beta)}{Z_{k,n}(\beta)} \prod_{i=1}^k \sigma_i^{-\beta r_i}.$$

**Proof:**

$$\begin{aligned} \mathbf{E} \left[ \prod_{i=1}^k \{A_{k,n}(i)\}_{r_i} \right] &= \frac{n!}{Z_{k,n}(\beta)} \sum_{a_1, \dots, a_k} \frac{1}{(n - \sum_1^k a_i)!} \prod_{i=1}^k \left\{ \frac{\sigma_i^{-\beta a_i}}{(a_i - r_i)!} \right\} \\ &= \frac{n!}{Z_{k,n}(\beta)} \prod_{i=1}^k \sigma_i^{-\beta r_i} \sum_{a_1, \dots, a_k} \frac{1}{(n - \sum_1^k a_i)!} \prod_{i=1}^k \left\{ \frac{\sigma_i^{-\beta(a_i - r_i)}}{(a_i - r_i)!} \right\}. \end{aligned}$$

The normalization condition (2.12) gives:

$$\sum_{a_1, \dots, a_k} \frac{1}{(n - \sum_1^k a_i)!} \prod_{i=1}^k \left\{ \frac{\sigma_i^{-\beta(a_i - r_i)}}{(a_i - r_i)!} \right\} = \frac{Z_{k-\kappa, n-r}(\beta)}{(n-r)!}.$$

Finally, we get

$$\mathbf{E} \left[ \prod_{i=1}^k \{A_{k,n}(i)\}_{r_i} \right] = \{n\}_r \frac{Z_{k-\kappa, n-r}(\beta)}{Z_{k,n}(\beta)} \prod_{i=1}^k \sigma_i^{-\beta r_i}. \quad \square$$

In particular, if all  $r_i = 0$ , except for one  $i$  for which  $r_i = r$ , then

$$(2.15) \quad \mathbf{E} [\{A_{k,n}(i)\}_r] = \{n\}_r \frac{Z_{k-i, n-r}(\beta) \sigma_i^{-\beta r}}{Z_{k,n}(\beta)}.$$

If all  $r_i = 0$ , except for one  $i$  for which  $r_i = r = 1$ , then

$$(2.16) \quad \mathbf{E}[A_{k,n}(i)] = n \frac{Z_{k-i,n-1}(\beta) \sigma_i^{-\beta}}{Z_{k,n}(\beta)} = n \mathbf{P}(K_{k,n}(1) = i).$$

This shows that the expected number of cells visited  $i$  times is  $n$  times the probability that there are  $i$  visits to (say) cell one. From this, we more generally get

**Proposition 5** *In the thermodynamic limit, with  $\rho \in (0, \rho_c)$ :*

$$\frac{1}{n} A_{\lfloor n\rho \rfloor, n}(i) \xrightarrow{a.s.} \mathbf{P}(K_\rho = i) \text{ as } n \nearrow \infty.$$

**Proof:** In the thermodynamic limit  $\kappa_n = \lfloor n\rho \rfloor$ ,  $n \nearrow \infty$ ,  $\mathbf{K}_{\kappa_n, n}$  is asymptotically iid with components law:  $\mathbf{P}(K_\rho = l) = \frac{\sigma_l^{-\beta} \cdot z_\beta(\rho)^l}{D_\beta(z_\beta(\rho))}$ ,  $l \in \mathbb{N}_0$ . The above statement therefore follows from  $A_{k,n}(i) = \sum_{m=1}^n \mathbf{I}(K_{k,n}(m) = i)$  and the strong law of large numbers.  $\square$

### 3 Grand canonical occupancies

In this Section, we investigate the grand-canonical occupancy distributions.

• **Gibbs randomization of sample size.** Assume sample size  $k$  is now random, say  $K_{z,n}$ . Assume further that  $K_{z,n}$  has distribution:

$$\mathbf{P}(K_{z,n} = k) = \frac{z^k Z_{k,n}(\beta)}{D_\beta(z)^n}, \quad k \in \mathbb{N}_0.$$

The randomized version  $K_{z,n}$  of  $k$  has generating function

$$\mathbf{E}(u^{K_{z,n}}) = \left( \frac{D_\beta(zu)}{D_\beta(z)} \right)^n$$

and so  $K_{z,n}$  is a sum of  $n$  independent  $\text{zeta}(\beta, z)$  integral-valued random variables with common generating function  $D_\beta(zu)/D_\beta(z)$ .

• **Grand-canonical occupancies.** Now indexing cell occupancies by  $z$  rather than by  $k$ , we can define the joint laws of cell occupancies vector  $\mathbf{K}_{z,n} := (K_{z,n}(m); m = 1, \dots, n)$  and  $K_{z,n}$  as

$$\begin{aligned} \mathbf{P}(\mathbf{K}_{z,n} = \mathbf{k}_n; K_{z,n} = k) &= \frac{z^k Z_{k,n}(\beta)}{D_\beta(z)^n} \mathbf{P}(\mathbf{K}_{k,n} = \mathbf{k}_n) \\ &= \frac{z^k}{D_\beta(z)^n} \prod_{m=1}^n \sigma_{k_m}^{-\beta}. \end{aligned}$$

Observing that  $K_{z,n} = \sum_{m=1}^n K_{z,n}(m)$ , regardless of the event  $K_{z,n} = k$ , we get

**Proposition 6** For all  $k_m \in \mathbb{N}_0$ ,  $m = 1, \dots, n$

$$\mathbf{P}(\mathbf{K}_{z,n} = \mathbf{k}_n) = \prod_{m=1}^n \frac{z^{k_m} \sigma_{k_m}^{-\beta}}{D_\beta(z)} = \prod_{m=1}^n \mathbf{P}(\xi_{z,m} = k_m).$$

The law of the  $\xi_{z,m}$ s depends on  $z$ ; further,  $z$  and  $\kappa := \mathbf{E}(K_{z,n})$  are easily seen to be related through  $\kappa = nz \frac{D'_\beta(z)}{D_\beta(z)}$  (under this model, the expected number of particles is proportional to  $n$ ). In this interpretation,  $\mathbf{P}(K_{z,n}(m) = k_m) = \mathbf{P}(\xi_{z,m} = k_m)$  and the law of  $\mathbf{K}_{z,n}$  turns out to be a mere *product* measure. Note that, if  $P_{z,n} = \sum_{m=1}^n \mathbf{I}(K_{z,n}(m) > 0)$  denotes the number of occupied boxes, with  $\mathbf{k}_p = (k_1, \dots, k_p) \in \mathbb{N}^p$

$$\mathbf{P}(\mathbf{K}_{z,p} = \mathbf{k}_p; P_{z,n} = p) = \binom{n}{p} \mathbf{P}(\xi_{z,n} = 0)^{n-p} \prod_{q=1}^p \mathbf{P}(\xi_{z,q} = k_q)$$

and, with  $\mathbf{P}(\xi_{z,n} = 0) = 1/D_\beta(z)$ , for  $p \in \{0, \dots, n\}$

$$\mathbf{P}(P_{z,n} = p) = \binom{n}{p} \mathbf{P}(\xi_{z,n} = 0)^{n-p} \mathbf{P}(\xi_{z,n} > 0)^p$$

giving a binomial distribution for  $P_{z,n}$ . As required, we have  $\mathbf{P}(P_{z,n} = 0) = 1/D_\beta(z)^n = \mathbf{P}(K_{z,n} = 0)$ . Next,  $\mathbf{E}(P_{z,n}) = n \left(1 - \frac{1}{D_\beta(z)}\right)$ .

When  $n \nearrow \infty$ ,  $\beta \nearrow \infty$  while  $n2^{-\beta} = \gamma > 0$ , the binomial/Poisson approximation gives  $P_{z,n} \xrightarrow{d} P_z \stackrel{d}{\sim} \text{Poisson}(\gamma z)$ . We shall come back to this zero temperature weak  $*$ -limit later.

Before that, let us first reconsider the frequency of frequencies problem after having randomized sample size as described above. Let then

$$\begin{aligned} \mathbf{P}(A_{z,n}(1) = a_1, \dots, A_{z,n}(k) = a_k; K_{z,n} = k) &= \\ \frac{z^k Z_{k,n}(\beta)}{D_\beta(z)^n} \mathbf{P}(A_{k,n}(1) = a_1, \dots, A_{k,n}(k) = a_k) &= \\ n! \prod_{i=0}^k \left\{ \left( \frac{z^i \sigma_i^{-\beta}}{D_\beta(z)} \right)^{a_i} \frac{1}{a_i!} \right\} &= n! \prod_{i=0}^k \left\{ \frac{(\mathbf{P}(\xi_{z,1} = i))^{a_i}}{a_i!} \right\} \end{aligned}$$

be the joint grand-canonical multinomial probability of the event  $A_{z,n}(1) = a_1, \dots, A_{z,n}(k) = a_k; K_{z,n} = k$ . In other words, for all sequences  $(a_i; i \geq 0)$  satisfying the single constraint  $\sum_{i \geq 0} a_i = n$ :

$$\mathbf{P}(A_{z,n}(1) = a_1, \dots, A_{z,n}(i) = a_i, \dots) = n! \cdot \prod_{i=0}^{\sum_{l \geq 1} l a_l} \left\{ \frac{(\mathbf{P}(\xi_{z,1} = i))^{a_i}}{a_i!} \right\}$$

only depends on  $z$ . We have  $A_{z,n}(i) = \sum_{m=1}^n \mathbf{I}(K_{z,n}(m) = i)$ , so that

$$\begin{aligned} \mathbf{E}(A_{z,n}(i)) &= \sum_{m=1}^n \mathbf{P}(K_{z,n}(m) = i) = \sum_{m=1}^n \mathbf{P}(\xi_{z,m} = i) \\ &= n\mathbf{P}(\xi_{z,1} = i) = n \frac{z^i \sigma_i^{-\beta}}{D_\beta(z)}. \end{aligned}$$

The expected number of boxes with  $i$  particles is  $n$  times the probability that (say) box number 1 has  $i$  particles. Next, consistently,

$$\kappa := \mathbf{E}(K_{z,n}) = \sum_{i \geq 1} i \mathbf{E}(A_{z,n}(i)) = n \sum_{i \geq 1} \frac{iz^i \sigma_i^{-\beta}}{D_\beta(z)} = nz\Phi'_\beta(z).$$

**Proposition 7** (i) for all sequences  $(a_i; i \geq 0)$  satisfying the single constraint  $\sum_{i \geq 0} a_i = n$  :

$$\mathbf{P}(A_{z,n}(1) = a_1, \dots, A_{z,n}(i) = a_i, \dots) = n! \cdot \prod_{i=0}^{\sum_{l \geq 1} la_l} \left\{ \frac{(\mathbf{P}(\xi_{z,1} = i))^{a_i}}{a_i!} \right\}.$$

(ii) in the weak  $*$ -limit  $n \nearrow \infty, \beta \nearrow \infty$  while  $n2^{-\beta} = \gamma > 0, (A_{z,n}(i); i \geq 1)$  converges to a sequence of independent random elements with limit law  $A_z(1) \stackrel{d}{\sim} \text{Poisson}(\gamma z)$  and  $A_z(i) \stackrel{d}{\sim} \delta_0$  when  $i \geq 2$ .

(iii) in the weak  $*$ -limit, the number of visited boxes  $P_{z,n}$  converges to  $P_z \stackrel{d}{\sim} \text{Poisson}(\gamma z)$ .

**Proof:** it remains to prove (ii) and (iii). Observing  $D_\beta(z)^n \sim_* e^{\gamma z}$ , we have

$$\mathbf{P}(A_{z,n}(1) = a_1, \dots, A_{z,n}(i) = a_i, \dots) = \frac{n!}{(n - \sum_{i \geq 1} a_i)!} \prod_{i=1}^{\sum_{l \geq 1} la_l} \left\{ \frac{(\mathbf{P}(\xi_{z,1} = i))^{a_i}}{a_i!} \right\}$$

$$\sim_* \frac{n^{\sum_{i \geq 1} a_i}}{D_\beta(z)^n} \prod_{i \geq 1} \left\{ \frac{(z^i \sigma_i^{-\beta})^{a_i}}{a_i!} \right\} \sim_* \frac{(\gamma z)^{a_1} e^{-\gamma z}}{a_1!} \prod_{i \geq 2} \left\{ \frac{(0)^{a_i}}{a_i!} \right\}$$

$$=: \mathbf{P}_*(A_z(1) = a_1, \dots, A_z(i) = a_i, \dots).$$

In the limit, with the convention  $(0)^{a_i} = \mathbf{1}(a_i = 0)$ ,  $A_z(i) \neq 0$  with some positive probability only when  $i = 1$  (singletons) and  $\mathbf{P}_*(A_z(1) = a_1) = \frac{\lambda_1^{a_1} e^{-\lambda_1}}{a_1!}$  is Poisson with intensity  $\lambda_1 = \gamma z$ . This (low temperature, large number of boxes)

asymptotic regime is the one of uniques or singletons. States whose occupancies cannot exceed 1 are currently obtained in a Fermi-Dirac context. To prove (iii), recall that  $P_{z,n} \stackrel{d}{\sim} \text{Bin}\left(n, 1 - \frac{1}{D_\beta(z)}\right)$  with  $1 - \frac{1}{D_\beta(z)} \underset{\beta \nearrow \infty}{\sim} z2^{-\beta}$ , giving the Poisson weak  $*$ -limit already discussed. Note finally that the limiting number of particles is  $K_z \stackrel{d}{\sim} \text{Poisson}(\gamma z)$ .  $\square$

## References

- [1] Bialas, P.; Burda, Z.; Johnston, D. Condensation in the Backgammon model. Nucl. Phys. B 493, 505-516, **1997**.
- [2] Bialas, P.; Burda, Z.; Johnston, D. Phase diagram of the mean field model of simplicial gravity. Nuclear Physics B, Volume 542, Issue 1-2, p. 413-424, **1999**.
- [3] Bialas, P.; Bogacz, L.; Burda, Z.; Johnston, D. Finite size scaling of the balls in boxes model. Nuclear Physics B, Volume 575, Issue 3, p. 599-612, **2000**.
- [4] Charalambides, C. A. *Combinatorial methods in discrete distributions*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, **2005**.
- [5] Comtet, L. *Analyse combinatoire*. Tomes 1 et 2. Presses Universitaires de France, Paris, **1970**.
- [6] Flajolet, P.; Odlyzko A. Singularity analysis of generating functions. SIAM J. Disc. Math., 3, No 2, 216-240, **1990**.
- [7] Flajolet, P. Singularity analysis and asymptotics of Bernoulli sums. Theoret. Comput. Sci. 215, No. 1-2, 371-381, **1999**.
- [8] Godrèche, C.; Luck, J.M. Nonequilibrium dynamics of the zeta urn model. Eur. Phys. J. B 23, 473-486, **2001**.
- [9] Gut, A. Some remarks on the Riemann zeta distribution. Rev. Roumaine Math. Pures et Appl. 51, 205-217, **2006**.
- [10] Ijiri, Y.; Simon, H.A. Some distributions associated with Bose-Einstein statistics. Proc. Nat. Acad. Sc. USA, 72, No 5, 1654-1657, **1975**.
- [11] Kolchin, V. F. *Random mappings*. Translated from the Russian. With a foreword by S. R. S. Varadhan. Translation Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York, **1986**.

- [12] Nadarajah, S.; Kotz, S. A generalized Planck distribution. *Test*, 15, No 2; 111-126, **2006**.
- [13] Steutel, F. W.; van Harn, K. *Infinite divisibility of probability distributions on the real line*. Monographs and Textbooks in Pure and Applied Mathematics, 259. Marcel Dekker, Inc., New York, **2004**.