Complex distributions emerging in filtering and compression

G. J. Baxter,[1] R. A. da Costa,[1] S. N. Dorogovtsev,[1,2] and J. F. F. Mendes[1,3]

[1] *Departamento de Física da Universidade de Aveiro & I3N,*
Campus Universitário de Santiago, 3810-193 Aveiro, Portugal
[2] *A.F. Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia*
[3] *School of Computer and Communication Sciences and School of Life Sciences,*
École Polytechnique Fédéral de Lausanne, 1015 Lausanne EPFL, Switzerland

In filtering, each output is produced by a certain number of different inputs. We explore the statistics of this degeneracy in an explicitly treatable filtering problem in which filtering performs the maximal compression of relevant information contained in inputs. The filter patterns in this problem conveniently allow a microscopic, combinatorial consideration. This allows us to find the statistics of outputs, namely the exact distribution of output degeneracies, for relatively large input sizes. We observe that the resulting degeneracy distribution of outputs decays as $e^{-c\log^{\alpha}d}$ with degeneracy d, where c is a constant and exponent $\alpha > 1$, i.e. faster than a power law. Importantly, its form essentially depends on the size of the input data set, appearing to be closer to a power-law dependence for small data set sizes than for large ones. We demonstrate that for sufficiently small input data set sizes typical for empirical studies, this distribution could be easily perceived as a power law.

I. INTRODUCTION

Compression, filtering, and cryptography are related areas in signal and information processing. By definition, a large number of possible inputs are mapped to a smaller number of possible outputs, so that a given output may correspond to multiple inputs, which number is the output's degeneracy. A similar problem emerges in cooperative systems with a large number of local minima in the energy landscape, in particular, in spin glasses and deep learning neural networks. The configuration space of a system of this sort can be divided into a set of domains (basins) of attraction of these minima. One can ask: what is the statistics of these domains of attraction, what is the distribution of their sizes? This is analogous to the degeneracy statistics problem.

These issues were explored in a recent series of works [1–5] which exploited the principle of maximum entropy, see also Refs. [6–10]. The finding of Refs. [1–5] is that optimal compression generates outputs with broad distributions. More specifically, their entropy optimization based theory predicts power-law like distributions of degeneracy of maximally informative outputs (minimal sufficient representations). The title of the paper "Minimum description codes are critical" [2] published in the journal Entropy highlights the message of that theory in the most succinct way. On the other hand, distributions markedly distinct from power laws can be observed in empirical distributions for these and related problems, for example, a collection of empirical curves in Fig. 1 of Ref. [2].

In the present work we demonstrate how such diversity can emerge by analysing a simple representative problem. We introduce and explore a reference filtering problem straightforwardly treatable through purely combinatorial techniques. This filter performs the maximal compression of relevant information contained in inputs extracting all positions of a given local pattern in the input, see Fig. 1. The chosen simple filter pattern enables us to uncover a direct relation to a statistical physics problem, namely, the statistics of dimers on a chain. We directly obtain the complex degeneracy distributions for outputs generated from various input sets. We develop efficient recursive methods that allow us to find this distribution for large input strings, as can be seen in Fig.2, which would not be accessible through empirical sampling methods. Due to the tractability of this problem, we are able to identify precisely how these distributions deviate from a power-law dependence. We find scaling forms that accurately describe the tail of the degeneracy distributions (see Fig. 3). We discover that these distributions are essentially shaped by the size of the input data. That is, input data sets of relatively small size, typical for empirical studies, produce degeneracy distributions of outputs that are closer to a power-law than the distributions of outputs from very large input data

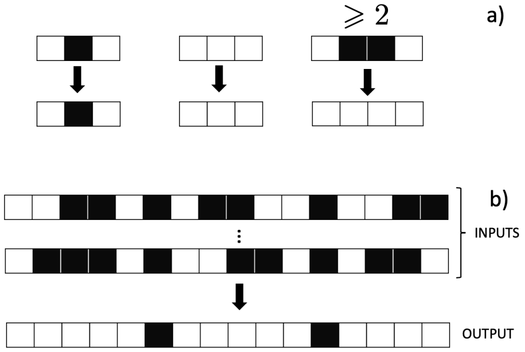

FIG. 1. (a) Filter extracting single ones and their positions from sequences of ones and zeroes. Here the filter pattern is a single one. Ones and zeroes are shown as black and white pixels, respectively. (b) A number of different inputs produce the same output. This number is the degeneracy of the output.

sets. Finally, we develop a mean field theory and obtain the asymptotics of the degeneracy distribution and the spectrum of degeneracies. Our findings indicate that the phenomena we observe should apply to more general filtering and compression problems.

Our paper is organized in the following way. In Sec. II we introduce our filter and the synthetic input data sets for it. In Secs. III, IV, and V we obtain the basic relations for the degeneracy of outputs, develop our algorithm, obtain the complex degeneracy distributions of outputs for the complete input data sets, and describe their features. We develop the mean-field theory of these distributions in filtering in Sec. VI. In Sec. VII we obtain the degeneracy distributions of outputs for uniformly random input data sets of various sizes. In Sec. VIII we discuss our results and indicate possible generalizations of our problem. The Appendix contains the combinatorial derivations of recursive relations used in Sec. III and explicit asymptotics. The exact results obtained by our algorithm and recursions are provided in the Supplementary Material [11].

II. A REFERENCE FILTERING PROBLEM

We study the distribution of outputs in a solvable filtering problem by implementing a purely combinatorial, microscopic approach not involving entropy considerations. Let the input data be a set of N strings of zeroes and ones (x_i), $x_i = 0, 1$, of length n, assuming the periodic condition $x_1 = x_{n+1}$. We consider two types of data set. The first set is the complete set of all possible unique inputs. Its size N is determined by the size n of inputs, $N = 2^n$. Second, we consider data sets of arbitrary size N consisting of strings of uniformly randomly generated zeroes and ones constrained by the same periodic condition as above. In the latter situation, some of the elements of a data set may coincide. Clearly, in the limit $N \to \infty$, we arrive at a situation equivalent to the complete data set. (We stress that the random data set of size $N = 2^n$ differs from the the complete data set.)

The filter works as follows: every instance of a specific pattern in the input is marked by a one in the corresponding position in the output. All other positions are marked with zeroes. This produces a minimal coding of the positions of the pattern occurrences in the input. For the sake of simplicity, we use the following filter. Each sequence of ones of length 1 in the input (i.e., every 1 whose neighbors are both zeroes) gives 1 at the same position in the output. All other sequences of ones or zeroes in the input produce zeroes in the corresponding places in the output, as shown in Fig. 1. In other words, the input vector (x_i), $i = 1, 2, ..., n$, $x_i = 0, 1$, is transformed to the output vector with the following components:

$$y_i = (1 - x_{i-1})x_i(1 - x_{i+1}). \tag{1}$$

Filters extracting more complex patterns may be constructed in a similar way.

III. OUTPUTS AND THEIR DEGENERACIES FOR COMPLETE INPUT DATASET

Let us begin by considering the complete input data set, all 2^n different configurations of zeroes and ones. The total number of generated outputs, $M(n)$, is the number of all possible combinations having no sequences of ones longer than one. This number coincides with the number of different configurations of dimers in a closed chain (ring) of length n:

$$M(n) = n \sum_{k=0}^{[n/2]} \frac{1}{n-k} \binom{n-k}{k}, \tag{2}$$

where $[n/2]$ is the integer part of $n/2$, see Appendix A 1. $M(n) = M(n-1) + M(n-2)$, $M(3) = 4$, $M(4) = 7$, which gives a sequence corresponding to the Lucas numbers [12–14]. The elements of the sequence may be written in terms of the roots of the characteristic equation

$$z^2 - z = 1. \tag{3}$$

Thus

$$M(n) = \Big(\frac{1+\sqrt{5}}{2}\Big)^n + \Big(\frac{1-\sqrt{5}}{2}\Big)^n \cong \Big(\frac{1+\sqrt{5}}{2}\Big)^n = z_g^n, \tag{4}$$

where the last expression gives the large n limit. Here the largest root $(1+\sqrt{5})/2 = 1.61803... \equiv z_g$ is the famous golden ratio.

Each output consists of isolated ones separated by strings of zeroes of various lengths. The key point for our study is that the degeneracy of an output is the product of the degeneracies of these strings of zeroes between ones in this output. As is clear from Eq. (1), the presence of a 1 in an output at position i fixes the digits of its inputs at positions $i-1$, i, and $i+1$. These digits must be 0, 1, and 0, respectively. This is true for each 1 in the output. On the other hand, each of the remaining digits of these inputs compatible with a given output must be either a 0 or be a 1 with one or two neighboring ones. All the degrees of freedom in the input corresponding to a given output correspond to these digits. Thus the degeneracy of a given output is given by the product of the degeneracies the strings of zeroes lying between the ones.

Let an output with $m \geq 1$ ones contain m strings of zeroes with lengths $\ell_1, \ell_2, ..., \ell_m$. Then the degeneracy of this output equals

$$d = \prod_{i=1}^{m} \tilde{d}(\ell_i). \tag{5}$$

Here $\tilde{d}(\ell)$ is the number of input strings of length ℓ, having the first and last digits 0, that generate an output string of ℓ zeroes. This number plays an important role in our problem, similar to prime numbers in number theory, so we call the $\tilde{d}(\ell)$ *prime degeneracies.* Below we shall find these prime degeneracies explicitly. The case

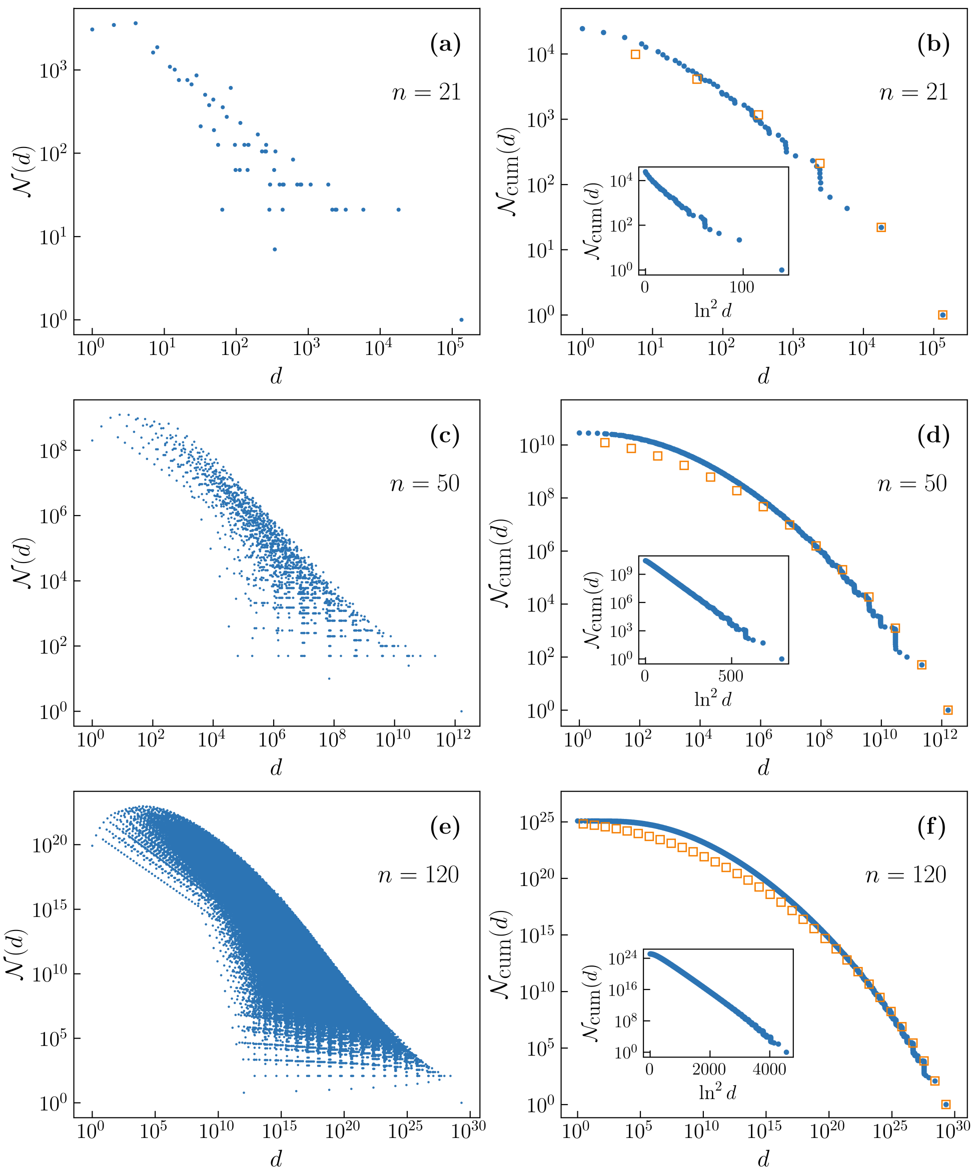

FIG. 2. (a,c,e) Degeneracy distribution for the complete input data set: number of outputs of a given degeneracy vs. degeneracy. (b,d,f) Cumulative degeneracy distribution for the complete input data set: number of outputs of degeneracy greater than or equal to a given degeneracy vs. degeneracy. The input length is $n = 21, 50, 120$.

$m = 0$ is special due to the periodicity of the inputs and outputs. Suppose that the output contains μ_ℓ strings of zeroes of length ℓ, $\ell = 1, 2, ...,$ where

$$m + \sum_{\ell \geq 1} \ell \mu_\ell = n. \tag{6}$$

Then Eq. (5) may be rewritten

$$d = \prod_{\ell \geq 1} [\tilde{d}(\ell)]^{\mu_\ell} \tag{7}$$

for $m \geq 1$.

Let us obtain $\tilde{d}(\ell)$. We consider input strings whose first and last digits are 0 and produce the output string of ℓ zeroes. The number of these configurations, i.e., the degeneracy $\tilde{d}(\ell)$ of the output string of ℓ zeroes, can be obtained recursively in the following way. We take into account three points.

(i) Relevant input configurations of length ℓ are obtained by inserting 0 or 1 into each relevant configuration of length $\ell - 1$ between the first and second positions of the sequence. (Recall that the first and last positions of the input sequence are fixed to 0.)

(ii) Input strings of length ℓ beginning and/or ending with 010 are irrelevant, and so they should be removed from the set generated at the previous step. These configurations can be obtained by inserting two digits 10 into each relevant input string of length $\ell - 2$ between its first and second positions.

(iii) Finally, there exist input strings, compatible with the output string of ℓ zeroes, that cannot be obtained by inserting a single digit into relevant input strings of length $\ell - 1$ between their first and second positions. These are the input strings of length ℓ beginning with 0110. These inputs can be obtained by inserting 110 into each relevant input string of length $\ell - 3$ between their first and second positions.

Following these rules, the degeneracy of a string of ℓ zeroes at the output, prime degeneracy $\tilde{d}(\ell)$, can be written recursively as a linear difference equation:

$$\tilde{d}(\ell) = 2\tilde{d}(\ell - 1) - \tilde{d}(\ell - 2) + \tilde{d}(\ell - 3) \tag{8}$$

with the initial condition $\tilde{d}(1) = \tilde{d}(2) = \tilde{d}(3) = 1$. Alternatively, applying this equality to the term with $\tilde{d}(\ell - 1)$ on the right-hand side of this equation we arrive at the equivalent difference equation

$$\tilde{d}(\ell) = \tilde{d}(\ell - 1) + \tilde{d}(\ell - 2) + \tilde{d}(\ell - 4) \tag{9}$$

with the initial condition $\tilde{d}(0) = 0$, $\tilde{d}(1) = \tilde{d}(2) = \tilde{d}(3) = 1$. The solution of Eq. (8) is explicitly expressed in terms of the complex roots z_1, z_2, and z_3 of the characteristic equation

$$z^3 = 2z^2 - z + 1, \tag{10}$$

given initial conditions $\tilde{d}(1) = \tilde{d}(2) = \tilde{d}(3) = 1$:

$$\tilde{d}(\ell) = C_1 {z_1}^\ell + C_2 {z_2}^\ell + C_3 {z_3}^\ell, \tag{11}$$

where

$$\begin{aligned} C_1 &= (z_1 - 1)/[(z_1 - z_2)(z_1 - z_3)], \\ C_2 &= (z_2 - 1)/[(z_2 - z_1)(z_2 - z_3)], \\ C_3 &= (z_3 - 1)/[(z_3 - z_1)(z_3 - z_2)]. \end{aligned} \tag{12}$$

One of the roots of Eq. (10), z_1, say, is real,

$$z_1 \equiv z_d = 1.75488\ldots. \tag{13}$$

It determines the large ℓ asymptotics of prime degeneracies $\tilde{d}(\ell)$:

$$\tilde{d}(\ell) \cong \frac{z_d}{z_d^4 - 2}\, z_d^\ell. \tag{14}$$

Here we used the identity $C_1 = z_d/(z_d^4 - 2)$. The two other roots are complex conjugate numbers,

$$z_{2,3} = 0.122561... \pm 0.744862...\, i. \tag{15}$$

The special case of the periodic output of length n with all digits 0 has to be considered separately. First let us take an arbitrary digit of the input. The number of input configurations where this digit is 0 and the resulting output has only zeroes is given by $\tilde{d}(n + 1)$, because the fixed 0 of the periodic input plays the role of first and last digit of the configurations of a string of $n + 1$ digits. If the chosen digit is 1, then the number of the relevant input configurations equals $1 + \sum_{i=2}^{n-1} i\tilde{d}(n - i)$, where the sum over i accounts for the configurations where the digit is in a group of i consecutive ones, plus one configuration with all input digits equal to 1. Consequently, we obtain the following expression for the degeneracy of the output with all zeroes:

$$d_D(n) = 1 + \tilde{d}(n + 1) + \sum_{i=2}^{n-1} i\tilde{d}(n - i), \tag{16}$$

which is the largest possible degeneracy of an output of a given length. Applying the recursion relation for prime degeneracies $\tilde{d}$, Eq. (8) [or (9)], to the terms on the right-hand side of Eq. (16) we find that the largest degeneracy $d_D(n)$ satisfies the same difference equation as Eq. (8) [or, equivalently, Eq. (9)] though with different initial condition, see, e.g., Ref. [15] and the On-line Encyclopedia of Integer Sequencies [16]. We present this equation here for future reference,

$$d_D(n) = 2d_D(n - 1) - d_D(n - 2) + d_D(n - 3) \tag{17}$$

with the initial condition $d_D(3) = 5$, $d_D(4) = 10$, and $d_D(5) = 17$ or, equivalently, $d_D(0) = 3$, $d_D(1) = 2$, and $d_D(2) = 2$. Here, of course, $D = D(n)$. With these initial condition, we get the explicit solution of this equation,

$$d_D(n) = {z_1}^\ell + {z_2}^\ell + {z_3}^\ell, \tag{18}$$

where $z_1 \equiv z_d, z_2, z_3$ are given by Eqs. (13) and (15), and its large n asymptotics

$$d_D(n) \cong z_d^n. \tag{19}$$

IV. CALCULATING THE EXACT DEGENERACY SPECTRUM

Let us explore the outputs generated by the complete input data set, 2^n inputs. In principle, we can find them directly. That is, for each of these inputs, one by one, we can obtain an output numerically by applying Eq. (1). It is convenient to treat each output as an n-digit binary number $x_1x_2...x_n$. This enables us to sort all generated outputs and easily find the degeneracy of each output. In practice, we use a more efficient algorithm described below. This algorithm focuses on outputs with a fixed number of ones and exploits the factorization of the output degeneracies, see Eq. (5).

Let us first describe how to generate the full list of degeneracies. This can be found from integer partitions in an explicit form, as follows. From Eqs. (8) and (14), we know the prime degeneracy $\tilde{d}(\ell)$ and $d_D(n)$, respectively. Let us introduce the operator $\mathcal{P}$, which generates all integer partitions of positive integer k into r integers, that is, $\mathcal{P}(k, r)$ is the matrix

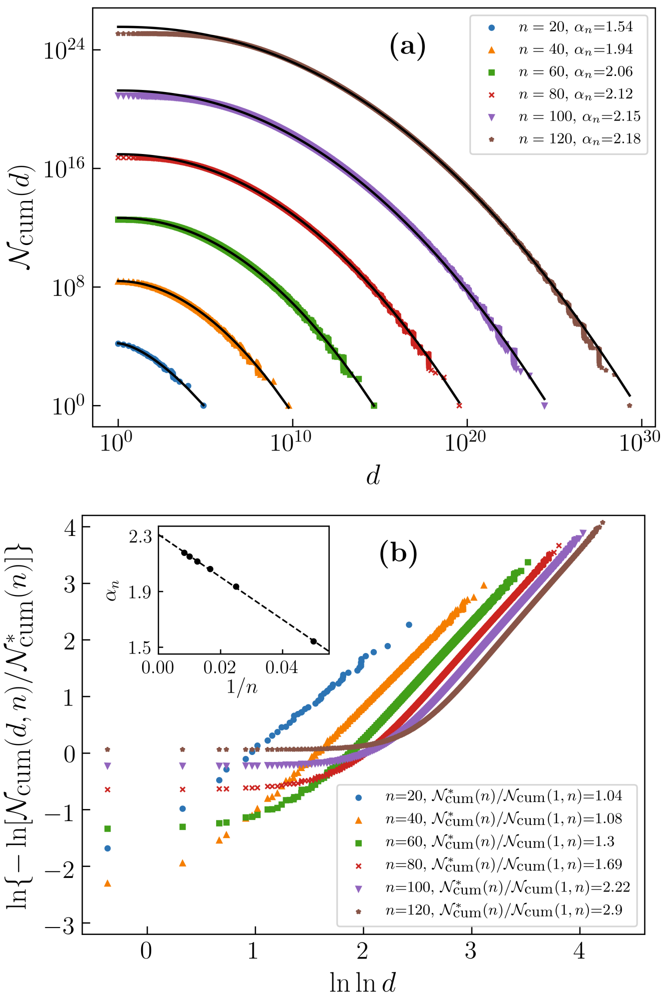

FIG. 3. (a) Cumulative degeneracy distribution for $n = 20, 40, 60, 80, 100, 120$. The black curves represent least-squares fittings of $\ln \mathcal{N}_{\mathrm{cum}}(d,n)$ as $\ln \mathcal{N}^*_{\mathrm{cum}}(n) + B_n \ln^{\alpha_n} d$ for each n. (b) Cumulative degeneracy distribution $\ln\{-\ln[\mathcal{N}_{\mathrm{cum}}(d,n)/\mathcal{N}^*_{\mathrm{cum}}(n)]\}$ vs. $\ln\ln d$ for $n = 20, 40, 60, 80, 100, 120$. Inset: exponent α vs. $1/n$.

whose rows $i = 1, 2, ..., r$ are all different integer partitions $\{\mathcal{P}_{i1}(k,r), \mathcal{P}_{i2}(k,r), ..., \mathcal{P}_{ij}(k,r), ..., \mathcal{P}_{ir}(k,r)\}$ of k, $\sum_{j=1}^r \mathcal{P}_{ij}(k,r) = k$ [17, 18]. For an output of length n, we consider all partitions of $n-m$ into m integers , for all possible $m = 1, 2, ..., [n/2]$. For each such partition, [i.e., a row of $\mathcal{P}(n-m,m)$], we find the degeneracy $d_i = \prod_j \tilde{d}(\mathcal{P}_{ij}(n-m,m))$ [see Eq. (5)]. Some of them coincide, so we find the union of them. Finally, we add the largest degeneracy $d_D(n)$ (corresponded to $m = 0$) to the resulting set. In summary, for the full set of degeneracies $\mathcal{D}_{\mathrm{full}}(n)$, we have

$$\mathcal{D}_{\mathrm{full}}(n) = d_D(n) \bigcup \left\{ \bigcup_{m=1}^{[n/2]} \left[\bigcup_i \prod_j \tilde{d}(\mathcal{P}_{ij}(n-m,m)) \right] \right\} \tag{20}$$

with $\tilde{d}(\ell)$ and $d_D(n)$ provided by Eqs. (8) and (14), respectively.

For each integer partition, i.e., for each row i of the matrix $\mathcal{P}(n-m,m)$ we introduce the number $\mu^{(i)}(\ell; n-m, m)$ of pieces of length ℓ present in this partition:

$$n - m = \sum_\ell \ell \mu_\ell^{(i)}(n-m,m),$$
$$m = \sum_\ell \mu_\ell^{(i)}(n-m,m). \tag{21}$$

So one can write

$$d_i = \prod_j \tilde{d}(\mathcal{P}_{ij}(n-m,m)) = \prod_\ell [\tilde{d}(\ell)]^{\mu_\ell^{(i)}(n-m,m)}. \tag{22}$$

The number of outputs that contain m chains of zeros with lengths specified by the integer partition $\mathcal{P}_i(n-m,m)$ is then obtained by considering the number of distinct permutations of the m strings of zeros, multiplying by n, and finally dividing by m, giving

$$\mathcal{N}[\mathcal{P}_i(n-m,m)] = n \frac{(m-1)!}{\prod_\ell \mu_\ell^{(i)}(n-m,m)!}, \tag{23}$$

where the product in the denominator is over the lengths of the parts of this partition. The total number of outputs with degeneracy d is then finally:

$$\mathcal{N}(d,n) = \delta[d, d_D(n)]$$
$$+ \sum_{m=1}^{[n/2]} \sum_i n \frac{(m-1)!}{\prod_\ell \mu_\ell^{(i)}(n-m,m)!} \, \delta\left[d, \prod_\ell [\tilde{d}(\ell)]^{\mu_\ell^{(i)}(n-m,m)}\right], \tag{24}$$

where $\delta(a,b)$ is the Kronecker symbol. This is the expression we use for computing $\mathcal{N}(d)$ in an efficient way. For the sake of brevity, hereafter we refer to $\mathcal{N}(d)$ as a degeneracy distribution, without normalization.

The distribution $\mathcal{N}(d,n)$ can also be built up recursively starting from small values of d and n, as we show in Appendix A 2. This technique is valid for any finite d and n.

V. OUTPUT DEGENERACY DISTRIBUTION FOR COMPLETE INPUT DATASETS

Using this algorithm we obtained the number of outputs $\mathcal{N}(d)$ for the full spectrum of degeneracies d for n up to 120, see Supplementary Material [11]. These results demonstrate that the degeneracies d_i , $i = 1, ..., D$, form a discrete spectrum of values where d_D is the largest degeneracy, and $d_1 = 1$.

Figure 2 shows the resulting distribution and the corresponding cumulative distribution $\mathcal{N}_{\mathrm{cum}}(d)$ for $n = 21$, 50, and 120. Here $\mathcal{N}_{\mathrm{cum}}(d_i) \equiv \sum_{j=i}^D \mathcal{N}(d_j)$. In particular, $\mathcal{N}_{\mathrm{cum}}(d_1) = M(n)$, i.e., the total number of outputs.

Figures 2(b), (c), and (d) demonstrate that the cumulative degeneracy distributions decay with d more rapidly than a power law. On the other hand, the decay of the cumulative distribution is well described by the function

$$\mathcal{N}_{\rm cum}(d) \propto e^{-c\ln^{\alpha} d} = d^{-c\ln^{\alpha-1} d}, \qquad (25)$$

where c is a positive number and exponent α approaches 2.3 as $n \to \infty$, see the inset in Fig. 3 (b). Notice that the degeneracy distribution for smaller n, Fig. 2 (a), appears more like a power law than the degeneracy distribution for larger n, Fig. 2(c), for example, since the range of d increases with n, while the exponent $c\ln^{\alpha-1}d$ of the function $d^{-c\ln^{\alpha-1}d}$ varies slowly. Similarly, the cumulative distribution plotted in log-log scale for $n = 21$ deviates from linear (power law behavior) noticeably less than for larger n. The wide range of degeneracies d that we observe enable us to present Fig. 3 (b) showing $\ln\{-\ln[\mathcal{N}_{\rm cum}(d)/\mathcal{N}_{\rm cum}(1)]\}$ vs. $\ln\ln d$. This plot supports the functional form given in Eq. (25). Note that in Fig. 3 we assumed that the coefficient factor of the asymptotics in Eq. (25) is close to $\mathcal{N}_{\rm cum}(1)$, which is justified by the results. Figure 4 shows how the set of degeneracies $\mathcal{D}_{\rm full}(n)$ varies with n, see below for more detail.

In Sec. III we derived the explicit expression for the largest degeneracy $d_D(n)$ corresponding to the output with all zeroes, Eq. (18), and found its large n asymptotics, $d_D(n) \cong z_d^n$, Eq. (19). As is natural, $\mathcal{N}(1,n) = 1$. The second largest degeneracy corresponds to an output with a single 1. The third largest degeneracy is for an output with two ones separated by a single 0. The fourth largest degeneracy is of the output with two ones separated by three 0. Clearly, $\mathcal{N}(D-1,n) = \mathcal{N}(D-2,n) = \mathcal{N}(D-3,n) = n$. Using the asymptotics of $\tilde{d}(\ell)$, Eq. (14), we find that, asymptotically, at large n,

$$d_{D-1}(n) \cong \frac{z_d^n}{7.48391...} = \frac{z_d^n}{z_d^4-2},$$

$$d_{D-2}(n) \cong \frac{z_d^{n-2}}{z_d^4-2},$$

$$d_{D-3}(n) \cong \frac{z_d^{n-3}}{z_d^4-2}. \qquad (26)$$

One may also notice a complex structure in the cumulative distributions Figs. 2(b), (d), and (f), $\mathcal{N}_{\rm cum}(d)$ resembling a staircase, with steep jumps between steps. The heights of these jumps are especially large in the region of high degeneracies. Inserting the asymptotics of $\tilde{d}(\ell)$, Eq. (5), into the expression for the degeneracy, Eq. (14), we see that outputs with the same number of ones have degeneracies exponentially close to $z_d^n/(z_d^4-2)^2$ if these ones are separated by many zeroes and n is large. The slight deviations from this asymptotic value mean that the degeneracies are split among many points falling in a narrow range. For example, in the rightmost step, corresponding to outputs with exactly two ones, the number of these outputs is $n^2/2$ because there are relatively few outputs in which the two ones are close together. This is the height of this jump. If n is finite, these degeneracies are split into a set of about $n/2$ distinct values, each one for n outputs corresponding to the location of the same structure in different parts of the ring. Other jumps are produced by outputs with m strongly separated ones, or by outputs with, e.g., a pair of ones separated by one 0 with $m-2$ ones far from each other and from that pair, and so on. This forms the rich staircase-like structure that we observe in Figs. 2(b), (d), and (e) and that is also reflected in Fig. 4.

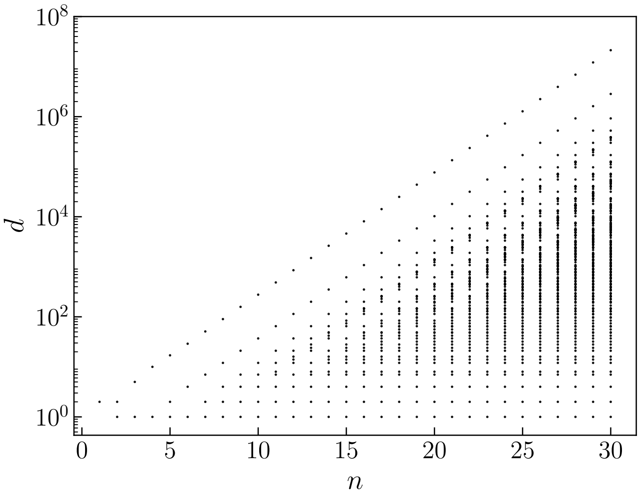

FIG. 4. The set of degeneracy values d for each n (logarithmic scale). (The range of n was selected to make individual points visible.)

Note that any filter of our sort for a finite input pattern will produce a similar complex structure in the degeneracy distribution, as the degeneracy associated to a chain of zeroes (in output) of a given length must asymptotically grow exponentially with the length of the chain, as in Eq. (19). We stress that Eq. (19) is asymptotic and not exact.

In Appendix A 2 we derive the chain of linear coupled recursion relations for the number of outputs of degeneracy d, $\mathcal{N}(d,n)$, see Eqs. (A3), (A5) and (A12), (A13). These recursions generate exact $\mathcal{N}(d,n)$ for finite d and n, and in this sense provide the exact full spectrum of degeneracies. (There recursions also provide us with the explicit leading large n asymptotics of $\mathcal{N}(d,n)$, Eq. (A14), for an arbitrary d.) In particular, Eqs. (A3) with the initial conditions, Eq. (A5), provide $\mathcal{N}(1,n)$ which is the number of configurations having only groups of zeroes of length 1, 2, and 3 between single ones. We find the large n asymptotics for this number:

$$\mathcal{N}(1,n) \cong z_a^n, \qquad (27)$$

where

$$z_a = 1.46557... \qquad (28)$$

is the real root of the characterstic equation $z^4 = z^2 + z + 1$.

One should note that there are three key constants in this problem, namely z_g, z_d, and z_a, which appear in the main asymptotics: $M(n) \cong z_g^n$, $d_D(n) \cong z_d^n$, and $\mathcal{N}(1, n) \cong z_a^n$. Indeed, z_g enters the distribution $\mathcal{N}(d, n)/M(n)$ after normalization, z_d enters the asymptotics for degeneracies, see, e.g., Eq. (14) for $\tilde{d}(\ell)$ and (26) for $d_i(n)$, and z_a enters the asymptotics for $\mathcal{N}(d, n)$, Eqs. (A10) and (A14). In fact, all moments of $\mathcal{N}(d, n)$ exponentially diverge with n, $\sum_{i=1}^{D} d_i^k \mathcal{N}(d_i, n) \cong c_k^n$, where the numbers c_k and their large k asymptotics are presented in Appendix B. Clearly, for the first moment we have $\sum_{i=1}^{D} d_i \, \mathcal{N}(d_i, n) = 2^n$.

Using Eq. (20) we obtain the full set of degeneracies $\mathcal{D}_{\text{full}}(n)$ for different n, shown in Fig. 4, and the series of total number of discrete values of degeneracy $D(n)$ versus input size, represented in Fig. 5, see also the Supplementary Material [11]. As is natural, $D(n)$ is smaller than the number $p(n)$ of integer partitions of n. Figure 5 demonstrates that $D(n)$ is close to $p(n)/n$ for large n. The well known large n asymptotics $p(n) \cong \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{2n/3}}$ [19–21] enables us to estimate $D(n) \sim e^{\pi\sqrt{2n/3}}$.

By ranking the full set of degeneracies $\mathcal{D}_{\text{full}}(n)$ for $n = 120$ we arrive at Fig. 6, where the main plot presents the number of different degeneracies lower than or equal to d vs. d, increasing roughly as $\exp[\pi\sqrt{2 \ln d / 3 \ln z_d}]$ in the region $1 \ll d \ll d_D$, and the inset shows the number of different degeneracies higher than or equal d vs. d. The latter demonstrates a staircase-like structure similar to that of $\mathcal{N}_{\text{cum}}(d)$.

The set of degeneracies occurring in the infinite system ($n \to \infty$) can be obtained from the combinations of integer powers of the prime degeneracies $\tilde{d}(\ell)$ for all ℓ. In Fig. 6 we plot the number of different degeneracies lower than or equal to d in the infinite system, i.e., the rank of each degeneracy d. To obtain the full set of different degeneracies for $n \to \infty$ up to some $d_{\max}$ we generate the set of prime degeneracies $\tilde{d} \leq d_{\max}$. We start by listing all powers $m \geq 1$ of the first prime degeneracy $\tilde{d}$ while $\tilde{d}^m \leq d_{\max}$. Then, we multiply each member of that list by increasing powers of the next prime degeneracy, while the product stays lower or equal to $d_{\max}$, and so on with all the remaining prime degeneracies $\tilde{d} \leq d_{\max}$. This procedure will result in duplicate degeneracies that should be removed, particularly for those values of d that have non-unique multiplicative partitions in terms of the prime degeneracies.

For example, the two smallest (larger than 1) prime degeneracies are $\tilde{d}(4) = 2$ and $\tilde{d}(5) = 4$, so a multiplicative partition of d with at least a part equal to $\tilde{d}(5) = 4$ is not unique because there is at least another partition of d where the contribution of the 4 for given in terms $\tilde{d}(4)$ as 2^2. However, apart from the partitions with parts equal to 4, the non-unique partitions are very rare, see Appendix A 2. In fact we were able to check that $\tilde{d}(4) = 2$ is the only prime degeneracy $\tilde{d}(\ell)$ that can be expressed as a product of lower $\tilde{d}$ for all $\ell \leq 5000$.

To explain why this property of primality is very likely to hold for all large enough $\tilde{d}(\ell)$, let us consider the number of conventional prime numbers smaller than some number d, which grows as $\sim d/\ln d$. Since $\tilde{d}(\ell) \sim z_d^\ell$, the number of conventional primes smaller than $\tilde{d}(\ell)$ grows exponentially with ℓ as $\sim z_d^\ell/\ell$. Additionally, the average number of conventional primes in the factorization of numbers of the magnitude of d grows as $\sim \ln\ln d$ [22]. A necessary condition for some $\tilde{d}(\ell)$ to not be prime is that all of its prime factors are also factors of at least another $\tilde{d} < \tilde{d}(\ell)$. The combination of the exponential increase of conventional primes smaller than $\tilde{d}(\ell)$, and the increase of the number of factors of $\tilde{d}(\ell)$ as ℓ increases, makes the probability of the degeneracy $\tilde{d}(\ell)$ not being prime approach 0 very rapidly. In the set of values $\tilde{d}(\ell \leq 200)$, the only ones that do not contain at least one conventional prime factor absent from all factorizations of smaller $\tilde{d}$ are $\tilde{d}(5) = 4$, $\tilde{d}(8) = 21$, $\tilde{d}(12) = 200$, $\tilde{d}(13) = 351$, and $\tilde{d}(24) = 170\,625$, and from these only $\tilde{d}(5)$ is actually not a prime degeneracy.

A similar argument can be used to derive recursion relations that allow the calculation of $\mathcal{N}(d, n)$ up to any finite n and d. We explain this calculation in Appendix A 2.

The rank of the degeneracies in the infinite system can be estimated with precision for large d, as shown by the black line in Fig. 6. For large ℓ the logarithms of prime degeneracies are uniformly distributed $\ln \tilde{d}(\ell) \approx \ell \ln z_d$. Assuming that all degeneracies d have a unique factorization in terms of the prime degeneracies $\tilde{d}$, the expect number of different degeneracies smaller than $d \sim z_d^m$ would be rank $d \approx \sum_{n \leq m} p(n)$, where $p(n)$ is the number of integer partitions of n. The vast majority of the degeneracies for which the assumption does not hold are the values of d which are multiples of 4, because that factor 4 can also be expressed as 2^2. We therefore remove from each term of the previous sum the number of integer partitions that have at least one factor equal to 2, which is given by $p(n-2)$, and write for large d:

$$
\begin{aligned}
\text{rank } d &\approx \sum_{n \leq \ln d/\ln z_d} [p(n) - p(n-2)] \\
&\approx p(\ln d/\ln z_d) + p(\ln d/\ln z_d - 1) \\
&\approx \frac{\exp\left(\pi\sqrt{2/3}\sqrt{\ln d/\ln z_d}\right)}{\pi 2\sqrt{3}(\ln d/\ln z_d)} \left[1 - \left(\frac{13\pi}{72} + \frac{1}{\pi}\right)\sqrt{\frac{3 \ln z_d}{2 \ln d}}\right],
\end{aligned}
\tag{29}
$$

where we have used the two leading terms of the asymptotic expansion of the number of integer partitions [19]

$$
p(n) \approx \frac{\exp\left(\pi\sqrt{2/3}\sqrt{n}\right)}{4\sqrt{3}n} \left[1 - \left(\frac{\pi}{72} + \frac{1}{\pi}\right)\sqrt{\frac{3}{2n}}\right]. \tag{30}
$$

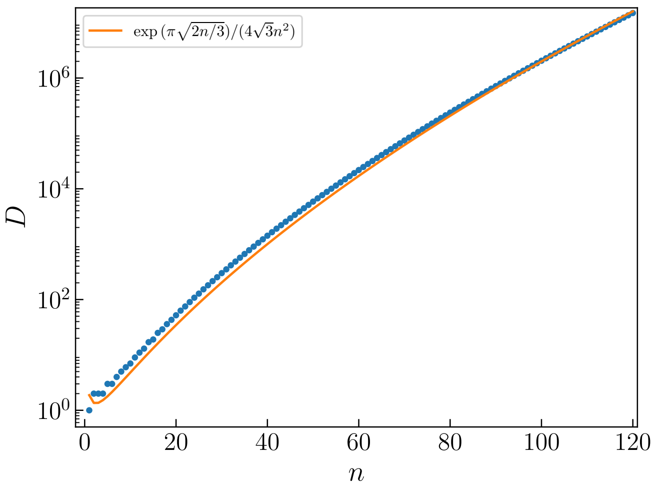

FIG. 5. Total number of different degeneracies D vs. n (symbols). The curve is the number of integer partitions of n divided by n.

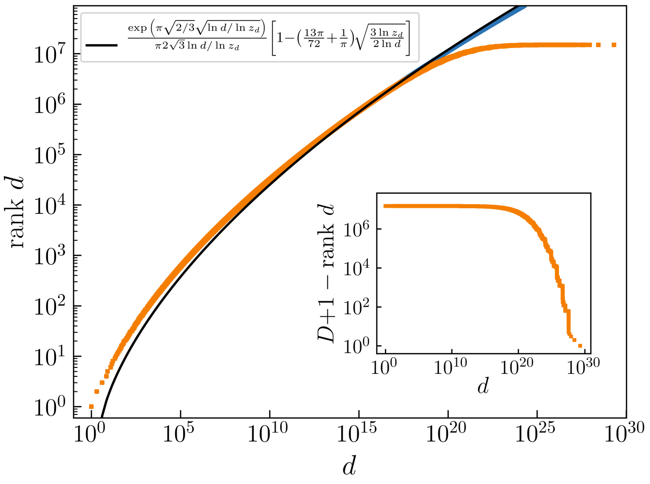

FIG. 6. Number of distinct degeneracies less than or equal to d vs. d. Inset: number of different degeneracies higher than or equal d vs. d. The orange square symbols are full results for $n = 120$. The circular blue symbols are values calculated in the limit $n \to \infty$ as described in Sect. V. Solid black curve is the asymptotic approximation Eq. (29).

Fig. 6 demonstrates that rank d as a function of d is well described by the estimate Eq. (29).

VI. MEAN-FIELD THEORY

Here we develop a mean-field theory enabling us to describe the cumulative distribution of degeneracies, $\mathcal{N}_{\rm cum}(d)$, in the region of large d where there are few ones in the outputs, and so the gaps of zeroes between them are typically large. In this situation one can assume that the ones exist in a sea (or a mean field) of zeroes, far from each other, so that the degeneracy of an output is completely determined by the number of its ones (and the output size n).

This ansatz is based on the observation that the three terms on the right-hand side of Eq. (11) behave very differently for increasing ℓ: The first term grows exponentially as z_d^ℓ, where $z_d \equiv z_1 = 1.754877...$ is real. The combined contribution of the other two terms is also real, since $z_2 = z_3^*$ (here * denotes the complex conjugate). It decays exponentially because $|z_2| = |z_3| = 0.754877... < 1$. Thus, for increasing ℓ, the deviation of the asymptotics

$$\tilde{d}(\ell) \cong C_1 {z_d}^\ell, \tag{31}$$

where the coefficient

$$C_1 = \frac{z_1 - 1}{(z_1 - z_2)(z_1 - z_3)} = \frac{z_d}{z_d^4 - 2} = \frac{1}{4.26463...} \tag{32}$$

from the exact value of $\tilde{d}(\ell)$ approaches 0 exponentially rapidly, making this approximation excellent for large ℓ. The simplicity of Eq. (31) enables us to find the degeneracy of outputs whose strings of zeroes are large. These are the typical outputs when there are sfew ones. Consider an output with m ones, separated by m strings of zeroes of length ℓ_1, ℓ_2, ..., ℓ_k. If all $\ell_i \gg 1$ the degeneracy $\overline{d}(m)$ of this output with m ones is accurately given by the asymptotic expression

$$\overline{d}(m) = \prod_{i=1}^{m} C_1 {z_d}^{\ell_i} = {C_1}^m {z_d}^{n-m}, \tag{33}$$

where we used the condition $\sum_{i=1}^m \ell_i = n - m$ that the number of ones plus the number of zeroes must be equal to the total number of digits n.

The total number of different outputs with m ones is

$$\overline{\mathcal{N}}(m) = \frac{n}{n-m}\binom{n-m}{m}, \tag{34}$$

see Eq. (A1) in Appendix A 1. Notice that $\overline{\mathcal{N}}(m)$ coincides with the m-th term of the sum in Eq. (2), as it must.

One should stress that although we derived the expressions for $\overline{d}(m)$ and $\overline{\mathcal{N}}(m)$ for $m > 0$, these expressions are equally valid in the special case of $m = 0$. Indeed, Eq. (34) gives $\overline{\mathcal{N}}(0) = 1$, which is exact. Further, Eq. (33) gives $\overline{d}(0) \cong z_d^n$ for large n coinciding with $d_D \cong {z_d}^n$, Eq. (19). Thus Eqs. (33) and (34) work very well also for $m = 0$. For $m = 1$ we find a similar situation in terms of accuracy. In this case there is a single string of zeroes of size $n - 1$, then every output with a single 1 has degeneracy equal to $\tilde{d}(n-1)$. The deviation between $\overline{d}(1) = C_1 z_1^{n-1}$ and $\tilde{d}(n-1)$ decays exponentially with n, making the estimation $\overline{d}(1)$ to $d(1)$ extremely accurate. For each value of $m > 1$ there is actually more than one possible value for the degeneracy, that depends on the particular distribution of lengths of the strings of zeroes. Nonetheless, for large n and $m \ll n$ most of the outputs contain only large strings of zeroes. Therefore the set of points $(\overline{d}, \overline{\mathcal{N}}_{\text{cum}})$ for $k = 0, 1, 2, ...$ describes the exact cumulative distribution $\mathcal{N}_{\text{cum}}(d)$ well at the points $d = \overline{d}(m)$ in the region of large d, which corresponds to $m \ll n$. There are jumps in the cumulative distribution at these points when n is large, and $\overline{\mathcal{N}}_{\text{cum}}(m)$ approaches the top points of these jumps.

The resulting mean-field theory expression for the distribution takes the following form:

$$\overline{\mathcal{N}}(d,n) = \delta(d, d_D(n)) \\ + \sum_{m=1}^{[n/2]} \frac{n}{n-m}\binom{n-m}{m}\,\delta\Big(d, \frac{z_d^n}{(z_d^4-2)^m}\Big). \quad (35)$$

This approximation could also be obtained from the exact result for $\mathcal{N}(d,n)$, Eq. (24), by approximating the weighted sum $\sum_i$ of Kronecker symbols in the expression for $\mathcal{N}(d,n)$ by a single Kronecker symbol with a factor.

Let us obtain an explicit formula for the asymptotics of $\overline{\mathcal{N}}$ in terms of the degeneracy d. We simply replace m in Eq. (34) by the inverse of the function $\overline{d}(m)$ in Eq. (33). The inverse function is $m(d) = (\ln d - n \ln z_d)/(\ln C_1 - \ln z_d)$ (where we have dropped the bar because d is now the independent variable), and substituting into Eq. (34) gives

$$\overline{\mathcal{N}} = \frac{n}{n-m}\binom{n-m}{m} \approx \frac{n^m}{m!} \\ = \frac{z_d^{\frac{n \ln n}{\ln(z_d/C_1)}}}{\Gamma\left[1 - \frac{\ln(d/z_d^n)}{\ln(z_d/C_1)}\right]}\, d^{-\frac{\ln n}{\ln(z_d/C_1)}} \quad (36)$$

for $m \ll n$. Note that the logarithmic derivative of $\overline{\mathcal{N}}$ over d at the point d_D (the highest degeneracy), $-\ln n/\ln(z_d^4-2)$, properly fits the two rightmost points of the spectrum, namely $d_D(n) \cong z_d^n$, $\mathcal{N}(d_D) = 1$ and $d_{D-1}(n) \cong z_d^n/(z_d^4-2)$, $\mathcal{N}(d_{D-1}) = n$.

VII. DEGENERACY DISTRIBUTIONS OF OUTPUTS FOR RANDOMLY GENERATED INPUT DATASETS

We observed that the outputs generated by complete input data sets in the previous section do not produce real power laws. One should note that in empirical studies, input data sets, typically, are not complete. Usually, when inputs (input size n) are sufficiently large, the input data set sizes, N, are much smaller than of complete data sets, $N \ll C^n$, $C > 1$. Based on the sets of outputs of complete data sets, which we obtained in the previous section (listed in the Supplementary Material [11]), here we find and explore the distribution of outputs of randomly generated data sets of various sizes. (In principle, these sizes can be also bigger than or equal to that of complete data sets.)

Let the size of a randomly generated input data set be N, i.e., we apply filtering to N randomly generated rings of n zeroes and ones. We assume that all degeneracies of the outputs of the corresponding complete input data set are known, namely the full set of pairs $\{d_i, \mathcal{N}(d_i)\}$, where $i = 1, 2, ..., D$, D is the total number of degeneracies for this n in the case of a complete input data set. Then for the randomly generated input data set we obtain the following expected number of outputs of degeneracy d:

$$\langle\mathcal{N}\rangle(d) = \binom{N}{d}\sum_{i=1}^{D}\mathcal{N}(d_i)\Big(\frac{d_i}{2^n}\Big)^d\Big(1-\frac{d_i}{2^n}\Big)^{N-d}. \quad (37)$$

The probability that a randomly generated input string produces a given output is simply $d_i/2^n$, where d_i is the degeneracy of the output with respect to the total input set. The probability that d of the N inputs produce this output is then simply given by a binomial expression. Summing over all total degeneracies (multiplying by the number of instances of a given degeneracy d_i and summing over i) gives the above expression. Here outputs of any degeneracy within the interval $1 \le d \le d_D$ are present with nonzero probability, in contrast to the case of the complete input data set. The results of application of this formula coincide with those obtained by recording statistics of outputs directly filtered from randomly generated inputs. Equation (37) leads to the following equality:

$$\sum_d d\,\langle\mathcal{N}\rangle(d) = N. \quad (38)$$

Let us apply Eq. (37), for instance, to the cases $n = 21$ [see Fig. 2(a) above for the complete input data set consisting of all 2^{21} configurations] and $n = 120$, and inspect the distributions of outputs of uniformly randomly generated input sets of different sizes. The results are shown in Fig. 7. For each size of the random input data set we present the degeneracy distribution and its cumulative counterpart. These figures demonstrate that for $N \ll 2^n$, the distributions indeed resemble a power law. The reason is that for sufficiently small N, the distribution does not approach the large values of d for which the variation of the exponent $c \ln^{\alpha-1} d$ in $\mathcal{N}_{\text{cum}}(d) \propto d^{-c\ln^{\alpha-1} d}$, [see Eq. (25)], becomes noticeable. As N approaches 2^n, the distributions become closer to their counterparts for the complete input data set. Clearly, the distributions obtained in the limit $N \to \infty$ will coincide with those found for the complete input data set. Curiously, one may obtain distributions with a very similar form for different values of n by choosing N in order to maintain the scaling variable

$$s = (z_d/2)^n N \quad (39)$$

constant. For example, compare the first with the third row and the second with the fourth in Fig. 7. This combination follows from the fact that the binomial product in Eq. (37), as a function of d, forms a narrow peak centered at $d_i N/2^n$, which produces $(z_d/2)^n N$ for the largest degeneracy $d_D \cong z_d^n$. This scaling disappears in the region of small degeneracies. Note that if $N \ll 2^n$, Eq. (37) gives $\langle\mathcal{N}\rangle(1) \approx N$ since $\sum_i d_i \mathcal{N}(d_i) = 2^n$. Furthermore, let $N \ll (4/c_2)^n$, where $c_2 = 3.139...$, $\sum_i d_i^2 \mathcal{N}(d_i) \cong c_2^n$, see Appendix B. Then $\langle\mathcal{N}\rangle(1) \approx N \gg \langle\mathcal{N}\rangle(2) \approx N(Nc_2^n/2^{2n+1})$. Thus the vast majority of outputs have degeneracy 1 in this situation.

VIII. DISCUSSIONS AND CONCLUSIONS

Our straightforward, purely combinatorial treatment reveals features of distributions of outputs hidden from other approaches. For complete input data sets passed through our filter, we have obtained degeneracy distributions markedly distinct from power laws. On the other hand, these distributions decay as $\mathcal{N}_{\text{cum}}(d) \propto e^{-c\ln^\alpha d}$, α>1, much slower than exponentially, and in this sense they can still be called "critical". We have observed that the entire form of these output distributions essentially depends on the input size n, which strongly differs, for example, from heavy tailed degree distributions of complex networks having exponential cutoffs [23, 24].

For randomly generated input data sets, we found degeneracy distributions which could easily be taken for power laws in empirical studies, if the data set size N is essentially smaller than 2^n. As $N \to \infty$, these distributions approach the clearly non-power-law shape of the distributions for the complete input data set. Thus we show that the size of an input data set matters for these problems.

Our model filter can be used as a convenient reference filtering problem. We focused on simple input sets which were uniformly random strings of zeroes and ones. Correlated inputs are more challenging for analytical treatment. Exploring the simplest filter patterns, we showed that the statistics of outputs is determined not by the form of filter patterns but rather by what occurs in the gaps of zeroes between them. The degeneracy corresponding to each such gap can be found using recursion relationships. We then used an integer partitions apparatus to aggregate the statistics of prime degeneracies from these gaps, finding the exact full spectrum of output degeneracies.

Considering the permutations of the integer partitions allows us to calculate the resulting exact degeneracy distribution $\mathcal{N}(d, n)$. Alternatively, using multiplicative partitions we derive coupled linear recursions providing the exact $\mathcal{N}(d, n)$ for any finite d and n and the explicit large n asymptotics. Finally we developed a mean-field theory which describes the approximate degeneracy spectrum, degeneracy distribution and their asymptotic behavior. Our mean-field theory of Sec. IV also derives from the gaps between outputs. These results show that in filter problems of this kind, the statistics of outputs is determined by the gaps between outputs, which are essentially determined by the size of a filter pattern, and the generalization of our results to larger patterns should be straightforward. Therefore, more complex, larger filter patterns than ours, Eq. (1), can be considered in a similar way, and we expect similar distributions. For the sake of simplicity, we studied inputs containing only zeroes and ones. We expect that our approach and the mean-field theory should be applicable to more rich inputs containing more degrees of freedom: larger sets of numbers, vectors, etc., as well as to more general cooperative systems with a large number of local minima. The next natural step after the mean field theory, namely a fluctuation theory, should be based on accounting for small gaps between ones and the fluctuations of the lengths of these gaps.

In summary, we suggest that our conclusions could be generalized to other filtering and compression problems involving more complex filter patterns and complex, not necessary synthetic, higher dimensional inputs.

ACKNOWLEDGMENTS

This work was supported by National Funds through FCTPortuguese Foundation for Science and Technology, I.P. Project No. IF/00726/2015. R. A. C. acknowledges the FCT grant SFRH/BPD/123077/2016.

Appendix A: Recursive relations

Here we present the derivations of the expression for the number of different values of degeneracy of outputs produced by complete input data sets, Eq. (2), and the recursions for $\mathcal{N}(d, n)$ used in Sec. III.

1. Derivation of Eq. (2) for $M(n)$

The total number $M(n)$ of different outputs in our problem is the number of all possible periodic (period n) combinations of zeroes and ones having no neighboring ones. Clearly, this number coincides with the total number of different combinations of dimers in a ring of length n. Then $M(n)$ is a sum of combinations of dimers in a string of length $n-2$ or n, respectively.

To find the number of different outputs with k ones, that is, the number of ways of placing k dimers in the ring, it is convenient to consider the cases when the first output digit is 1 and when it is 0 separately. This is equivalent to fixing the state of two chosen neighboring units: either there is a dimer connecting this pair or this dimer is absent.

(i) When the first digit of an output is 1, both the second and the last digits must be 0, and the number of outputs in this case is given by the binomial coefficient $\binom{n-k-1}{k-1}$, which corresponds to starting with the sequence $1, \{0\}_{n-k}$ (where $\{0\}_{n-m}$ is a sequence of $n-m$ zeroes), and choosing $k-1$ out of $n-k-1$ distinct zeroes to be

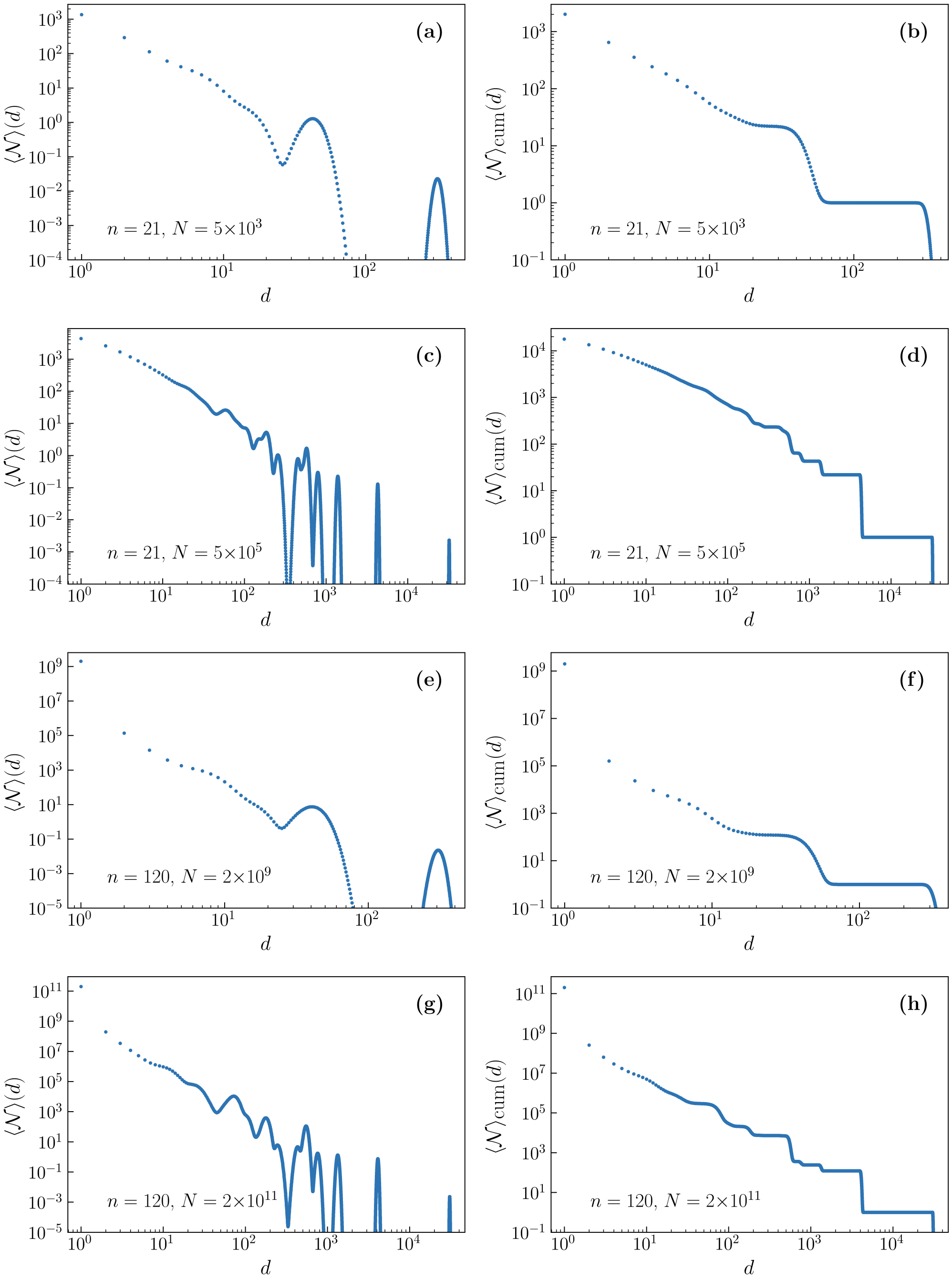

FIG. 7. Double logarithmic scale plots of (a,c,e,g) the degeneracy distributions and (b,d,f,h) the cumulative degeneracy distributions obtained for a randomly generated input data sets of different sizes N for $n = 21$ and $n = 120$. The specific sizes of input data sets for $n = 120$ are chosen to produce distributions similar to those for $n = 21$.

replaced by 01. Due to the periodic boundary condition, the last 0 cannot be selected for replacement, hence the number of zeroes that can be replaced is $n - k - 1$.

(ii) When the first digit of an output is 0, we start from the sequence $\{0\}_{n-k}$ and replace k zeroes by 01. In this case the replacement may be made at any of the $n - k$ zeroes, and the number of outputs is $\binom{n-k}{k}$.

The total number of different outputs is obtained by summing these two contributions

$$\overline{\mathcal{N}}(k) = \binom{n-k-1}{k-1} + \binom{n-k}{k} = \frac{n}{n-k}\binom{n-k}{k}, \tag{A1}$$

where we used the notation $\overline{\mathcal{N}}(k)$ from Sec. VI.

Summing over all possible values of k gives

$$M(n) = 1 + \sum_{k=1}^{[n/2]} \Big[\frac{(n-k-1)!}{(k-1)!(n-2k)!} + \frac{(n-k)!}{k!(n-k)!} \Big]$$
$$= n \sum_{k=0}^{[n/2]} \frac{1}{n-k} \binom{n-k}{k}, \qquad \text{(A2)}$$

i.e., Eq. (2).

2. Recursions for $\mathcal{N}(d,n)$

Input strings containing no strings of length greater than one give outputs identical to the input. No other input can produce the same output, so these outputs have degeneracy 1. Such outputs, of degeneracy 1, contain no chains of zeroes of length greater than 3. For the input size n, the number of such outputs, $\mathcal{N}(1,n)$, can be obtained recursively. Here we derive the recursive relation for $\mathcal{N}(1,n)$ and indicate the recursions for $\mathcal{N}(d,n)$ for higher degeneracy d.

Sequences of degeneracy 1 can be regarded as assortments of three kinds of building blocks, 01, 001, and 0001, put together in a ring of length n. All configurations of blocks are allowed, as long as the total number of binary digits is n. Let us consider a particular position i in the ring and the block to which i belongs. If we add a block 01 between the block of i and one of its neighbor blocks (say, the one to the right) to every possible configuration of length $n-2$, we get every possible configuration of length n that has a block 01 to the right of the block of i. Doing the same with configurations of length $n-3$ and blocks 001, we get all configurations with a block 001 to the right of the block of i. Finally, repeating the procedure for configurations of length $n-4$ and blocks 0001, gives all configurations with a block 0001 to the right of the block of i. Since every block must be 01, 001, or 0001, the union of these three sets is the full set of configurations of degeneracy 1 in system of n digits. Thus, for the number of configurations in this set, $\mathcal{N}(1,n)$, we can write

$$\mathcal{N}(1,n) = \mathcal{N}(1,n-2) + \mathcal{N}(1,n-3) + \mathcal{N}(1,n-4). \quad \text{(A3)}$$

The explicit solution of this linear difference equation is given in terms of the roots, z_1, z_2, z_3, and z_4, of the characteristic equation $z^4 = z^2 + z + 1$:

$$\mathcal{N}(1,n) = {z_1}^n + {z_2}^n + {z_3}^n + {z_4}^n, \qquad \text{(A4)}$$

where the coefficients of the powers of the roots z_i are found through the initial condition,

$$\mathcal{N}(1,1){=}0,\ \mathcal{N}(1,2){=}2,\ \mathcal{N}(1,3){=}3,\ \mathcal{N}(1,4){=}6, \quad \text{(A5)}$$

and are equal to 1. The root $z_1 \equiv z_a = 1.46557...$ determines the large n asymptotics of $\mathcal{N}(1,n)$, Eq. (27).

We can expand the analysis to higher values of degeneracy. For example, the second smallest value, $d_2 = 2$, corresponds to the outputs with blocks of the types 01, 001, and 0001 and a single string 00001. Applying the same reasoning as above, we get three terms similar to the ones of Eq.(A3). We must, additionally, consider configurations with degeneracy 1, to which we add a string 00001 in the same way, instead of one of the other blocks. Then we get

$$\mathcal{N}(2,n) = \mathcal{N}(2,n{-}2) + \mathcal{N}(2,n{-}3) + \mathcal{N}(2,n{-}4) + \mathcal{N}(1,n{-}5) \quad \text{(A6)}$$

with the initial condition

$$\mathcal{N}(2,1) = \mathcal{N}(2,2) = \mathcal{N}(2,3) = \mathcal{N}(2,4) = 0,\ \ \mathcal{N}(2,5) = 5. \quad \text{(A7)}$$

In a similar way, taking into account that the prime degeneracy $\tilde{d}(\ell)$ correspond to the inserted block of $\ell + 1$ (chain of ℓ zeroes and 1 on the right), i.e., $2, 4, 7, 12, 21, 37, 65, ...$ correspond to $\ell + 1 = 5, 6, 7, 8, 9, 10, 11, ...$, respectively, we derive the recursion relations for higher $\mathcal{N}(d,n)$ [here $d \neq d_D(n)$, for which, clearly, $\mathcal{N}(d_D(n),n) = 1$],

$$\mathcal{N}(4,n) = \mathcal{N}(4,n{-}2) + \mathcal{N}(4,n{-}3) + \mathcal{N}(4,n{-}4) + \mathcal{N}(2,n{-}5) + \mathcal{N}(1,n{-}6) \quad \text{(A8)}$$

with the initial condition

$$\mathcal{N}(4,1){=}...{=}\mathcal{N}(4,5){=}0,\ \mathcal{N}(4,6){=}6, \qquad \text{(A9)}$$

and so on. See the list of all recursive relations for $d < 100$ with initial conditions in the Supplementary Material [11]. The corresponding large n asymptotics are

$$\mathcal{N}(1,n) \cong z_a^n,$$
$$\mathcal{N}(2,n) \cong \frac{1}{z_a(2z_a^2 + 3z_a + 4)} n z_a^n = 0.05376 n z_a^n,$$
$$\mathcal{N}(4,n) \cong \frac{1}{2!} \Big[\frac{1}{z_a(2z_a^2 + 3z_a + 4)} \Big]^2 n^2 z_a^n = 0.001445 n^2 z_a^n, \quad \text{(A10)}$$

and so on. See the list of the resulting large n asymptotics of $\mathcal{N}(d{<}100, n)$ in the Supplementary Material [11].

More generally, for any value of degeneracy d one can write the recursion relation for $\mathcal{N}(d,n)$ in terms of the multiplicative partition of d into prime degeneracies. Let us present the full set of recursion relations for $\mathcal{N}(d,n)$. The reasoning is as follows.

(i) The recursion relation for $\mathcal{N}(1,n)$ is given by Eq. (A3) with the initial condition Eq. (A5).

(ii) For the largest degeneracies, $\mathcal{N}(d_D(n),n) = 1$.

(iii) All prime degeneracies $\tilde{d}(\ell)$ except $\tilde{d}(5) = 4$ are not multiplicatively separable into other prime degeneracies (primality property). The only exception is $\tilde{d}(5) = \tilde{d}^2(4) = 2^2$. See Sec. V above for more detail.

(iv) Any degeneracy in the spectrum except d_D (and $d = 1$) is multiplicatively separable into the prime degeneracies $\tilde{d}(\ell)$, namely,

$$\begin{aligned} d &= \tilde{d}^{\mu_4}(4)\tilde{d}^{\mu_5}(5)\tilde{d}^{\mu_6}(6)\tilde{d}^{\nu_7}(7)... \\ &= 2^{\mu_4}4^{\mu_5}7^{\mu_6}12^{\mu_7}... \\ &= 2^{\mu_4+2\mu_5}7^{\mu_6}12^{\mu_7}..., \end{aligned} \tag{A11}$$

where the powers $\mu_\ell \equiv \mu(\ell)$ are non-negative integers. Due to the exception $\tilde{d}(5) = \tilde{d}^2(4) = 2^2$ and the coincidence $\tilde{d}(5)\tilde{d}(8) = 4 \times 21 = 7 \times 12 = \tilde{d}(6)\tilde{d}(7)$, etc., the multiplicative partition into the prime degeneracies $\tilde{d}(\ell)$, Eq. (A11), may not be unique, see below. With all possible integers $\mu(4) \geq 0$, $\mu(6) \geq 0$, $\mu(7) \geq 0$ ... we generate the full list $\mathcal{D}$ of all degeneracies except 1 and d_D in the spectrum for $n \to \infty$. We place the generated degeneracies in ascending order, ignoring repetitions for some degeneracies. For a finite n, only some on the degeneracies from this list are present in the spectrum. In practice, we need degeneracies d only up to some, maybe, large but finite value. So, generating this list, we use a finite set of non-negative integers $\mu_4, \mu_6, \mu_7, ...$ respectively restricted from above.

(v) For any degeneracy d in this list $\mathcal{D}$ we define the vector $\mathcal{L}(d) \equiv (\ell_1, \ell_2, ..., \ell_{\max})$, indicating that the prime degeneracies $(\tilde{d}(\ell_1), \tilde{d}(\ell_2), ..., \tilde{d}(\ell_{\max}))$, given in ascending order, are present in at least one of the partitions of d, Eq. (A11). That is, if a prime partition $\tilde{d}(\ell)$ is present in $\mathcal{L}(d)$, then the ratio $d/\tilde{d}(\ell)$ also belongs to $\mathcal{D}$. Then the recursion relation for $\mathcal{N}(d, n)$ has the form

$$\begin{aligned} \mathcal{N}(d,n) = \mathcal{N}(d,n-2) + \mathcal{N}(d,n-3) + \mathcal{N}(d,n-4) \\ + \sum_{\ell \in \mathcal{L}(d)} \mathcal{N}(d/\tilde{d}(\ell), n-\ell-1), \end{aligned} \tag{A12}$$

where ℓ in the sum takes the values $\ell_1, \ell_2, ..., \ell_{\max}$, i.e., the components of the vector $\mathcal{L}(d)$. This formula sums the number of configurations $\mathcal{N}(d/\tilde{d}(\ell), n-\ell-1)$ of smaller systems to which a block of ℓ zeros (followed by a one) can be added in order to form a configuration of size n and degeneracy d. We should not explicitly include in this sum configurations that achieve degeneracy d by inserting more than one block of zeros. The insertion of additional blocks besides a block of length l, say, into smaller configurations are already accounted for in the calculation of the degeneracy of the configuration of size $n-\ell-1$ into which the block of length l will be inserted. Note that the ratio $d/\tilde{d}(\ell)$ in each of the terms in the sum is one of the degeneracies, smaller than d, in the list $\mathcal{D}$ with the degeneracy $d = 1$ added, so this set of recursions together with the recursion for $\mathcal{N}(1, n)$, Eq. (A3), is closed. The initial conditions for these recursions are

$$\begin{aligned} \mathcal{N}(d,1) = ... = \mathcal{N}(d,\ell_{\max}) = 0, \\ \mathcal{N}(d,\ell_{\max}+1) = (\ell_{\max}+1)\delta(d,\tilde{d}(\ell_{\max})), \end{aligned} \tag{A13}$$

where $\delta(i, j)$ is the Kronecker symbol. Note than when d is a prime degeneracy $d = \tilde{d}(\ell) = \tilde{d}(\ell_{\max})$ except $d = 4 = \tilde{d}(5)$, the sum on the left-hand side of Eq. (A12) has only one term, $\mathcal{N}(1, n-\ell_{\max}-1)$. If $d = 4$, Eqs. (A12) and (A13) properly give the recursion for from Eq. (A8) with the initial condition from Eq. (A9). Finally, for $d = 1$, use Eq. (A3) with the initial condition Eq. (A5). [In fact, if $d = 1$, the list $\mathcal{L}(d = 1)$ is empty, and Eq. (A12) is reduced to Eq. (A3). We still need the initial condition, Eq. (A5), for $d = 1$, since Eq. (A13) is not defined in this case.]

(vi) Equations (A3), (A5) and (A12), (A13), in particular, provide all degeneracies present in the spectrum for a given n. If some d is absent, these recursions produce $\mathcal{N}(d, n) = 0$. For instance, let $d = 64$. The initial condition is $\mathcal{N}(64,1)$=...=$\mathcal{N}(64,6)$=0. The recursions produce $\mathcal{N}(64,7)$=...=$\mathcal{N}(64,17)$=0, $\mathcal{N}(64,18)$=6, $\mathcal{N}(64,19)$=0, $\mathcal{N}(64,20)$=20, and so on.

(vii) Our recursions, Eq. (A12), with their initial conditions, Eq. (A13), lead to the following large n asymptotics of $\mathcal{N}(d, n)$:

$$\begin{aligned} &\mathcal{N}(d,n) \\ &\cong \frac{1}{\prod_{\ell \neq 5} \mu_\ell!} \left[\frac{n}{z_a(2z_a^2 + 3z_a + 4)}\right]^{\sum_{\ell \neq 5} \mu_\ell} z_a^{n - \sum_{\ell \geq 6}(\ell-4)\mu_\ell} \end{aligned} \tag{A14}$$

for d with the multiplicative partition into prime degeneracies, Eq. (A12), chosen in such a way that the power of 2, μ_4, in this partition is maximal.

Let us discuss this point in more detail. We stressed above that a multiplicative partition of $d \in \mathcal{D}$, Eq. (A12), may not be unique. For large n, the contribution of each of the possible multiplicative partitions of d to $\mathcal{N}(d, n)$ is about $n^{\sum_\ell \mu_\ell} z_a^{n-\sum_\ell(\ell-4)\mu_\ell}$. The leading asymptotics, Eq. (A14), is with the maximal power of n, i.e., it originates from the partition with the maximal sum $\sum_\ell \mu_\ell$. Let us check whether the multiplicative partition of d with the maximal sum $\sum_\ell \mu_\ell$ is unique and that this maximum corresponds to the maximal μ_4. In other words, if there exist a number of different multiplicative partitions of d, that only one of them has the maximal μ_4, and that this partition has the maximal sum $\sum_\ell \mu_\ell$. We inspected all products of prime degeneracies $\tilde{d}(\ell \leq 200) \approx 1.7\, 10^{48}$ e focusing on the products producing non-unique multiplicative partitions. Apart from $2^2 = 4$ discussed above, we find only three such combinations with $\ell \leq 200$ (into this number, we do not include the products of these combinations and arbitrary prime degeneracies). Namely, $2^2 \times 21 = 7 \times 12$, $2^9 \times 200 \times 351^2 = 12^6 \times 65^2$, and $2^2 \times 12^2 \times 170625 = 7 \times 200^2 \times 351$. For each of these combinations we confirm that the left side of the equality corresponds to the maximal $\sum_\ell \mu_\ell$ and that this partitions is unique. Clearly, the same is true for the products of these combinations and arbitrary prime degeneracies. Thus partitions contributing to the leading asymptotics of $\mathcal{N}(d, n)$ are unique, and the gauge is fixed by demanding $\mu_4 = \max$, which leads to Eq. (A14). For grasping the form of Eq. (A14), see also Eqs. (A10), especially the asymptotics of $\mathcal{N}(84, n)$.

Thus we have the chain of conveniently coupled linear recursion relations that one can easily process starting from $d = 1$. These recursions generate exact $\mathcal{N}(d, n)$ for finite d and n, and in this sense provide the exact solution of the problem. Moreover, the leading large n asymptotics of $\mathcal{N}(d, n)$ are found explicitly. Note that the time of computing all $\mathcal{N}(d < d_0, n)$, where d_0 is fixed, by using our recursions, is proportional to n, i.e. the computation for each next size takes the same time.

Appendix B: Moments of $\mathcal{N}(d, n)$

Here we list the large n asymptotics of the moments of $\mathcal{N}(d, n)$. The leading asymptotics of the moments are of the following form:

$$\mathcal{M}_k(n) \equiv \sum_{i=1}^{D} d_i^k \mathcal{N}(d_i, n) \cong c_k^n, \tag{B1}$$

where

$$\begin{aligned}
c_0 &= z_g, \\
c_1 &= 2, \\
c_2 &= 3.13899009933542, \\
c_3 &= 5.41762864130976, \\
c_4 &= 9.48696631140060, \\
c_5 &= 16.6438119672308, \\
c_6 &= 29.2067717071942, \\
c_7 &= 51.2540583046806, \\
c_8 &= 89.9445429351823, \\
&\ \ ...
\end{aligned} \tag{B2}$$

In their turn, the numbers c_k have the following large k asymptotics:

$$c_k \cong z_d^k [1 + (z_d^4 - 2)^{-k} + ...], \tag{B3}$$

which is close to the numerical values listed above, Eq. (B2), already for $k = 2$. The leading term z_d^k results from the largest degeneracy in the spectrum, $\mathcal{N}(d_D \cong z_d^n, n) = 1$, the next term originates from the second largest degeneracy, etc.

[1] J. Song, M. Marsili, and J. Jo, "Emergence and relevance of criticality in deep learning," arXiv preprint arXiv:1710.11324 (2017).

[2] R. Cubero, M. Marsili, and Y. Roudi, "Minimum description length codes are critical," Entropy **20**, 755 (2018).

[3] A. Haimovici and M. Marsili, "Criticality of mostly informative samples: A Bayesian model selection approach," J. Stat. Mech.: Theory and Experiment **2015**, P10013 (2015).

[4] R. J. Cubero, J. Jo, M. Marsili, Y. Roudi, and J. Song, "Minimally sufficient representations, maximally informative samples and Zipf's law," arXiv preprint arXiv:1808.00249 (2018).

[5] J. Song, M. Marsili, and J. Jo, "Resolution and relevance trade-offs in deep learning," J. Stat. Mech.: Theory and Experiment **2018**, 123406 (2018).

[6] G. Bianconi, "Entropy of network ensembles," Phys. Rev. E **79**, 036114 (2009).

[7] G. Bianconi, "A statistical mechanics approach for scale-free networks and finite-scale networks," Chaos: An Interdisciplinary Journal of Nonlinear Science **17**, 026114 (2007).

[8] A. G. Bashkirov and A. V. Vityazev, "Information entropy and power-law distributions for chaotic systems," Physica A **277**, 136 (2000).

[9] Y. Dover, "A short account of a connection of power laws to the information entropy," Physica A **334**, 591 (2004).

[10] M. Visser, "Zipf's law, power laws and maximum entropy," New Journal of Physics **15**, 043021 (2013).

[11] "Supplementary Material," `https://???.com`.

[12] V. E. Hoggatt Jr., *Fibonacci and Lucas Numbers* (Houghton Mifflin, Boston, MA, 1969).

[13] R. L. Graham, D. E. Knuth, O. Patashnik, and S. Liu, *Concrete Mathematics: A Foundation for Computer Science* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1994).

[14] T. Koshy, *Fibonacci and Lucas Numbers with Applications* (John Wiley & Sons, Inc., Hoboken, New Jersey, 2019).

[15] H. C. Austin and R. Guy, "Binary sequences without isolated ones," The Fibonacci Quarterly **16**, 84 (1978).

[16] "The On-Line Encyclopedia of Integer Sequences," `https://oeis.org`.

[17] G. E. Andrews, *The Theory of Partitions* (Cambridge University Press, Cambridge, 1998).

[18] G. E. Andrews and K. Eriksson, *Integer Partitions* (Cambridge University Press, Cambridge, 2004).

[19] G. H. Hardy and S. Ramanujan, “Asymptotic formulae in combinatory analysis,” Proc. London Math. Soc. **17**, 75 (1918).

[20] J. V. Uspensky, “Asymptotic formulae for numerical functions which occur in the theory of partitions,” Bull. Acad. Sci. URSS **14**, 199 (1920).

[21] H. S. Wilf, “Lectures on Integer Partitions,” `https://www.math.upenn.edu/~wilf/PIMS/PIMSLectures.pdf` (2000).

[22] P. Erdös and M. Kac, “The Gaussian law of errors in the theory of additive number theoretic functions,” Am. J. Math. **62**, 738 (1940).

[23] A. Clauset, C. R. Shalizi, and M. E. J. Newman, “Power-law distributions in empirical data,” SIAM Review **51**, 661 (2009).

[24] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).

SUPPLEMENTARY MATERIAL: 1. RECURSIONS AND ASYMPTOTICS FOR $\mathcal{N}(d < 100, n)$. 2. TABLES OF EXACT RESULTS FOR COMPLETE INPUT DATA SETS

1. Recursions and asymptotics for $\mathcal{N}(d < 100, n)$

Here we show the recursion relations for $\mathcal{N}(d, n)$ with degeneracy d up to 100, initial conditions for them, and their large n asymptotics. These formulas are particular cases of Eqs. (A12), (A13), and (A10), respectively.

$$\begin{aligned}
&\mathcal{N}(1,n) = \mathcal{N}(1,n-2) + \mathcal{N}(1,n-3) + \mathcal{N}(1,n-4),\\
&\mathcal{N}(2,n) = \mathcal{N}(2,n-2) + \mathcal{N}(2,n-3) + \mathcal{N}(2,n-4) + \mathcal{N}(1,n-5),\\
&\mathcal{N}(4,n) = \mathcal{N}(4,n-2) + \mathcal{N}(4,n-3) + \mathcal{N}(4,n-4) + \mathcal{N}(2,n-5) + \mathcal{N}(1,n-6),\\
&\mathcal{N}(7,n) = \mathcal{N}(7,n-2) + \mathcal{N}(7,n-3) + \mathcal{N}(7,n-4) + \mathcal{N}(1,n-7),\\
&\mathcal{N}(8,n) = \mathcal{N}(8,n-2) + \mathcal{N}(8,n-3) + \mathcal{N}(8,n-4) + \mathcal{N}(4,n-5) + \mathcal{N}(2,n-6),\\
&\mathcal{N}(12,n) = \mathcal{N}(12,n-2) + \mathcal{N}(12,n-3) + \mathcal{N}(12,n-4) + \mathcal{N}(1,n-8),\\
&\mathcal{N}(14,n) = \mathcal{N}(14,n-2) + \mathcal{N}(14,n-3) + \mathcal{N}(14,n-4) + \mathcal{N}(7,n-5) + \mathcal{N}(2,n-7),\\
&\mathcal{N}(16,n) = \mathcal{N}(16,n-2) + \mathcal{N}(16,n-3) + \mathcal{N}(16,n-4) + \mathcal{N}(8,n-5) + \mathcal{N}(4,n-6),\\
&\mathcal{N}(21,n) = \mathcal{N}(21,n-2) + \mathcal{N}(21,n-3) + \mathcal{N}(21,n-4) + \mathcal{N}(1,n-9),\\
&\mathcal{N}(24,n) = \mathcal{N}(24,n-2) + \mathcal{N}(24,n-3) + \mathcal{N}(24,n-4) + \mathcal{N}(12,n-5) + \mathcal{N}(2,n-8),\\
&\mathcal{N}(28,n) = \mathcal{N}(28,n-2) + \mathcal{N}(28,n-3) + \mathcal{N}(28,n-4) + \mathcal{N}(14,n-5) + \mathcal{N}(7,n-6) + \mathcal{N}(4,n-7),\\
&\mathcal{N}(32,n) = \mathcal{N}(32,n-2) + \mathcal{N}(32,n-3) + \mathcal{N}(32,n-4) + \mathcal{N}(16,n-5) + \mathcal{N}(8,n-6),\\
&\mathcal{N}(37,n) = \mathcal{N}(37,n-2) + \mathcal{N}(37,n-3) + \mathcal{N}(37,n-4) + \mathcal{N}(1,n-10),\\
&\mathcal{N}(42,n) = \mathcal{N}(42,n-2) + \mathcal{N}(42,n-3) + \mathcal{N}(42,n-4) + \mathcal{N}(21,n-5) + \mathcal{N}(2,n-9),\\
&\mathcal{N}(48,n) = \mathcal{N}(48,n-2) + \mathcal{N}(48,n-3) + \mathcal{N}(48,n-4) + \mathcal{N}(24,n-5) + \mathcal{N}(12,n-6) + \mathcal{N}(4,n-8),\\
&\mathcal{N}(49,n) = \mathcal{N}(49,n-2) + \mathcal{N}(49,n-3) + \mathcal{N}(49,n-4) + \mathcal{N}(7,n-7),\\
&\mathcal{N}(56,n) = \mathcal{N}(56,n-2) + \mathcal{N}(56,n-3) + \mathcal{N}(56,n-4) + \mathcal{N}(28,n-5) + \mathcal{N}(14,n-6) + \mathcal{N}(8,n-7),\\
&\mathcal{N}(64,n) = \mathcal{N}(64,n-2) + \mathcal{N}(64,n-3) + \mathcal{N}(64,n-4) + \mathcal{N}(32,n-5) + \mathcal{N}(16,n-6),\\
&\mathcal{N}(65,n) = \mathcal{N}(65,n-2) + \mathcal{N}(65,n-3) + \mathcal{N}(65,n-4) + \mathcal{N}(1,n-11),\\
&\mathcal{N}(74,n) = \mathcal{N}(74,n-2) + \mathcal{N}(74,n-3) + \mathcal{N}(74,n-4) + \mathcal{N}(37,n-5) + \mathcal{N}(2,n-10),\\
&\mathcal{N}(84,n) = \mathcal{N}(84,n-2)+\mathcal{N}(84,n-3)+\mathcal{N}(84,n-4)+\mathcal{N}(42,n-5)+\mathcal{N}(21,n-6)+\mathcal{N}(12,n-7)+\mathcal{N}(7,n-8)+\mathcal{N}(4,n-9),\\
&\mathcal{N}(96,n) = \mathcal{N}(96,n-2) + \mathcal{N}(96,n-3) + \mathcal{N}(96,n-4) + \mathcal{N}(48,n-5) + \mathcal{N}(24,n-6) + \mathcal{N}(8,n-8),\\
&\mathcal{N}(98,n) = \mathcal{N}(98,n-2) + \mathcal{N}(98,n-3) + \mathcal{N}(98,n-4) + \mathcal{N}(49,n-5) + \mathcal{N}(14,n-7)
\end{aligned} \tag{B4}$$

with the initial conditions

$$\begin{aligned}
&\mathcal{N}(1,1){=}0,\ \mathcal{N}(1,2){=}2,\ \mathcal{N}(1,3){=}3,\ \mathcal{N}(1,4){=}6,\\
&\mathcal{N}(2,1){=}...{=}\mathcal{N}(2,4){=}0,\ \mathcal{N}(2,5){=}5,\\
&\mathcal{N}(4,1){=}...{=}\mathcal{N}(4,5){=}0,\ \mathcal{N}(4,6){=}6,\\
&\mathcal{N}(7,1){=}...{=}\mathcal{N}(7,6){=}0,\ \mathcal{N}(7,7){=}7,\\
&\mathcal{N}(8,1){=}...{=}\mathcal{N}(8,6){=}0,\\
&\mathcal{N}(12,1){=}...{=}\mathcal{N}(12,7){=}0,\ \mathcal{N}(12,8){=}8,\\
&\mathcal{N}(14,1){=}...{=}\mathcal{N}(14,7){=}0,\\
&\mathcal{N}(16,1){=}...{=}\mathcal{N}(16,6){=}0,\\
&\mathcal{N}(21,1){=}...{=}\mathcal{N}(21,8){=}0,\ \mathcal{N}(21,9){=}9,\\
&\mathcal{N}(24,1){=}...{=}\mathcal{N}(24,8){=}0,\\
&\mathcal{N}(28,1){=}...{=}\mathcal{N}(28,7){=}0,\\
&\mathcal{N}(32,1){=}...{=}\mathcal{N}(32,6){=}0,\\
&\mathcal{N}(37,1){=}...{=}\mathcal{N}(37,9){=}0,\ \mathcal{N}(37,10){=}10,\\
&\mathcal{N}(42,1){=}...{=}\mathcal{N}(42,9){=}0,\\
&\mathcal{N}(48,1){=}...{=}\mathcal{N}(48,8){=}0,\\
&\mathcal{N}(49,1){=}...{=}\mathcal{N}(49,7){=}0,\\
&\mathcal{N}(56,1){=}...{=}\mathcal{N}(56,7){=}0,\\
&\mathcal{N}(64,1){=}...{=}\mathcal{N}(64,6){=}0,\\
&\mathcal{N}(65,1){=}...{=}\mathcal{N}(65,10){=}0,\ \mathcal{N}(65,11){=}11,\\
&\mathcal{N}(74,1){=}...{=}\mathcal{N}(74,10){=}0,\\
&\mathcal{N}(84,1){=}...{=}\mathcal{N}(84,9){=}0,\\
&\mathcal{N}(96,1){=}...{=}\mathcal{N}(96,8){=}0,\\
&\mathcal{N}(98,1){=}...{=}\mathcal{N}(98,7){=}0. \qquad \text{(B5)}
\end{aligned}$$

The corresponding large n asymptotics are

$$
\begin{aligned}
&\mathcal{N}(1,n) \cong z_a^n,\\
&\mathcal{N}(2,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^n = 0.05376 n z_a^n,\\
&\mathcal{N}(4,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^n = 0.001445 n^2 z_a^n,\\
&\mathcal{N}(7,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^{n-2} = 0.02503 n z_a^n,\\
&\mathcal{N}(8,n) \cong \frac{1}{3!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^3 n^3 z_a^n = 0.00002589 n^3 z_a^n,\\
&\mathcal{N}(12,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^{n-3} = 0.01708 n z_a^n,\\
&\mathcal{N}(14,n) \cong \Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^{n-2} = 0.001345 n^2 z_a^n,\\
&\mathcal{N}(16,n) \cong \frac{1}{4!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^4 n^4 z_a^n = 3.480\times 10^{-7} n^4 z_a^n,\\
&\mathcal{N}(21,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^{n-4},\\
&\mathcal{N}(24,n) \cong \Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^{n-3},\\
&\mathcal{N}(28,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^3 n^3 z_a^{n-2},\\
&\mathcal{N}(32,n) \cong \frac{1}{5!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^5 n^5 z_a^n,\\
&\mathcal{N}(37,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^{n-5},\\
&\mathcal{N}(42,n) \cong \Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^{n-4},\\
&\mathcal{N}(48,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^3 n^3 z_a^{n-3},\\
&\mathcal{N}(49,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^{n-4},\\
&\mathcal{N}(56,n) \cong \frac{1}{3!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^4 n^4 z_a^{n-2},\\
&\mathcal{N}(64,n) \cong \frac{1}{6!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^6 n^6 z_a^n,\\
&\mathcal{N}(65,n) \cong \frac{1}{z_a(2z_a^2+3z_a+4)}\, n z_a^{n-6},\\
&\mathcal{N}(74,n) \cong \Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^2 n^2 z_a^{n-5},\\
&\mathcal{N}(84,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^3 n^3 z_a^{n-4},\\
&\mathcal{N}(96,n) \cong \frac{1}{3!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^4 n^4 z_a^{n-3},\\
&\mathcal{N}(98,n) \cong \frac{1}{2!}\Big[\frac{1}{z_a(2z_a^2+3z_a+4)}\Big]^3 n^3 z_a^{n-4},\\
&\qquad\qquad \ldots .
\end{aligned}
\tag{B6}
$$

2. Tables of exact results for complete input data sets

Table I contains the complete data on the spectrum of degeneracy of outputs for complete input data sets of length n up to $n = 21$, namely all pairs $\{d_i, \mathcal{N}_i\}$, $i = 1, ..., D$, where d_i is the i-th value of degeneracy and $\mathcal{N}_i$ is the corresponding number of outputs having this degeneracy. This data for $n = 21$ is presented in graphical form in Fig. 2(a,b). Table II presents the total number of different values of degeneracy D for each input length $n \leq 120$.

TABLE I. Spectrum of degeneracy produced by complete input data sets with input strings of length n, i.e., the pairs $\{d_i, \mathcal{N}_i\}$, $i = 1, ..., D$, where d_i is the i-th value of degeneracy and $\mathcal{N}_i$ is the corresponding number of outputs having this degeneracy, D is the total number of different values of degeneracy.

n	d_j $\mathcal{N}_j$																												
3	1	5
	3	1
4	1	10
	6	1
5	1	2	17
	5	5	1
6	1	4	29
	11	6	1
7	1	2	7	51
	14	7	7	1
8	1	2	4	12	90
	22	8	8	8	1
9	1	2	4	7	21	158
	30	18	9	9	9	1
10	1	2	4	7	12	37	277
	47	20	25	10	10	10	1
11	1	2	4	7	8	12	21	65	486
	66	44	22	22	11	11	11	11	1
12	1	2	4	7	12	14	16	21	37	114	853
	99	60	60	24	24	12	6	12	12	12	1
13	1	2	4	7	8	12	21	24	28	37	65	200	1497
	143	104	78	52	26	26	26	13	13	13	13	13	1
14	1	2	4	7	8	12	14	16	21	37	42	48	49	65	114	351	2627
	212	154	147	70	28	56	28	14	28	28	14	14	7	14	14	14	1
15	1	2	4	7	8	12	14	16	21	24	28	37	65	74	84	114	200	616	4610
	308	255	210	120	80	75	30	15	60	30	30	30	30	15	30	15	15	15	1
16	1	2	4	7	8	12	14	16	21	24	28	37	42	48	49	65	114	130	144	147	148	200	351	1081	8090
	454	384	376	176	96	128	80	56	80	32	32	64	32	32	16	32	32	16	8	16	16	16	16	16	1
17	1	2	4	7	8	12	14	16	21	24	28	32	37	42	48	49	65	74	84	114	200	228	252	259	260	351	616	1897	14197
	663	612	561	289	238	187	102	51	136	85	102	17	85	34	34	17	68	34	68	34	34	17	17	17	17	17	17	17	1
18	1	2	4	7	8	12	14	16	21	24	28	37	42	48	49	56	64	65	74	84	114	130	144	147	148	200	351	400	441
	974	936	936	432	342	306	234	171	198	108	108	144	90	108	45	36	6	90	36	72	72	36	18	36	36	36	36	18	9
	444	455	456	616	1081	3329	24914
	18	18	18	18	18	18	1
19	1	2	4	7	8	12	14	16	21	24	28	32	37	42	48	49	65	74	84	96	98	112	114	130	144	147	148	200
	1425	1463	1444	684	665	456	342	228	323	247	304	57	209	114	114	57	152	95	209	38	19	19	95	38	19	38	38	76
	228	252	259	260	351	616	702	777	780	798	800	1081	1897	5842	43721
	38	38	38	38	38	38	19	19	19	19	19	19	19	19	1
20	1	2	4	7	8	12	14	16	21	24	28	32	37	42	48	49	56	64	65	74	84	114	130	144	147	148	168
	2091	2240	2340	1040	1040	720	640	505	480	360	420	60	340	260	320	130	120	20	220	120	240	160	100	50	100	120	80
	192	196	200	228	252	259	260	351	400	441	444	455	456	616	1081	1232	1365	1368	1369	1400	1404	1897	3329	10252	76725
	20	20	100	40	40	40	40	80	40	20	40	40	40	40	40	20	20	20	10	20	20	20	20	20	1
21	1	2	4	7	8	12	14	16	21	24	28	32	37	42	48	49	56	64	65	74	84	96	98	112	114	130
	3062	3465	3633	1617	1876	1092	1008	756	756	672	861	210	504	378	441	189	126	21	357	273	609	126	63	63	231	126
	144	147	148	200	228	252	259	260	288	294	296	336	343	351	400	441	444	455	456	616	702	777	780	798	800	1081	1897
	63	126	126	168	105	105	105	126	21	42	42	63	7	105	42	21	42	42	42	84	42	42	42	42	42	42	42
	2162	2394	2400	2405	2457	2464	3329	5842	17991	134643
	21	21	21	21	21	21	21	21	21	1

TABLE II. The total number of different values of degeneracy D for each input length n.

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16
D	2	2	3	3	4	5	6	7	9	11	13	17	19	25
n	17	18	19	20	21	22	23	24	25	26	27	28	29	30
D	29	36	43	52	63	75	90	108	128	153	181	215	253	300
n	31	32	33	34	35	36	37	38	39	40	41	42	43	44
D	351	415	485	569	665	777	904	1054	1223	1421	1645	1905	2200	2543
n	45	46	47	48	49	50	51	52	53	54	55	56	57	58
D	2929	3375	3879	4461	5114	5868	6716	7686	8782	10030	11437	13040	14841	16888
n	59	60	61	62	63	64	65	66	67	68	69	70	71	72
D	19190	21799	24727	28043	31761	35960	40667	45973	51913	58600	66080	74482	83876	94416
n	73	74	75	76	77	78	79	80	81	82	83	84	85	86
D	106179	119365	134072	150524	168868	189358	212176	237646	265977	297558	332666	371756	415165	463454
n	87	88	89	90	91	92	93	94	95	96	97	98	99	100
D	517032	576564	642571	715835	796997	887002	986631	1096998	1219086	1354211	1503550	1668712	1851106	2052643
n	101	102	103	104	105	106	107	108	109	110	111	112	113	114
D	2275056	2520605	2791384	3090106	3419284	3782133	4181719	4621841	5106175	5639272	6225525	6870327	7578971	8357846
n	115	116	117	118	119	120								
D	9213269	10152854	11184127	12316088	13557775	14919808								