J Math Chem (2016) 54:1942–1951
DOI 10.1007/s10910-016-0658-z

ORIGINAL PAPER

Doppelgänger entropies

A. D. Kirwan Jr.[1] · **William Seitz**[2]

Received: 20 January 2016 / Accepted: 20 June 2016 / Published online: 30 June 2016

Abstract We report on a systematic study of Boltzmann entropy as a function of state space size. As the state space, characterized by the number of objects N, is increased we find that identical entropies are shared by many different state space configurations. These degenerate states are called *doppelgänger* states. A calculus is developed to predict the occurrence of these states. Theoretical and numerical analysis shows that for large N almost all configurations are *doppelgängers*. Boltzmann entropy is fundamental to disparate disciplines such as statistical mechanics, mixing theory, combinatorics, and information theory. Our analysis then may have some broad interest.

Keywords Entropy · Statistical mechanics · Boltzmann systems · Mixing · Classical ensemble theory

1 Introduction

In a study of the evolution of a Boltzmann system of dimension N from lowest to maximum entropy [1] we noticed that for $N \geq 7$ certain iso-entropy states are produced by incomparable microstate configurations. For lack of a better term we call these *doppelgänger* entropy states, or perhaps just *doppelgängers*. Since the number of available states of a Boltzmann system goes as the integer partitions of N, (IP[N]), we first speculated this was a rare pathological phenomenon of no significance. However,

✉ A. D. Kirwan Jr.
adk@udel.edu

1 School of Marine Science and Policy, University of Delaware, Newark, DE 19716, USA

2 Department of Marine Sciences, Texas A&M University at Galveston, P.O. Box 1675, Galveston, TX 77553, USA

it turns out that even for a relatively modest value of $N = 50$ where there are IP[50] = 204,226 states, only 10,417 (approximately 5 %) are *not* doppelgängers.

The Boltzmann distribution defines the *microcanonical* ensemble in statistical mechanics. This ensemble describes the distribution in state space of N particles in an isolated system of volume V, with total energy E. This function is well known as

$$\Omega(N, V, E) = \frac{N!}{\prod_{i=1}^{N} n_i!}. \tag{1}$$

Ω has a simple physical interpretation. It is the number of ways N distinguishable objects can be distributed in n_i cells. It is fundamental to all classical canonical ensembles, [2].

The entropy of an isolated system is well known as

$$S = k \ln \Omega(N, V, E). \tag{2}$$

We have not found any discussion of iso-entropy states of (2) in the statistical mechanical literature. Since this distribution is so pervasive in science and technology it may be that doppelgäger Boltzmann entropies are of interest to a wide audience. This motivates our goal here to develop a calculus to predict the occurrence of doppelgäger entropies.

The next section introduces the necessary notation and definitions. This is applied to some simple examples of multiple entropy states in low N systems. Section 3 provides a detailed numerical analysis of cases up to $N = 50$. The paper concludes with a brief synopsis and discussion.

2 Theory

Our starting point is the Boltzmann entropy given by (2), normalized by k

$$S(n_i) = \ln\left(\frac{N!}{\prod_1^N n_i!}\right). \tag{3}$$

It is helpful to use a simpler notation for $S(n_i)$ as

$$S(n_i) = [n_1, n_2, \ldots]. \tag{4}$$

For large N many cells may contain the same number of objects. Then (4) is abbreviated as

$$S(n_i) = [n_1, \ldots n_p^k, \ldots] \tag{5}$$

where k is now the number of cells with n_p objects. Equations (4) or (5) are the signatures of the entropy states.

Now consider two distributions $S(n_i)$ and $S(m_i)$ with the same number of objects. Clearly the entropies will be the same iff

$$\prod n_i! = \prod m_i!. \tag{6}$$

The goal here is to develop a calculus for determining cases when (6) is true. The key is to find seed numbers whose factorials are the product of other factorials. Thus, we look for some n_j in (6) that can be expressed as the product of two or more factorials. That is let

$$n_j! = n_{j-a}!n_{j-b}! \ldots . \tag{7}$$

Then

$$[n_1, \ldots n_j, \ldots] = [n_1, \ldots, (n_{j-a}, n_{j-b}, \ldots), \ldots]. \tag{8}$$

Now $n_{j-a} + n_{j-b} + \cdots - n_j = N_d$ so N_d cells with 1 or more objects must be added to $S(n_i)$ to achieve a mass balance with $S(m_i)$ in (6). An important ramification is that more than N_d cells can be added to $S(n_i)$ as long as the additional cells are added in the same combination to $S(m_i)$.

This idea is illustrated for small n_j where the seeds are easy to establish. Consider the following seeds

$$\begin{aligned}
&4! = 3!2!2!, && 6! = 5!3!, && 8! = 7!2!2!2!\\
&9! = 7!3!3!2!, && 10! = 7!6!, && 10! = 7!5!3!\\
&12! = 11!3!2!, && 16! = 15!2!2!2!2!, && 16! = 14!5!2!\\
&24! = 23!4!
\end{aligned} \tag{9}$$

The simplest case is $4! = 3!2!2!$. In (7) take $j = 4$, $a = 1$, $b = c = 2$. Then the deficit number in this case is $N_d = -3$. So a mass balance is achieved if 3 or more objects are added to left hand side of (8) to achieve

$$[4, 1^3] = [3, 2^2]. \tag{10}$$

The total number of particles is $N = 7$, which is this lowest dimensional state space that has a doppelgänger entropy! In the same fashion the seeds in (9) give rise to other doppelängers. Inspection shows that another doppelgänger occurs at $N = 8$ two at $N = 13$ and another two at $N = 15$, just to name a few.

Note that (10) is a DNA fingerprint that will reappear for state dimensions $N > 7$. It is readily seen that

$$[4, \mathbf{q}^{\mathbf{k}}, 1^3] = [3, \mathbf{q}^{\mathbf{k}}, 2^2] \tag{11}$$

where $\mathbf{q}^{\mathbf{k}}$ is any combination of integer $\mathbf{q}$ and $\mathbf{k}$ that satisfy

$$\mathbf{q}^{\mathbf{k}} = N - 7. \tag{12}$$

Here $\mathbf{q}^{\mathbf{k}} = q_1^{k_1}, q_2^{k_2}, \ldots$. Thus the $\mathbf{q}^{\mathbf{k}}$ in (11) are seeds for doppelgängers that grow with N as IP$[N - 7]$. The same analysis applies to all seeds. Of course as N increases new doppelgänger seeds will arise. For example when $N = 10$, in addition to the five doppelgängers that come from $N = 7$ and $N = 8$, the doppelgänger $[6, 2^2] = [5, 4, 1]$ arises.

Since the number of doppelgängers with the same entropy increases as an IP it is obvious that for even a modest value of N the number of states with the same entropy will exceed the state size. Recall that for $N = 13$ there are two new doppelgängers, six more arising from the smallest seed $N - 7$, and five additional from the second seed $N - 8$. For $N = 14$ the number of doppelgängers arising from just the seed $N - 7$ is 15.

In addition, we have found that doppelgängers from different seeds will intersect at sufficiently large N to produce vielgänger entropy states. This is illustrated by the intersection of the two lowest seeds in (9); namely $[4, \mathbf{q^k}, 1^3] = [3, \mathbf{q^k}, 2^2]$ and $[6, \mathbf{q^k}, 1^2] = [5, \mathbf{q^k}, 3]$. One intersection occurs when $[3, q_1^{k_1}, 2^2] = [5, q_2^{k_2}, 3]$. This is solved by $q_1^{k_1} = 5$ and $q_2^{k_2} = 2^2$. Then

$$[5, 4, 1^3] = [5, 3, 2^2] = [6, 2^2, 1^2] \tag{13}$$

a dreigänger! Note that another dreigänger occurs at the intersection given by the solution to $[4, q_1^{k_1}, 1^3] = [6, \mathbf{q_2^{k_2}}, 1^2]$. This is $q_1^{k_1} = 6$ and $\mathbf{q_2^{k_2}} = 4, 1$. The former first occurs at $N = 12$, while the latter occurs at $N = 13$. Apparently $N = 12$ is the smallest state size where a dreigänger occurs. As the previous analysis shows the DNA of these dreigängers grow with N as $IP[N - 12]$ and $IP[N - 13]$ respectively. Of course as N increases dreigängers will intersect with doppelgängers and other dreigängers to produce vielgänger states all of which propagate into higher dimensional states as appropriate IPs. Moreover, for even modest values of N the number of of vielgängers increases dramatically as readily seen for just the lowest seed (11).

These conclusions raise the question: how many states will have unique entropies for large N? To answer this recall the famous asymptotic formula given by Hardy and Ramanujan [3] for the number of partitions of the integer N:

$$IP[N] \sim 1/(4N\sqrt{3}) \exp\left(\pi\sqrt{2N/3}\right). \tag{14}$$

The fraction of independent entropy states of size N is

$$\mathcal{N} = N_*(IP[N])^{-1} \tag{15}$$

where

$$N_* = IP[N] - \sum_{j=7} \Gamma_j IP[N-j] = IP[N]\left(1 - \sum_{j=7} \Gamma_j \frac{IP[N-j]}{IP[N]}\right) \tag{16}$$

is the number of independent entropy states, j is a doppelgänger seed number, i.e. $j = 7, 8, 12, 13, \ldots$, and Γ_j is the number of vielgänger states for individual seed numbers. Using (14) it is straightforward to establish

$$\frac{IP[N-j]}{IP[N]} \approx \left(1-\frac{j}{N}\right)^{-1} \exp\left[\pi\sqrt{2N/3}\left(\sqrt{1-j/N}-1\right)\right]. \tag{17}$$

Substituting into (15) gives

$$\mathcal{N} \approx 1-\sum_{j=7}\Gamma_j\left(1-\frac{j}{N}\right)^{-1} \exp\left[\pi\sqrt{2N/3}\left(\sqrt{1-j/N}-1\right)\right]. \tag{18}$$

For $N \gg j$ one has $\exp\left[\pi\sqrt{2N/3}\left(\sqrt{1-j/N}-1\right)\right] \approx 1-\pi j(6N)^{-1/2}$. Thus the fraction of independent entropy states approximates

$$\mathcal{N} \approx \pi\,(6N)^{-1/2}\sum_{j=7}^{j_{max}\ll N} j\Gamma_j. \tag{19}$$

We are surprised that the number of unique entropy states asympotes as $N^{-1/2}$. Nevertheless for a finite N it is always possible to find unique entropy states.

Clearly low doppelänger seeds such as $N-7$ and $N-8$ have more impact on $\mathcal{N}$ than higher doppelänger seeds. However, veilgänger seeds, which don't occur until $N \geq 12$ have a multiplicative effect. Thus the cumulative effect on the fraction of independent entropy states for any N will be the agreegate of all doppelgänger and veilgängers states.

Conceptually the theory outlined above is simple. For a given N one can calculate the number of vielgänger states iteratively from lower values of N and from any new seeds that become available. But for even modest values of N the calculations quickly become exponentially tedious. In the next section we focus on numerical results for cases up to $N = 50$.

3 Numerical results

In this section we summarize numerical calculations for finite N up to $N = 50$. We used an extremely fast recursive C++ algorithm developed by Monsi Terdex to generate integer partitions [4]. This code generated the 204,226 partitions for $N = 50$ in under a second on a MacBook Pro.

The entropies for all partitions were obtained from (3), sorted and the vielgängers identified. Table 1 shows the total number of partitions that do not have at least one doppelgänger as a function of N. In Fig. 1 we plot the fraction of these unique entropy states versus $N^{-1/2}$. This figure supports (19).

Next we consider the frequencies of occurrence of the various vielgänger states. Figure 2 shows the frequencies of occurrence for the vielgängers for $N = 50$. The top panel suggests an approximate exponential decrease in vielgänger frequency as a function of degeneracy as expected from our analysis. The bottom panel depicts the same information on a semi-log scale. This shows some modest variability at large degeneracies from an exponential fit. We attribute this to a sampling bias since small degeneracies are much more frequent than large degeneracies.

Table 1 Non-degenerate entropy state counts

N	# of partitions	# of non-degenerate entropy states	
10	42	30	0.71429
20	627	214	0.34131
30	5604	950	0.16952
40	37,338	3362	0.09004
50	204,226	10,417	0.05101

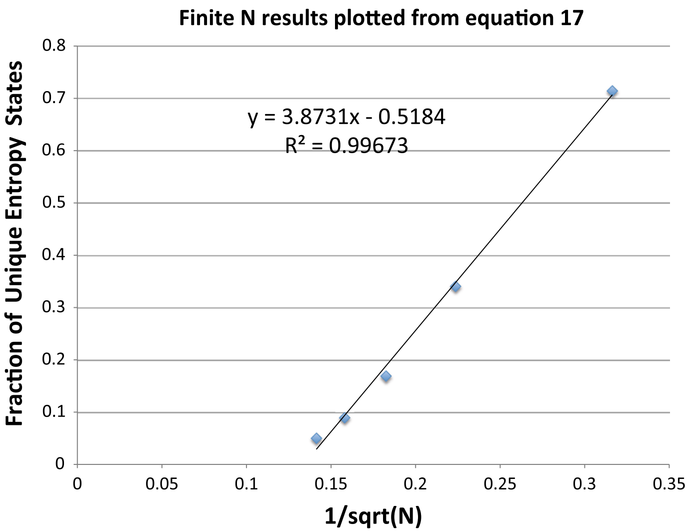

Fig. 1 Fraction of non-degenerate, unique entropy states, N/ IP[N]—*solid line* is fit to Eq. (17)

We next consider the maximum degeneracy values. The results for finite N up to 50 are listed in Table 2 and plotted in Fig. 3. In view of Fig. 2 we observe that the maximum degeneracy increases exponentially with $\sqrt{N}$, while its frequency of occurrence decreases exponentially with $-\sqrt{N}$.

4 Discussion

The analytic analysis in Sect. 2 and the numerical analysis in Sect. 3 demonstrated that the number of veilgängers increases as the $IP[N]$, which asymptotes as exp $\sqrt{N}$. Moreover, as shown by (19) the non degenerate states asymptote as $N^{-1/2}$. Thus for even a modest state size of $N = 50$ only 5 % are non degenerate. Surprisingly the numerical analysis showed a similar exp $\sqrt{N}$ increase in the maximum degeneracy d_{max}. For $N > 40$ d_{max} exceeds the state size! Simultaneously, for a given N, there is a near exponential decrease in the number of vielegängers as a function of their degeneracy.

All of these attributes are purely mathematical characteristics of the Boltzmann distribution. This distribution arose from Boltzmann's pioneering work on statistical

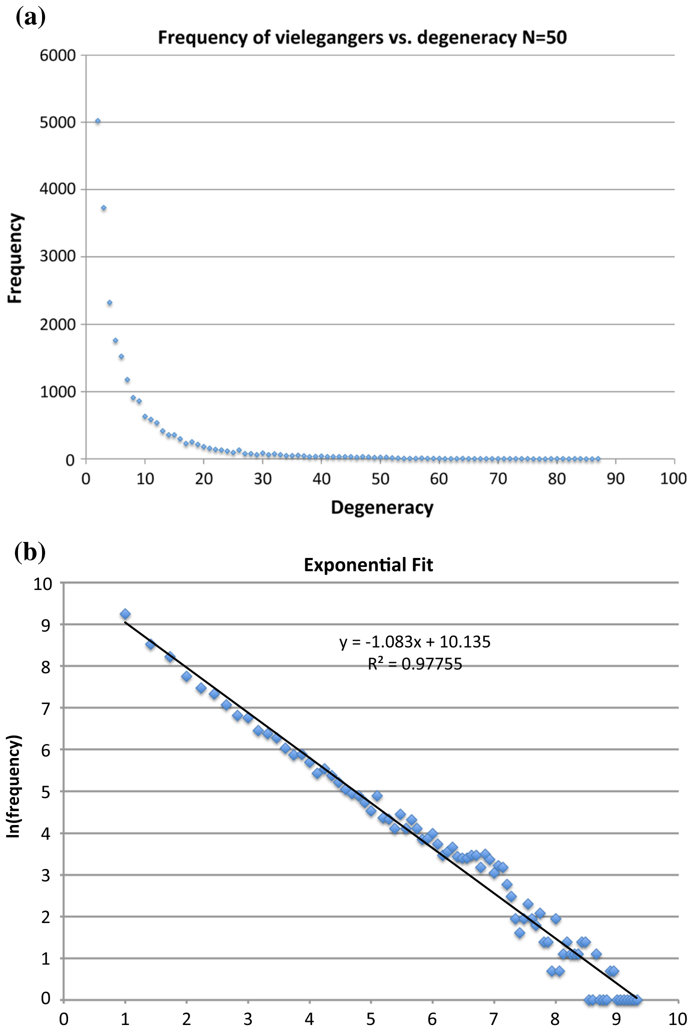

Fig. 2 **a** Frequency decrease with degeneracy, **b** Logarithmic plot showing frequency decreasing as $exp(-\sqrt{N})$

mechanics where N may be is at least of the order of Avogrado's number and phase space is partitioned by energy levels. We show here that for large N there is a vanishingly small percentage of Boltzmann states with unique entropies. The vast number of entropy states are degenerate with many degeneracies exceeding N. Energy and entropy are fundamental concepts in statistical mechanics and thermodynamics. We are unaware of any study that relates energy partitions to these mathematical entropy degeneracies.

In addition to statistical mechanics and thermodynamics, the Boltzmann entropy is relevant to a wide variety of areas that include communications theory, probability,

Table 2 Largest degeneracy, d_{max}

N	d_{max}
10	2
20	6
30	16
40	37
50	87

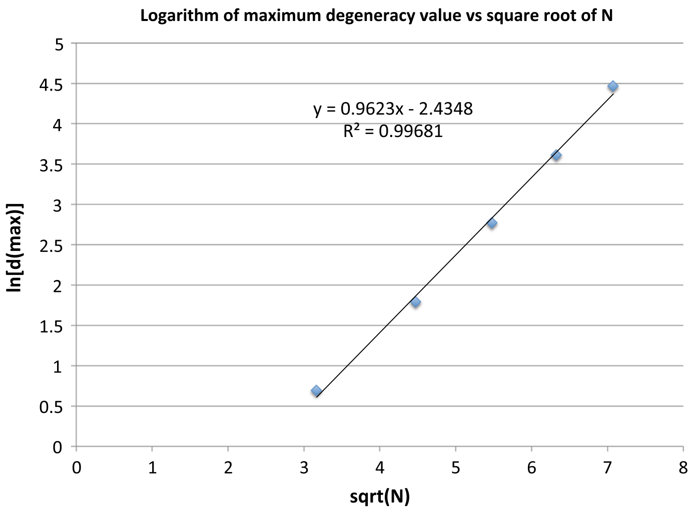

Fig. 3 Exponential growth of d_{max}

mixing, and even biology and ecology. Consequently it is presumptuous for us to gauge the impact these results may have on these disciplines. Instead we now comment on one of interest to us, namely mixing.

When a set of N objects is classified according to some principle, one obtains a subdivision in subsets without objects in common. Each of these subsets can be represented by an integer partition vector, which can be equivalently displayed as a diagram. These diagrams, sometimes called *Young Diagrams*, are in 1:1 correspondence with the Boltzmann macrostates. Ruch [5] proved that the appropriate measure for comparing mixedness of partitions is the majorization or dominance partial order. Thus, majorization describes the *fundamental* mixing character of Boltzmann states. The majorization partial order is sometimes represented as a lattice of partitions with the least mixed state [N] at the top and the most mixed state $[1^N]$ at the bottom, [6]. This lattice is known as the Young Diagram Lattice (YDL) or simply Young's lattice.

Previously [1] we pointed out that mixedness and entropy of macrostates are complementary concepts. The fact that mixing is only partially ordered by majorization necessarily implies that many Boltzmann entropy states are incomparably mixed. We argued that for $N \geq 6$, the extent of the incomparability of each state (i.e. the number of states to which it is not comparable by mixing) of the system provides a qualitative measure of the complexity of the system [1]. Seitz [7] previously had suggested

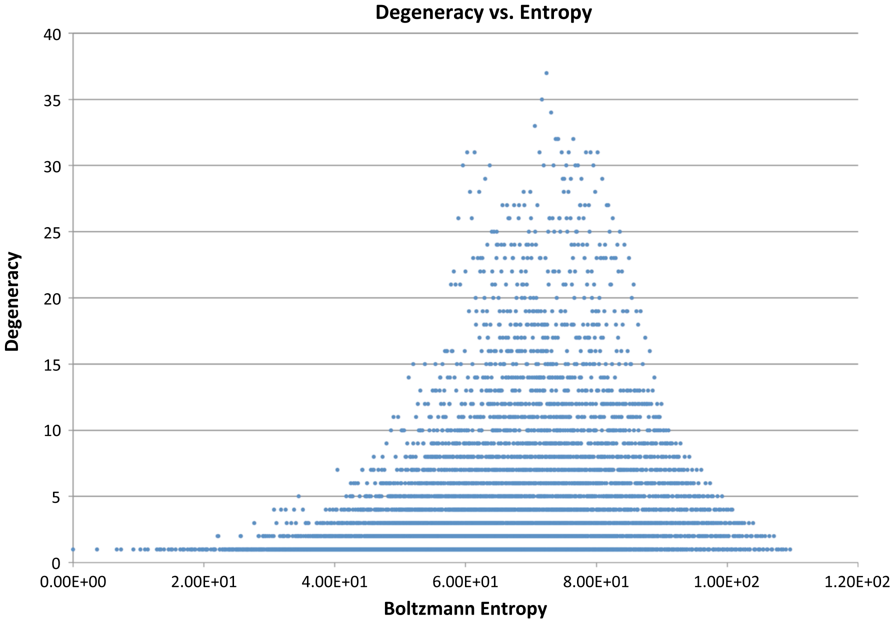

Fig. 4 Growth of d_{max}

that complexity of a state is related to the length of the maximal antichain containing that state (An anti-chain is defined mathematically as a set of *mutually incompatible* partitions [8].). Whether either maximal antichain length or incomparability number is plotted versus entropy one obtains a plot of complexity versus entropy in close agreement with the complexity arguments of Huberman and Hogg [9]. Thus, both incomparability number and maximal antichain length arguments show the complementarity of complexity and entropy. It is noted that maximum incomparability and maximal anti-chain lengths occur for Boltzmann states with entropies that are intermediate between total order and disorder.

Further, vielgängers obviously form an anti-chain in the Young Diagram Lattice. Finding the largest anti-chain in the lattice is a long-standing open mathematical question. Unfortunately, the maximum degeneracy vilegänger, while an anti-chain, is not maximal as is easily seen from $N = 15$ where the maximal antichain length is known to be 9 while 4 is the maximum degeneracy vielgänger. Nevertheless, we find that the largest degeneracy vilegängers occur at intermediate entropies. Early has shown [10] that the length of the maximal antichain for N is bounded by $N^{-5/2}e^{\pi\sqrt{2N/3}} \leq a_N \leq N^{-1}e^{\pi\sqrt{2N/3}}$. This is consistent with Fig. 3, which also showed that the logarithm of the maximum degeneracy value increases with $\sqrt{N}$. However the results here are limited to $N \leq 50$.

Regardless of whether incomparability numbers or maximal antichain lengths are employed, [1] incomparability qualitatively "measures" complexity. It is trivial to establish from the formulae developed in Sect. 2 that the members of every set of vielgängers are incomparable by mixing (as are the members of maximal antichains). Thus, the findings here suggest a deep relationship between degeneracy of Boltzmann entropy states and complexity since the degeneracy of an vielegänger is related to it's incomparability number. Figure 4 shows is a plot of degeneracy vs. entropy for $N = 40$. We note that this general form is similar to other curves that seek to qualitatively relate complexity to entropy, ([1,9,11]). This view of complexity regards completely ordered (low entropy) systems as "simple". Also high entropy (very mixed) systems are similarly simple since they as are essentially random. Consequently "complexity" occurs at intermediate entropies. The roots of complexity in nature are embedded in the the Boltzmann distribution.

Acknowledgments Discussions with Professors Kevin Mclhany and San X. Liang provided useful insight into degenerate entropies.

References

1. W. Seitz, A.D. Kirwan Jr., Entropy **16**, 3793 (2014)
2. D.A. McQuarrie, *Statistical Mechanics* (University Science Books, Sausalito, 2000)
3. G. Hardy, S. Ramanujan, Proc. Lond. Math. Soc. **17**, 75 (1918)
4. M. Terdex, Generating integer partitions in "c/c++" (2013). https://monsiterdex.wordpress.com/2013/04/12/integer-partition-in-cc/
5. E. Ruch, Theoret. Chim. Acta (Berl.) **38**, 167 (1975)
6. A. Young, Proc. Lond. Math. Soc. **33**, 97 (1900)
7. W.A. Seitz, in *Partial Orders, in Environmental Sciences and Chemistry*, ed. by P.D.L.C. Dr, R. Brüggemann (Springer, The Netherlands, 2006), pp. 367–382
8. W.T. Trotter, *Combinatorics and Partially Ordered Sets: Dimension Theory*. Johns Hopkins Studies in the Mathematical Sciences (Book 6) (Johns Hopkins University Press, Baltimore, 2001)
9. B. Huberman, T. Hogg, Phys. D **2**, 376 (1986)
10. E. Early, Discrete Math. **313**(20), 2168 (2013)
11. Scholarpedia. Complexity (2014). http://www.scholarpedia.org/article/Complexity