

PAPER • OPEN ACCESS

Classification of complex systems by their sample-space scaling exponents

To cite this article: Jan Korbelt *et al* 2018 *New J. Phys.* **20** 093007

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.



PAPER

Classification of complex systems by their sample-space scaling exponents

OPEN ACCESS

RECEIVED

28 June 2018

REVISED

14 August 2018

ACCEPTED FOR PUBLICATION

24 August 2018

PUBLISHED

10 September 2018

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Jan Korbel^{1,2} , Rudolf Hanel^{1,2} and Stefan Thurner^{1,2,3,4,5}¹ Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria² Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1080 Vienna, Austria³ Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, United States of America⁴ IIASA, Schlossplatz 1, A-2361 Laxenburg, Austria⁵ Author to whom any correspondence should be addressed.E-mail: jan.korbel@meduniwien.ac.at, rudolf.hanel@meduniwien.ac.at and stefan.thurner@meduniwien.ac.at**Keywords:** scaling expansion, extensive entropy, super-exponential systems, complex systems, sample space**Abstract**

The nature of statistics, statistical mechanics and consequently the thermodynamics of stochastic systems is largely determined by how the number of states $W(N)$ depends on the size N of the system. Here we propose a scaling expansion of the phasespace volume $W(N)$ of a stochastic system. The corresponding expansion coefficients (exponents) define the universality class the system belongs to. Systems within the same universality class share the same statistics and thermodynamics. For sub-exponentially growing systems such expansions have been shown to exist. By using the scaling expansion this classification can be extended to *all* stochastic systems, including correlated, constraint and super-exponential systems. The extensive entropy of these systems can be easily expressed in terms of these scaling exponents. Systems with super-exponential phasespace growth contain important systems, such as magnetic coins that combine combinatorial and structural statistics. We discuss other applications in the statistics of networks, aging, and cascading random walks.

1. Introduction

Classical statistical physics typically deals with large systems composed of weakly interacting components, which can be decomposed into (practically) independent sub-systems. The phasespace volume W or the number of states of such systems grows exponentially with system size N . For example, the number of configurations in a spin system of N independent spins is $W(N) = 2^N$. For more complicated systems, however, where particles interact strongly, which are path-dependent, or whose configurations become constrained, exponential phasespace growth no-longer occurs, and things become more interesting. For example, in black holes the accessible number of states does not scale with the volume but with surface, which leads to non-standard entropies and thermodynamics [1–3]. A version of entropy that depends on the surface and the volume was recently suggested in [4].

Other examples include systems with interactions on networks, path-dependent processes, co-evolving systems, and many driven non-equilibrium systems. These systems are often non-ergodic and are referred to as *complex systems*. For these systems, in general, the classical statistical description based on Boltzmann–Gibbs statistical mechanics fails to make correct predictions with respect of the thermodynamic, the information theoretic, or the maximum entropy related aspects [5]. Often the underlying statistics is then dominated by fat-tailed distributions, and power-laws in particular. There have been considerable efforts to understand the origin of power-law statistics in complex systems. Some progress was made for systems with sub-exponentially growing phasespace. It was shown that systems whose phasespace grow as power laws, $W(N) \sim N^b$, are tightly related to so-called Tsallis statistics [6].

The tremendous variety and richness of complex systems has led to the question whether it is possible to classify them in terms of their statistical behavior. Given such a classification, is it possible to arrive at a generalized concept of the statistical physics of complex systems, or do we have to establish the statistical physics

framework for every particular system independently? For sub-exponentially growing systems such a classification was attempted by characterizing stochastic systems in terms of two scaling exponents of their extensive entropy [7]. The first scaling exponent is recovered from the relation $\frac{S(\lambda W)}{S(W)} \sim \lambda^c$, which is valid if the first three Shannon–Khinchin (SK) axioms (see appendix) are valid (the fourth, the composition axiom, can be violated), and if the entropy is of so-called trace form, which means that it can be expressed as $S = \sum_i^W g(p_i)$, where p_i is the probability for state i , and g some function. The second scaling exponent d is obtained from a scaling relation that involves the re-scaling of the number of states $W \rightarrow W^a$. With these two scaling exponents c and d it becomes possible to classify sub-exponentially growing systems that fulfil the first three SK axioms [7]. Further, the exponents c and d characterize the extensive entropy, $S_{c,d} \sim \sum \Gamma(1 + d, c \log(p_i))$. Practically all entropies that were suggested within the past three decades, are special cases of this (c, d) -entropy, including Boltzmann–Gibbs–Shannon entropy ($c = 1, d = 1$), Tsallis entropy ($d = 0$), Kaniadakis entropy ($c = 1, d = 1$) [8], Anteneodo–Plastino entropy ($c = 1, d > 0$) [9], and all others that fulfil the first three SK axioms. In [10] it was then shown that the exponents c and d are tightly related with phasespace growth of the underlying systems. In fact, they can be derived from the knowledge of $W(N)$, $1/(1 - c) = \lim_{N \rightarrow \infty} NW'/W$, and $d = \lim_{N \rightarrow \infty} \log W(W/(NW')) + c - 1$.

For super-exponential systems such a classification is hitherto missing. These systems include important examples of stochastic complex systems that form new states as a result of the interactions of elements. These are systems that—besides their combinatorial number of states (e.g. exponential)—form additional states that emerge as *structures* from the components. The total number of states then grows super-exponentially with respect to system size, e.g. the number of elements. Stochastic systems with elements that can occupy several states (more than one) and that can form structures with other elements, are generally super-exponential systems. It was pointed out in [11] that such systems might exhibit non-trivial thermodynamical properties.

An example for such systems are magnetic coins of the following kind. Imagine a set of N coins that come in two states, up and down. There are 2^N states. However, these coins are ‘magnetic’, and any two of them can stick to each other, forming a new bond state (neither up nor down). If there are $N = 2$ coins, there are five states: the usual four states, uu, ud, du, dd, and a fifth state ‘bond’. If there are $N = 3$ coins, there are 14 states, the 2^3 combinatorial states, and six states involving bond states: state 9 is bond between coin 1 and 2, with the third coin up, state 10 is the same bond state with the third coin down, state 11 is a bond between 1 and 3 with the second coin up, 12 the same bond with the second coin down, state 13 is a bond between 2 and 3, with the first state up, and finally, state 14 is the bond between 2 and 3 with the first coin down. It can be easily shown that the recursive formula for the number of states is, $W(N + 1) = 2W(N) + NW(N - 1)$, which, for large N , grows as $W(N) \sim N^{N/2} e^{2\sqrt{N}}$, see [11].

In this paper we show that it is indeed possible to find a complete classification of complex stochastic systems, including the super-exponential case. By expanding a generic phasespace volume $W(N)$ in a Poincaré expansion, we will see that for any possibility of phasespace growth, there exists a sequence of unique expansion coefficients that are nothing but scaling exponents that describe systems in their large size limit. The *set* of scaling exponents gives us the full classification of complex systems in the sense that two systems belong to the same universality class, if it is possible to rescale one into the other with exactly these exponents. The framework presented here has been proposed in the appendix of [12] and generalizes the classification approach of [7, 10]. It includes the sub-exponential systems as a special case. We show further that these exponents can be used straight forwardly to express—with a few additional requirements—the corresponding extensive entropy, which is the basis for the thermodynamic properties of the system. Finally, we see in several examples that many systems are fully characterized by a very few exponents. Technical details and auxiliary results are presented in the appendix. We reference the appendix in the corresponding parts of the main text. However, readers may also go through the appendix before they continue reading. We use the following notation for applying a function f for n times, $f^{(n)}(x) = \underbrace{f(\dots(f(x))\dots)}_{n \text{ times}}$.

2. Rescaling phasespace

Suppose that phasespace volume depends on system size N (e.g. number of elements) as $W(N)$. We use the Poincaré asymptotic expansion for the $l + 1$ th logarithm of W ,

$$\log^{(l+1)} W(N) = \sum_{j=0}^n c_j \phi_j(N) + \mathcal{O}(\phi_n(N)), \quad (1)$$

where $\phi_j(N) = \log^{(j+1)}(N)$ for $N \rightarrow \infty$. A uniqueness theorem (see e.g. [13]) states that the asymptotic expansion exists and is uniquely determined for any $W(N)$ for which $\log^{(l+1)} W(N) = \mathcal{O}(\phi_0(N))$, see appendix.

To see how the exponents c_j correspond to scaling exponents, let us define a sequence of re-scaling operations,

$$r_\lambda^{(n)}(x) = \exp^{(n)}[\lambda \log^{(n)}(x)]. \tag{2}$$

For example $r_\lambda^{(0)}(x) = \lambda x$, $r_\lambda^{(1)}(x) = x^\lambda$, etc. Obviously, $r_1^{(n)}(x) = x$. The scaling operations obey the composition rule

$$r_\lambda^{(n)}[r_{\lambda'}^{(n)}(x)] = r_{\lambda\lambda'}^{(n)}(x). \tag{3}$$

We can now investigate the scaling behavior of the phasespace volume in the thermodynamic limit, $N \gg 1$. The leading order of the scaling is given by the first rescaling r_0 . We show in the appendix that the rescaling of phasespace is asymptotically described by

$$W(r_\lambda^{(0)}(N)) \sim r_{\lambda^{c_0^{(l)}}}^{(l)}(W(N)) \Rightarrow \frac{\log^{(l)} W(\lambda N)}{\log^{(l)} W(N)} \sim \lambda^{c_0^{(l)}}, \tag{4}$$

where $c_0^{(l)} \in \mathbb{R}$ is the leading exponent, and l is determined from the condition that $c_0^{(l)}$ should be finite. Thus, to leading order, the sample space grows as $W(N) \sim \exp^{(l)}(N^{c_0^{(l)}})$. We now identify the scaling laws for the sub-leading corrections through higher-order rescalings $W(r_\lambda^{(k)}(N))$. We get (see appendix)

$$\frac{\log^{(l)} W(r_\lambda^{(k)}(N))}{\log^{(l)} W(N)} \prod_{j=0}^{k-1} \left(\frac{\log^{(j)}(r_\lambda^{(k)}(N))}{\log^{(j)}(N)} \right)^{-c_j^{(l)}} \sim \lambda^{c_k^{(l)}}. \tag{5}$$

Equivalently, one can express this relation as, $W(r_\lambda^{(k)}(N)) \sim r_{\sigma_k(N)}^{(l)}(W(N))$, where $\sigma_k(N) = \prod_{j=0}^k \left(\frac{\log^{(j)}(r_\lambda^{(k)}(N))}{\log^{(j)}(N)} \right)^{c_j^{(l)}}$.

To extract $c_j^{(l)}$, take the derivative of equation (4) w.r.t. λ , set $\lambda = 1$ and consider the limit $N \rightarrow \infty$. For the leading scaling exponent we obtain

$$c_0^{(l)} = \lim_{N \rightarrow \infty} \frac{NW'(N)}{\prod_{i=0}^l \log^{(i)} W(N)}. \tag{6}$$

The scaling exponent corresponding to the k th order is obtained in a similar way and reads,

$$c_k^{(l)} = \lim_{N \rightarrow \infty} \log^{(k)}(N) \left(\log^{(k-1)}(N) \left(\dots \left(\log(N) \left(\frac{NW'(N)}{\prod_{i=0}^l \log^{(i)} W(N)} - c_0^{(l)} \right) - c_1^{(l)} \right) \dots \right) - c_{k-1}^{(l)} \right). \tag{7}$$

This expression is not identically equal to zero, because the expression on the rhs of equation (6) becomes $c_0^{(l)}$ only in the limit. As a result, the phasespace volume grows as

$$W(N) \sim \exp^{(l)} \left[\prod_{j=0}^n (\log^{(j)}(N))^{c_j^{(l)}} \right], \tag{8}$$

which is nothing but the Poincaré asymptotic expansion in equation (1). In the appendix we show that the formulas for c_j , given by the theory of asymptotic expansions, correspond to the formulas for scaling exponents $c_j^{(l)}$ and therefore it is indeed possible to express *any* $W(N)$ in terms of an asymptotic expansion that is based on the sequence $\phi_n(N)$. The expansion coefficients are scaling exponents determined by the rescaling of phasespace. Here n denotes the minimal number of expansion terms. In the typical situations, only a few scaling exponents are non-zero. If all exponents are non-zero, we can truncate the expansion after a few terms and still preserve a high level of precision. In many realistic situations it is enough to consider $n = 2$. The estimation of the leading order exponent can be tricky, because looking for the order l incorporates calculation of several infinite limits. Therefore, it is convenient to use an approach based on the corresponding extensive entropy.

3. The extensive entropy

The extensive entropy can be obtained by following an idea exposed in [7, 10]. Let us assume a so-called trace form entropy for some probability distribution $P = (p_1, \dots, p_W)$

$$S_g(p) = \sum_{i=1}^W g(p_i), \tag{9}$$

where g is some function. The aim is to find such a function g , for which the entropy functional S_g is *extensive* for a given $W(N)$. Assuming that no prior information about the system is given, we consider uniform probabilities $p_i = 1/W$. The extensivity condition can be expressed by an equation for g , which is [10]

$$S_g(W(N)) = W(N)g(1/W(N)) \sim N \quad \text{for } N \gg 1. \quad (10)$$

Alternatively, it is possible to define the extensive entropy as the solution of Euler's differential equation, see also [4],

$$N \frac{dS(W(N))}{dN} = S(W(N)). \quad (11)$$

The question now is, how the scaling exponents of $W(N)$ are related to scaling exponents of $S_g(W)$. We begin with the first scaling operation $r^{(0)}$. One can show that for $N \gg 1$, we have

$$S_g(r_\lambda^{(0)}(W)) \sim r_{\lambda^{d_0}}^{(0)}(S_g(W)) \Rightarrow \lambda \frac{g\left(\frac{1}{\lambda W(N)}\right)}{g\left(\frac{1}{W(N)}\right)} \sim \lambda^{d_0}. \quad (12)$$

Thus, $g(x) \sim (1/x)^{d_0-1}$ for $x \rightarrow 0$. Again, it is possible to determine the relation for the n th scaling exponent

$$\frac{g(1/r_\lambda^{(n)}(W))r_\lambda^{(n)}(W)}{g(1/W)W} \prod_{j=0}^{k-1} \left(\frac{\log^{(j)}(r_\lambda^{(n)}(W))}{\log^{(j)}(W)} \right)^{-d_j} \sim \lambda^{d_n}, \quad (13)$$

or equivalently, $S_g(r_\lambda^{(n)}(W)) \sim r_{\rho_n(W)}^{(0)}(S_g(W))$, where $\rho_n(W) = \prod_{j=0}^n \left(\frac{\log^{(j)}(\lambda^{c_k^{(l)}}(W))}{\log^{(j)}(W)} \right)^{d_j}$. We can extract the scaling exponents d_n by the same procedure as for $c_k^{(l)}$ by taking the derivative w.r.t. λ , setting $\lambda = 1$ and performing the limit. For the first exponent we get

$$d_0 = \lim_{W \rightarrow \infty} \left(1 - \frac{g'(1/W)}{Wg(1/W)} \right). \quad (14)$$

De L'Hospital's rule and applying the extensivity condition of equation (10) gives $g'(W(N)) \sim N$, and

$$d_0 = \lim_{N \rightarrow \infty} \frac{W(N)}{NW'(N)}. \quad (15)$$

We mentioned this result already above. The n th term can be found analogously to be

$$d_n = \lim_{N \rightarrow \infty} \log^{(n)}(W) \left(\log^{(n-1)}(W) \left(\dots \left(\log(W) \left(\frac{W(N)}{NW'(N)} - d_0 \right) - d_1 \right) \dots \right) - d_{n-1} \right). \quad (16)$$

We can now relate the scaling exponents $c_k^{(l)}$ and d_n by comparing equations (7) and (16). For this we use a similar notation as for the exponents $c_k^{(l)}$ and assign $d_0^{(l)} \equiv d_1$ to the first non-zero exponent, $d_1 \neq 0$. All higher terms are denoted by $d_k^{(l)} = d_{l+k}$. Using the fact that $N \sim (\log^{(l)} W)^{1/c_0^{(l)}}$, we finally obtain

$$\begin{aligned} d_0^{(l)} &= \frac{1}{c_0^{(l)}} \\ d_k^{(l)} &= -\frac{c_k^{(l)}}{c_0^{(l)}}, \quad k = 1, 2, \dots \end{aligned} \quad (17)$$

The corresponding extensive entropy can now be characterized by the function $g(x)$, which scales as

$$g^{(l,n)}(x) \sim x \prod_{j=0}^n \left(\log^{(j+l)} \frac{1}{x} \right)^{d_j^{(l)}} \quad \text{for } x \rightarrow 0. \quad (18)$$

the corresponding entropy scales as

$$S_g^{(l,n)}(W) \sim \prod_{j=0}^n (\log^{(j+l)} W)^{d_j^{(l)}}. \quad (19)$$

This equation is nothing but the asymptotic expansion of $\log S_g$ in terms of $\phi_{n+l}(N) = \log^{(n+l+1)}(N)$; the coefficients are again the scaling exponents that correspond to the rescaling of the entropy.

Note that the entropy approach allows us to obtain additional restrictions for the scaling exponents if further information about the system is available. For example, many systems fulfil the first three of the four SK axioms, see appendix. There we also show that it is possible to find a representation of the entropy that obeys the three axioms and the scaling in equation (19). In this case $g(x)$ can be expressed as

$$g_{(d_0^{(l)}, \dots, d_n^{(l)})}^{(l,n)}(x) = \int_0^x \prod_{i=0}^n \left[a_i + [1 + \log]^{(i+l)} \left(\frac{1}{y} \right) \right]^{d_i^{(l)}} dy, \quad (20)$$

where a_i are constants. One possible choice for those is

$$a_i = \max \left\{ -1 - \frac{d_i^{(l)}}{(n-l)d_0^{(l)}}, 0 \right\}. \quad (21)$$

The axioms impose restrictions on the range of scaling exponents. (SK2) requires that $d_0^{(l)} > 0$; (SK3) requires that $d_0^{(0)} \equiv d_0 < 1$. The resulting entropy can be expressed by equation (12). One can trivially adjust the entropy minimal value, such that for the totally ordered state, $\mathcal{S}_g(1) = 0$. This is obtained by rescaling

$$\mathcal{S}_g(P) = r_\lambda^{(-1)}(\mathcal{S}_g(P)) = \left(\sum_{i=1}^W g(p_i) \right) - g(1), \quad (22)$$

where $\lambda = \exp(g(1))$. Note that the form of the entropy in equation (20) is equivalent to (c, d) -entropy for $c = 1 - d_0$ and $d = d_1$, and $d_j = 0$ for all $j \geq 2$.

4. Examples

We conclude with several examples of systems that are characterized by different sets of scaling exponents.

4.1. Exponential growth: the random walk

Imagine the ordinary random walk with two possibilities at any timestep—a step to the left, or to the right. The number of possible configurations (i.e. possible paths) after N steps is

$$W(N+1) = 2W(N), \quad (23)$$

which means exponential phasespace growth, $W(N) = 2^N$. We obtain $l = 1$, $c_0^{(1)} = 1$ and $c_j^{(1)} = 0$, for $j \geq 1$, and for the exponents of the entropy $d_0 = 0$, $d_1 \equiv d_0^{(1)} = 1$ and $d_j = 0$, for $j \geq 2$. This set of exponents belongs to the class of (c, d) -entropies described in [7] for $c = 1 - d_0 = 1$, and $d = d_1 = 1$. They correspond to the scaling exponents of the Shannon entropy: from (18) we obtain that $g(x) \sim x \log x$ and from (19) we get $S(W) \sim \log W$, which is Boltzmann entropy. It is not immediately apparent what the entropy of a random walk should be. However, the random walk is equivalent to spin system of N independent spins, the 2^N different paths correspond one-to-one to the 2^N configurations in the spin model, where the role entropy of it is clear. Obviously, for the random walk, (SK1–3) are applicable.

4.2. Sub-exponential growth: the aging random walk

In this variation of the random walk we impose correlations on the walk. After the first random choice (left or right) the walker goes one step in that direction. The second random choice is followed by two steps in the same direction, the next step is followed by three steps in the same direction, etc. For k independent choices, one has to make $N = \sum_{i=1}^{k-1} i = 1/2k(k-1)$ steps. For this walk, we get that the number of possible paths is

$$W(N+k) = 2W(N), \quad (24)$$

which leads to $W(N) = 2^{N/k} \sim 2^{k/2}$. For $N \gg 1$, we have $k \approx \sqrt{N}$, and we obtain a stretched exponential (sub-exponential) asymptotic behavior, $W(N) \sim 2^{\sqrt{N}}$. The order is again $l = 1$ and the exponents are $c_0^{(1)} = 1/2$ and $c_j^{(1)} = 0$, for $j \geq 1$. In terms of the d exponents we have $d_0 = 0$ and $d_1 \equiv d_0^{(1)} = 2$. Therefore, the three SK axioms are applicable and the resulting extensive entropy belongs to the class of entropies characterized by the Anteneodo–Plastino entropy, since we have $g(x) \sim x(\log x)^2$ and $S(W) \sim (\log W)^2$. This entropy is the special case of the (c, d) -entropy for $c = 1$ and $d = 2$, see [7].

4.3. Super-exponential growth: magnetic coins

Consider N coins with two states (up or down). These coins are magnetic, so that any two can stick to each other to create a pair which is a third state obtained by interactions of elements (one possible configuration). As mentioned before, in [11] it is shown that the phasespace volume can be obtained recursively

$$W(N+1) = 2W(N) + NW(N-1). \quad (25)$$

For $N \gg 1$, we get $W(N) \sim N^{N/2} e^{2\sqrt{N}}$, which yields $l = 1$, and the scaling exponents $c_0^{(1)} = 1$, $c_1^{(1)} = 1$ and $c_j^{(1)} = 0$, for $j \geq 2$. The scaling exponents of the entropy are $d_0 = 0$, $d_1 \equiv d_0^{(1)} = 1$, and $d_2 \equiv d_1^{(1)} = -1$. For the entropy this means, that $g(x) \sim x \log x / \log(\log x)$ and $S(W) \sim \log W / \log(\log W)$. This case is *not* contained in the class of (c, d) -entropies, because the third exponent, corresponding to the doubly-logarithmic correction, is not zero. Actually we obtain $c = 1$ and $d = 1$, which would naively indicate Shannon entropy. However, the correction makes the system clearly super-exponential. The SK axioms are still applicable, the class of accessible entropy formulas is restricted by (SK2). For example, for the representative entropy equation (20) we find that $a_0 \geq 0$ and $a_1 \geq 0$, see appendix.

4.4. Super-exponential growth: random networks

Imagine a random network with N nodes. When a new node is added, there emerge N new possible links, which gives us 2^N new possible configurations for each configuration of the network with N links. We obtain the recursive growth equation

$$W(N+1) = 2^N W(N), \quad (26)$$

which leads to $W(N) = 2^{\binom{N}{2}}$, as expected. For this phasespace growth, we obtain $l = 1$, $c_0^{(1)} = 2$ and $c_j^{(1)} = 0$ for $j \geq 1$, and $d_0 = 0$ and $d_1 \equiv d_0^{(1)} = \frac{1}{2}$. The corresponding entropy can be expressed by $g(x) \sim x(\log x)^{1/2}$, and $S(W) \sim (\log W)^{1/2}$. The entropy corresponds to the class of compressed exponentials, which are super-exponential, however, the entropy belongs to the class of (c, d) -entropies for $c = 1$ and $d = 1/2$. Because all exponents are positive the entropy observes the SK axioms.

4.5. Super-exponential growth: the cascading random walk

Consider a generalization of the random walk, where a walker can take a left or right step, but it can also split into two walkers, one of which then goes left, the other to the right. Each walker can then go left, right, or split again (multiple walkers can occupy the same position). The number of possible paths after N steps is

$$W(N+1) = 2W(N) + W(N)^2, \quad (27)$$

where the first term reflects the left/right decisions, the second the splittings. We have $W(N) = 2^{(2^{N-1})} - 1$, and find that $l = 2$, $c_0^{(2)} = 1$ and $c_j^{(2)} = 0$, for $j \geq 1$, and $d_0 = 0$, $d_1 = 0$ and $d_2 \equiv d_0^{(2)} = 1$. The corresponding extensive entropy is $g(x) \sim x \log \log(x)$ and scales as $S(W) \sim \log \log W$. Because the coefficients are not negative, SK axioms are applicable. However, even though all correction scaling exponents are zero, the system cannot be described in terms of (c, d) -entropies, because $l = 2$. We would naively obtain that $c = 1$ and $d = 0$, which would wrongly correspond to Tsallis entropy. Alternatively, we can think of an example of a spin system with the same scaling exponents. In this case, N would not describe the size of a system, but its dimension. For $N = 1$, we would have two particles on the line, for $N = 2$ we have 4 particles forming a square, for $N = 3$ we have a cube with 8 particles in its vertices, etc. In general, we can think of a spin system of particles sitting on the vertices of a N -dimensional hypercube. The number of particles is naturally 2^N and for two possible spins we obtain $W(N) = 2^{(2^N)}$.

5. Conclusions

We introduced a comprehensive classification of complex systems in the thermodynamic limit based on the rescaling properties of their phasespace volume. From a scaling-expansion of the phasespace growth with system size, we obtain a set of scaling exponents, which uniquely characterize the statistical structure of the given system. Restrictions on the scaling exponents can be obtained with further information about the system. In this context we discuss the first three SK axioms, which are valid for many complex systems. The set of exponents further determine the scaling exponents of the corresponding extensive entropy, which plays a central role in the thermodynamics of statistical systems. Thermodynamics is not the only context where entropy appears. As was shown in [5] for many complex systems the functional expressions for entropy depend on the context, in particular if one talks about the thermodynamic (extensive) entropy, the information theoretic entropy, or the entropy that appears in the maximum entropy principle. It remains to be seen if for super-exponential systems there exists an underlying relation between the scaling exponents of the extensive entropy, and the exponents obtained from an information theoretic, or maximum entropy description of the same complex systems.

Acknowledgments

We thank the participants of CSH workshop for helpful initial discussion, in particular Henrik Jeldtoft Jensen, Tamás Sándor Biró, Piergiulio Tempesta and Jan Naudts. This work was supported by the Austrian Science Fund (FWF) under project I3073.

Appendix

A.1. SK axioms

The SK axioms read:

- (SK1) Entropy is a continuous function of the probabilities p_i only, and should not explicitly depend on any other parameters.

- (SK2) Entropy is maximal for the equi-distribution $p_i = 1/W$.
- (SK3) Adding a state $W + 1$ to a system with $p_{W+1} = 0$ does not change the entropy of the system.
- (SK4) Entropy of a system composed of 2 sub-systems A and B , is $S(A + B) = S(A) + S(B|A)$.

They state requirements that must be fulfilled by any entropy. For ergodic systems all four axioms hold. For non-ergodic ones the composition axiom (SK4) is explicitly violated, and only the first three (SK1–SK3) hold. If all four axioms hold the entropy is uniquely determined to be Shannon's; if only the first three axioms hold, the entropy is given by the (c, d) -entropy [7, 10]. The SK axioms were formulated in the context of information theory but are also sensible for many physical and complex systems.

Given a trace form of the entropy as in equation (9), the SK axioms imply the restrictions on $g(x)$: (SK1) implies that g is a continuous function, (SK2) means that $g(x)$ is concave, and (SK3) that $g(0) = 0$. For details, see [7].

A.2. Rescaling in the thermodynamic limit

We first prove a theorem which determines the general form of rescaling relations in the thermodynamic limit for any general function.

Theorem. Let $g(x)$ be a positive, continuous function on \mathbb{R}^+ . Let us define the function $z(\lambda): \mathbb{R}^+ \rightarrow \mathbb{R}^+$

$$z(\lambda) := \lim_{x \rightarrow \infty} \frac{g(r_\lambda^{(n)}(x))}{g(x)}. \quad (\text{A.1})$$

Then, $z(\lambda) = \lambda^c$ for some $c \in \mathbb{R}$.

Proof. From the definition of $z(\lambda)$, it is straightforward to show that $z(\lambda\lambda') = z(\lambda)z(\lambda')$, because

$$\begin{aligned} z(\lambda\lambda') &= \lim_{x \rightarrow \infty} \frac{g(r_{\lambda\lambda'}^{(n)}(x))}{g(x)} = \lim_{x \rightarrow \infty} \frac{g(r_{\lambda\lambda'}^{(n)}(x))}{g(r_\lambda^{(n)}(x))} \frac{g(r_\lambda^{(n)}(x))}{g(x)} \\ &= \lim_{r_\lambda^{(n)}(x) \rightarrow \infty} \frac{g(r_{\lambda'}^{(n)}[r_\lambda^{(n)}(x)])}{g(r_\lambda^{(n)}(x))} \lim_{x \rightarrow \infty} \frac{g(r_\lambda^{(n)}(x))}{g(x)} = z(\lambda')z(\lambda). \end{aligned}$$

For the computation we used the group property of rescaling in equation (3) and the continuity of g . The only class of functions satisfying the functional equation above are power functions, $z(\lambda) = \lambda^c$. \square

Let us take the first scaling relation of the sample space $W(r_\lambda^{(0)}(N)) = W(\lambda N)$. From the previous theorem we obtain

$$\frac{W(\lambda N)}{W(N)} \sim \lambda^{c_0} \Rightarrow W(r_\lambda^{(0)}(N)) \sim r_{\lambda^{c_0}}^{(0)}(W(N)). \quad (\text{A.2})$$

It may happen that c_0 is infinite. Thus, we may need to use higher-order scaling for the sample space, i.e., $r_{\lambda^{c_0}}^{(l)}(W(N))$, as shown in the main text. l is determined by the condition that the scaling exponent should be finite. The first correction term is given by the scaling $W(r_\lambda^{(1)}(N)) = W(N^\lambda)$. To obtain the sub-leading correction, we have to factor out the leading growth term. This means that the scaling relation for the first sub-leading correction looks like

$$\frac{(\log^{(l)} W(N^\lambda))/N^{c_0^{(l)}\lambda}}{(\log^{(l)} W(N))/N^{c_0^{(l)}}} \sim \lambda^{c_1^{(l)}}, \quad (\text{A.3})$$

which is again a consequence of the above theorem. To obtain the corresponding scaling relations for higher-order scaling exponents for the sample space (A.4), we need to factor out all previous terms corresponding to lower-order scalings, so the scaling relation looks like

$$\frac{\log^{(l)} W(r_\lambda^{(k)}(N))}{\log^{(l)} W(N)} \prod_{j=0}^{k-1} \left(\frac{\log^{(j)}(r_\lambda^{(k)}(N))}{\log^{(j)}(N)} \right)^{-c_j^{(l)}} \sim \lambda^{c_k^{(l)}}. \quad (\text{A.4})$$

Because the left-hand side of this relation has the form of the function z appearing in the theorem, the validity of the relation is satisfied for $N \rightarrow \infty$. Similarly, we can deduce the relations for scaling exponents that are associated with the extensive entropy.

A.3. Asymptotic expansion in terms of nested logarithms

The asymptotic representation of $W(N)$ is obtained by the rescaling that corresponds to the Poincaré asymptotic expansion [13] of $\log^{(l+1)}(W)$ in terms of $\phi_n(N) = \log^{(n+1)}(N)$ for $N \rightarrow \infty$. Let us consider a function $f(x)$ with a singular point at x_0 . It is possible to express its asymptotic properties in the neighborhood of x_0 in terms of the asymptotic series of functions $\phi_n(x)$, if $f(x) = \mathcal{O}(\phi_0(x))$ and $\phi_{n+1}(x) = \mathcal{O}(\phi_n(x))$. The series is given as

$$f(x) = \sum_{j=0}^k c_j \phi_j(x) + \mathcal{O}(\phi_k(x)). \tag{A.5}$$

The coefficients can be calculated from the formulas in [13]

$$c_k = \lim_{x \rightarrow x_0} \frac{f(x) - \sum_{j=0}^{k-1} c_j \phi_j(x)}{\phi_k(x)}. \tag{A.6}$$

In our case, i.e., for $N \rightarrow \infty$ and $\phi_n(N) = \log^{(n+1)}(N)$ the function $\log^{(l+1)}(W)$ can be expressed (for appropriate l) in terms of this series, and the coefficients $c_k^{(l)}$ are given by

$$\begin{aligned} c_k^{(l)} &= \lim_{N \rightarrow \infty} \frac{\log^{(l+1)}(W) - \sum_{j=0}^{k-1} c_j^{(l)} \log^{(j+1)}(N)}{\log^{(k+1)}(N)} \\ &= \lim_{N \rightarrow \infty} \frac{\log(\log^{(l)}(W)) / \prod_{j=0}^{k-1} \log^{(j)}(N)^{c_j^{(l)}}}{\log^{(k+1)}(N)}. \end{aligned}$$

Using L'Hospital's rule and the derivative of the nested logarithm

$$\frac{d \log^{(n)}(x)}{dx} = \frac{1}{\prod_{j=0}^{n-1} \log^{(j)}(x)}, \tag{A.7}$$

a straightforward calculation yields equation (7).

A.4. Derivation of $g_{(d_0^{(l)}, \dots, d_n^{(l)})}^{(l,n)}$

Which entropy functional that fulfills axioms (SK1–3)? The choice is not unique, but a concrete entropy functional serves as a representative of the class in the thermodynamic limit. The requirements imposed by the first three SK axioms are: $g(x)$ is continuous, $g(x)$ is concave, and $g(0) = 0$. From equation (18) we have,

$$g(x) \sim x \prod_{j=0}^n \left[\log^{(j+1)}\left(\frac{1}{x}\right) \right]^{d_j^{(l)}} \text{ for } x \rightarrow 0, \text{ which gives us the scaling for the values around zero.}$$

Unfortunately, the presented form cannot be extended to the full interval $[0, 1]$, because the domain of $\log^{(n)}(1/x)$ is $(0, 1/\exp^{(n-2)}(1))$. This can be fixed by replacing $\log^{(n)}$ by $[1 + \log]^{(n)} = 1 + \log(1 + \log(\dots))$, which is defined on the whole domain $(0, 1]$, where $\lim_{x \rightarrow 0} [1 + \log]^{(n)}(1/x) = +\infty$ and $[1 + \log]^{(n)}(1) = 1$. The scaling remains unchanged for $x \rightarrow 0$.

The second problem is that in general the function is not concave. For this we introduce the transformation

$$f^*(x) = \int_0^x \frac{f(y)}{y} dy. \tag{A.8}$$

The original function can be obtained by

$$f(x) = x \frac{df^*(x)}{dx}. \tag{A.9}$$

This transform turns an increasing/decreasing function to a convex/concave function, while the scaling for $x \rightarrow 0$ remains unchanged. Let us write the function g in the form of the transform

$$g(x) \sim \int_0^x \prod_{j=0}^n \left[[1 + \log]^{(j)}\left(\frac{1}{y}\right) \right]^{d_j} dy. \tag{A.10}$$

Axiom (SK3) means $g(0) = 0$. This requires that the integrand should not diverge faster than $1/x$ for $x \rightarrow 0$. This can be fulfilled for $d_0 \equiv d_0^{(0)} < 1$.

Because $[1 + \log]^{(n)}(1/x)$ is a decreasing function, $g(x)$ is automatically concave if $d_n \geq 0$, since a product of positive, decreasing functions is also decreasing. However, for $d_n < 0$, $[1 + \log]^{(n)}(1/x)^{d_n}$ is an increasing function from zero to one and the whole product may not be decreasing. In order to solve this issue, we introduce a set of constants a_i and write $g(x)$ in the form

$$g(x) = \int_0^x \prod_{j=0}^n \left[a_j + [1 + \log]^{(n)} \left(\frac{1}{y} \right) \right]^{d_j} dy. \tag{A.11}$$

The constants a_i can be chosen to ensure that the integrand is a decreasing function. We assume $a_i \geq -1$ to avoid problems with powers of negative numbers. The second derivative of $g(x)$, i.e., the first derivative of the integrand is an increasing function and $\frac{d^2g(x)}{dx^2}|_{x \rightarrow 0^+} = -\infty$ for $d_l > 0$. For $d_l < 0$, the entropy cannot be concave, so $d_l > 0$ is the restriction given by (SK2). To obtain a negative second derivative on the whole domain $[0, 1]$, it is therefore enough to investigate $\frac{d^2g(x)}{dx^2}|_{x=1}$, which leads to the condition

$$\left(\prod_{j=l}^n (1 + a_j)^{d_j-1} \right) \left(-\sum_{j=l}^n \frac{d_j}{1 + a_j} \right) \leq 0. \tag{A.12}$$

Because $d_0^{(l)} \equiv d_l > 0$, we can choose $a_l = 0$. In the following terms, i.e., for $i > l$, d_i can be both positive and negative. Positive d_i pose no problem, because the term corresponding to d_i , i.e. $-d_i/(1 + a_i)$ is negative, so we can choose $a_i = 0$. When all d_i are negative we can compensate the positive contribution of the negative terms by diminishing them through choice of appropriate a_i . If we choose

$$1 + a_i = -\frac{d_i^{(l)}}{nd_0^{(l)}}, \tag{A.13}$$

then equation (A.12) becomes zero. If this is given together with previous results and we summarize it as

$$a_i = \max \left\{ -1 - \frac{d_i^{(l)}}{nd_0^{(l)}}, 0 \right\}, \tag{A.14}$$

which has been presented in equation (21) in the main text. Clearly, this is not the only possible choice. Note that for all $d_i^{(l)} > 0$, one may even choose $a_i = -1$. On the other hand, for the case of the magnetic coin model, one obtains that for $a_0 = 0$, $a_1 = 0$ as well.

Finally, let us show the connection to (c, d) -entropy derived in [7]. In this case, we assume only d_0 and d_1 can be non-zero, which leads to

$$g_{(d_0, d_1)}^{(0,1)}(x) = \int_0^x (1/y)^{d_0} (1 + a_1 + \log(1/x))^{d_1} dy. \tag{A.15}$$

By the choice $a_1 = -1 + \frac{1}{1-d_0}$, we get

$$g_{(c,d)}(x) = \frac{e}{c^{d+1}} \Gamma(1 + d, 1 - c \log x), \tag{A.16}$$

for $c = 1 - d_0$ and $d = d_1$, which is nothing else than the gamma entropy of [7].

A.5. Ordering of processes and classes of equivalence

The set of scaling exponents form natural classes of equivalence with natural ordering. Consider two discrete random processes $X(N)$ and $Y(N)$ with sample spaces $W_X(N)$ and $W_Y(N)$, respectively. The corresponding sets of scaling exponents are denoted by $\mathcal{C}_X = \{c_0^{(l)}, c_1^{(l)}, \dots\}$, and $\mathcal{C}_Y = \{\tilde{c}_0^{(l)}, \tilde{c}_1^{(l)}, \dots\}$. One can introduce an ordering based on the scaling exponents. We write

$$X \prec Y \ (\mathcal{C}_X \prec \mathcal{C}_Y) \text{ if } \begin{cases} l < l' \\ l = l', c_0^{(l)} < \tilde{c}_0^{(l')} \\ l = l', c_0^{(l)} = \tilde{c}_0^{(l')}, c_1^{(l)} < \tilde{c}_1^{(l')} \\ \text{etc.} \end{cases} \tag{A.17}$$

This is equivalent to lexicographic ordering. One can also introduce an ordering, which takes into account only certain a number of correcting terms. So, for example

$$X \prec_0 Y \ (\mathcal{C}_X \prec_0 \mathcal{C}_Y) \text{ if } \begin{cases} l < l' \\ l = l', c_0^{(l)} < \tilde{c}_0^{(l')} \end{cases} \tag{A.18}$$

Similarly, one can define \prec_k , which takes into account only k correction terms. Additionally, it is possible to introduce an equivalence relation

$$X \sim Y \text{ if } \mathcal{C}_X \equiv \mathcal{C}_Y \Rightarrow l = l'; c_i^{(l)} = \tilde{c}_i^{(l')} \ \forall i \tag{A.19}$$

and also equivalence up to certain correction

$$X \sim_k Y \text{ if } \mathcal{C}_X \equiv \mathcal{C}_Y \Rightarrow l = l'; c_i^{(l)} = \tilde{c}_i^{(l')} \ \forall i \leq k. \tag{A.20}$$

As an example, for magnetic coin model and random walk we have that $X_{MC} \sim_0 X_{RW}$, but $X_{MC} \not\sim X_{RW}$.

A.6. Construction of a ‘representative process’

To understand the mechanism of how the scaling exponents correspond to the structure of a random process, let us discuss a simple procedure to generally obtain processes with given scaling exponents $c_k^{(l)}$. We start with a random variable X_0 with N possible outcomes, so that $W_{X_0}(N) = \{1, \dots, N\}$. The scaling exponents of this process are naturally $c_0^{(0)} = 1$ and $c_k^{(0)} = 0$ for $k \geq 1$. Let us construct a new variable by choosing subsets of $W_{X_0}(N)$.

First we can create all possible subsets of $W_{X_0}(N)$. This defines a new variable X_1 with $W_{X_1}(N) = 2^{W_{X_0}(N)}$, and we get $c_0^{(1)} = 1$. Generally, the transform

$$2: X \rightarrow 2^X, \tag{A.21}$$

where 2^X denotes a variable on all subsets of X . One can easily show that this results in a shift of scaling exponents $c_k^{(l)} \rightarrow c_k^{(l+1)}$, and $d_k^{(l)} \rightarrow d_k^{(l+1)}$, because $W_{2^X}(N) = 2^{W_X(N)}$. The interpretation of this transformation is the following: consider an ordinary random walk with two possible steps. If $X_0(N)$ denotes a number of steps of a random walker, then $X_1(N) = 2^{X_0(N)}$ denotes the number of possible paths. When we apply the transform again, we obtain $X_2(N) = 2^{X_1(N)}$. This denotes the number of possible configurations of a random walk cascade, etc. As a result, by more applications of 2, we obtain processes with more complicated structure of the respective phasespace.

To construct processes with arbitrary exponents, let us think about a procedure, where we create only partial subsets, which number $p(N)$ can be between N (no partitioning) and 2^N (full partitioning). We denote this procedure by \mathfrak{P} . This can be understood as a process corresponding to a correlated random walk. This means that not every step of the walk is independent, but some steps can be determined by the previous steps, which diminishes the number of possible configurations when compared to the uncorrelated random walk. The resulting random process is obtained as the composition of l uncorrelated random walks (full partitioning) and a correlated random walk

$$X = 2^{(l)}[\mathfrak{P}(X_0)]. \tag{A.22}$$

Let us now focus on the construction of correlated random walk with a pre-determined number of states given by $p(N)$.

First we consider the full set of subsets of N elements with natural ordering,

$$W_{2^X} = \{\{\}, \{1\}, \{2\}, \dots, \{1, \dots, n\}\}. \tag{A.23}$$

The correlations can be represented by merging subsets to $p(N)$ sequences of length $\{s(1), \dots, s(p(N))\}$, i.e.,

$$W_{\mathfrak{P}(X)} = \left\{ \underbrace{\{\{\}, \{1\}, \{2\}, \dots\}}_{s(1)}, \dots, \underbrace{\{\dots, \{1, \dots, n\}\}}_{s(p(N))} \right\}. \tag{A.24}$$

This means that after one independent step, there are $s(1) - 1$ dependent steps, after the second independent step, there are $s(2) - 1$ dependent steps, etc. Let us determine the form of function s for given $p(N)$. The function s can be obtained from

$$\sum_{i=1}^{p(N)} s(i) = 2^N. \tag{A.25}$$

In the limit of large N we can assume that the function s does not depend on N , i.e., is *a priori* given by the scaling exponents of the system. Let us also assume, without loss of generality, that s is an increasing function (we can neglect the last cell, because its size is determined by the size of previous cells). For $N \gg 1$, we approximate the sum by the integral and obtain

$$\int_0^{p(N)} s(i) di \sim 2^N. \tag{A.26}$$

Denoting $S(m) = \int_0^m s(y) dy$, and substituting $x = p(N)$, we recast the previous equation as, $S(x) = 2^{p^{-1}(x)}$, where p^{-1} denotes the inverse function of p . The function $s(x)$ can be therefore determined as

$$s(x) = \frac{d 2^{p^{-1}(x)}}{dx} = \frac{2^{p^{-1}(x)}}{p'(p^{-1}(x))}. \tag{A.27}$$

Some examples for $s(x)$ for a corresponding $p(N)$ are

- $p(N) = 2^N$, i.e., full partitioning corresponding to uncorrelated random walk. In this case, we obtain that $s(x) = \text{const.}$, as expected.
- $p(N) = N$, i.e., no partitioning to maximally correlated random walk. We obtain that $s(x) \sim 2^x$, which can be seen from the relation $\sum_i^N 2^i \sim 2^N$.

- $p(N) = N \log N$, which corresponds to the correction in the magnetic coin model.
In this case, $s(x) \sim 2^{W(x)} / \log(W(x))$, where $W(x)$ is the Lambert W -function.

ORCID iDs

Jan Korbel  <https://orcid.org/0000-0002-5371-5320>

References

- [1] Bekenstein J D 1974 *Phys. Rev. D* **9** 3292–300
- [2] Hawking S W 1974 *Nature* **248** 30–1
- [3] Thirring W 1970 *Z. Phys. A* **235** 339–52
- [4] Biró T S, Czinner V G, Iguchi H and Ván P 2018 *Phys. Lett. B* **782** 228–31
- [5] Thurner S, Corominas-Murtra B and Hanel R 2017 *Phys. Rev. E* **96** 032124
- [6] Tsallis C, Gell-Mann M and Sato Y 2005 *Proc. Natl Acad. Sci.* **102** 15377–82
- [7] Hanel R and Thurner S 2011 *Europhys. Lett.* **93** 20006
- [8] Kaniadakis G 2002 *Phys. Rev. E* **66** 056125
- [9] Anteneodo C and Plastino A R 1998 *J. Phys. A: Math. Theor.* **32** 1089
- [10] Hanel R and Thurner S 2011 *Europhys. Lett.* **96** 50003
- [11] Jensen H J, Pazuki R H, Pruessner G and Tempesta P 2018 *J. Phys. A: Math. Theor.* **51** 375002
- [12] Thurner S and Hanel R 2011 Appendix in arXiv: 1104.2070
- [13] Copson E T 1965 *Asymptotic Expansions Cambridge Tracts in Mathematics* (Cambridge: Cambridge University Press)