2-14-2022

Maximum Entropy and Constraints in Composite Systems

John D. Ramshaw
Portland State University, jdramshaw@yahoo.com

Citation Details
Ramshaw, J. D. (2022). Maximum entropy and constraints in composite systems. Physical Review E, 105(2), 024138.

Maximum entropy and constraints in composite systems

John D. Ramshaw
Department of Physics, Portland State University, Portland, Oregon 97207, USA

(Received 9 January 2022; accepted 14 February 2022; published 24 February 2022)

The principle of maximum entropy (PME), as expounded by Jaynes, is based on the maximization of the Boltzmann-Gibbs-Shannon (BGS) entropy subject to linear constraints. The resulting probability distributions are of canonical (exponential) form. However, the rationale for linear constraints is nebulous, and probability distributions are not always canonical. Here we show that the correct noncanonical distribution for a system in equilibrium with a finite heat bath is implied by the *unconstrained* maximization of the *total* BGS entropy of the system and bath together. This procedure is shown to be equivalent to maximizing the BGS entropy of the system alone subject to a contrived nonlinear constraint which reduces to (a) the usual linear constraint for an infinite heat bath, and (b) a previously enigmatic logarithmic constraint which implies a power-law distribution for a large but finite heat bath. This procedure eliminates the uncertainty as to the proper constraints, and easily generalizes to arbitrary composite systems, for which it provides a simpler alternative to the Jaynes PME.

DOI: 10.1103/PhysRevE.105.024138

The principle of maximum entropy (PME) [1–7] is a powerful and general method of statistical inference with numerous applications in a wide variety of disciplines, including information theory, probability theory, data analysis, linguistics, and most if not all of the physical and biological sciences. The versatility and continued vitality of the PME are amply attested by the published proceedings of the annual International Workshops on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, the 41st of which is scheduled to be held in Paris in July 2022. Those volumes often bear the abbreviated title *Maximum Entropy and Bayesian Methods*.

The standard version of the PME is based on maximizing the well-known Boltzmann-Gibbs-Shannon (BGS) entropy

$$S = -\sum_i p_i \log p_i, \tag{1}$$

where p_i is the probability that the system occupies state i. The maximization is performed subject to constraint conditions that represent known or hypothetical information about the system, typically expressed as mean or expectation values of the form $\bar{A} = \sum_i p_i A_i$. The resulting probability distribution p_i can then be used to predictively compute whatever statistical averages are of interest. The PME has a strong intuitive appeal, but there is no *a priori* assurance that its predictions will be accurate. However, inaccurate predictions do not necessarily imply that the PME itself has failed, as its detractors sometimes claim, but merely that the constraints imposed are inappropriate or insufficient and that different or additional constraints may be required [1]. Unfortunately, the form of those constraints is unlikely to be obvious, for if it were they would presumably have been imposed at the outset.

Thus the most conspicuous limitation of the PME is that the appropriate constraints to be imposed in any particular application are not always obvious. Our purpose here is to critically explore this issue within the context of statistical thermodynamics, to which the PME was first applied in two classic papers by Jaynes [8,9]. Those papers have been highly influential, and are widely regarded as providing a simple and elegant alternative derivation of the canonical and grand canonical probability distributions. For present purposes it suffices to restrict attention to the canonical distribution

$$p_i = \frac{\exp(-\beta E_i)}{Z_c(\beta)}, \tag{2}$$

where E_i is the energy of the system in state i and $Z_c(\beta) = \sum_j \exp(-\beta E_j)$ is the canonical partition function. As is well known, Eq. (2) describes a system in thermal equilibrium with an infinite heat bath at temperature $T = 1/\beta$ (in energy units). Jaynes showed that Eq. (2) is an immediate consequence of maximizing S subject to the constraint

$$\sum_j p_j E_j = \bar{E}, \tag{3}$$

where $\bar{E} = -(\partial/\partial\beta)\log Z_c(\beta)$. This constraint simply specifies the mean energy of the system, which intuitively seems very natural, especially since the thermodynamic state of the system is uniquely defined and determined by its mass, volume, and energy. On further reflection, however, a clear and compelling justification for Eq. (3) is not obvious [10,11]. In the absence of such a justification, the PME derivation of Eq. (2) can hardly be regarded as satisfactory, nor would it likely have been widely accepted as such if the validity of Eq. (2) had not previously been established on other grounds.

Equation (3), or at least its generality, is further called into question by the fact that the probability distribution p_i is no longer canonical when the heat bath is large but finite. In that case p_i assumes the more general form [12–14]

$$p_i = q_i \equiv \frac{W_b(E_0 - E_i)}{\sum_j W_b(E_0 - E_j)} = \frac{W_i}{W_0}, \tag{4}$$

where E_0 is the total energy of the system and bath together, $W_b(E)$ is the number of equally probable bath states with energies in the interval $(E, E + \Delta E)$, ΔE is a macroscopically negligible but finite energy tolerance, $W_i \equiv W_b(E_0 - E_i)$, and $W_0 \equiv \sum_j W_j$ is the total number of states in the system-bath composite system consistent with E_0. In order for the PME to remain valid in such systems, it would evidently be necessary to replace Eq. (3) by a different constraint, but neither the form of such a constraint nor its justification is apparent.

In the special case of ideal heat baths, $W_b(E)$ is typically proportional to E^α, where $\alpha \gg 1$ is of the order of the number of particles or degrees of freedom of the bath [12–14]. Equation (4) then reduces to [15]

$$p_i = \frac{(1 - \beta E_i/\alpha)^\alpha}{\sum_j (1 - \beta E_j/\alpha)^\alpha}, \tag{5}$$

where $\beta \equiv \alpha/E_0$. It was subsequently shown [16] that Eq. (5) is not restricted to ideal heat baths but is more generally valid as an asymptotic approximation to Eq. (4) for arbitrary $W_b(E)$. Equation (5) reduces to Eq. (2) in the limit $E_0 = \alpha/\beta \to \infty$, but for finite E_0 it is of power-law rather than exponential form. Power-law probability distributions are commonplace [17], but they do not seem to emerge from the PME in a natural way. They can, however, be obtained by maximizing S subject to logarithmic constraints (e.g., Refs. [16,18–24]), but such constraints seem artificial and often lack a clear physical interpretation or justification. More elaborate constraints to the same end have also been proposed (e.g., Ref. [25]).

A curious aspect of the Jaynes formulation of the PME, as applied to thermodynamic systems, is that it is based on maximizing the entropy S of the system of interest alone, exclusive of the entropy of the heat bath with which the system interacts and exchanges energy. In this respect it differs from the analogous thermodynamic extremum principle, which states that the *total* thermodynamic entropy of a composite system is a maximum at equilibrium [26]. Moreover, the statistical equilibrium between two macroscopic systems is similarly characterized by the maximum value of their *total* statistical entropy [7,27]. The system and heat bath together can be conceptually regarded as a single composite system, so the preceding observations suggest that it may be of interest to explore the consequences of maximizing the total entropy S_0 of the system and heat bath together, rather than the entropy S of the system alone as was done by Jaynes. Such an analysis was previously performed by Lee and Pressé [28,29], but with the limited objective of showing how the Jaynes PME and Eq. (3) emerge from the maximization of S_0. In what follows, the further implications of maximizing S_0 are explored in greater depth, detail, and generality.

To forestall confusion, a word about constraints is in order. The system-bath composite system is presumed to be isolated, so its energy E_0 is constant. This condition is intrinsic to isolated systems, so we do not regard or treat it as a constraint per se. In contrast, the normalization condition $\sum_i p_i = 1$ must be enforced as a constraint, even though it is merely a mathematical requirement that all probability distributions must satisfy by definition. It therefore has an entirely different character from physical constraints, so it will henceforth be understood that "unconstrained" is an abbreviation for "unconstrained except for normalization."

Of course, the system and heat bath are not statistically independent, so their separate BGS entropies are not simply additive, but their total entropy S_0 is nevertheless given by [30]

$$S_0 = S + \sum_i p_i S_i^b, \tag{6}$$

where S_i^b is the conditional BGS entropy of the heat bath when the system of interest is in state i. The energy of the heat bath is then simply $E_0 - E_i$, and the number of bath states with that energy is $W_b(E_0 - E_i)$. Those states are presumed equally probable, so S_i^b reduces to the Boltzmann-Planck entropy

$$S_i^b = \log W_b(E_0 - E_i) = \log W_i. \tag{7}$$

Combining Eqs. (1), (4), (6), and (7), we obtain

$$S_0 = \sum_i p_i \log \frac{q_i}{p_i} + \log W_0. \tag{8}$$

The well-known inequality $\log x \leqslant x - 1$ [5,7] implies that

$$\sum_i p_i \log \frac{q_i}{p_i} \leqslant 0, \tag{9}$$

which combines with Eq. (8) to imply that $S_0 \leqslant \log W_0$. Equation (8) further shows that the probability distribution p_i for which S_0 attains its maximum value $\log W_0$ is simply $p_i = q_i$, which is precisely the known correct probability distribution given by Eq. (4). Thus the *unconstrained* maximization of the total entropy S_0 implies the correct p_i for a system in equilibrium with a finite heat bath with an arbitrary density of states. The maximization of S_0 rather than S is much simpler than the Jaynes PME, as it eliminates the need to impose an associated constraint condition, such as Eq. (3) or some uncertain or unknown generalization thereof. Indeed, that maximization has now already been performed in advance for an arbitrary $W_b(E)$, with the general result $p_i = q_i$, so the need to repeat it for different choices of $W_b(E)$ has also been eliminated.

Nevertheless, the question naturally arises as to whether a general constraint exists and can be determined such that the constrained maximization of S alone produces the same result as the unconstrained maximization of S_0. To this end, we observe that if the mean bath entropy $\sum_i p_i S_i^b$ in Eq. (6) were held constant, the maximization of S_0 would reduce to the maximization of S. Maximizing S subject to the constraint

$$\sum_i p_i S_i^b = \sum_i p_i \log W_i = \sigma \tag{10}$$

should therefore be equivalent to the unconstrained maximization of S_0, provided the value of the constant σ is properly chosen. This constraint was previously inferred by Pressé *et al.* [29] via a different argument. Its validity is easily confirmed by the method of Lagrange multipliers, in which the maximization of S subject to the constraints of Eq. (10) and normalization is accomplished by maximizing the auxiliary quantity

$$\hat{S} \equiv -\sum_j p_j \log p_j + \lambda \sum_j p_j + \mu \sum_j p_j \log W_j, \quad (11)$$

where the values of the Lagrange multipliers λ and μ are implicitly determined by the constraints. Setting $\partial\hat{S}/\partial p_i = 0$ and imposing the normalization constraint, we obtain, after a little algebra,

$$p_i = \frac{W_i^{\mu}}{Z(\mu)}, \quad (12)$$

where

$$Z(\mu) \equiv \sum_j W_j^{\mu} \quad (13)$$

is a generalized partition function. It then follows from Eqs. (10), (12), and (13) that

$$\frac{\partial}{\partial\mu} \log Z(\mu) = \sigma, \quad (14)$$

which determines the functional relation between σ and μ. Comparison of Eqs. (12) and (13) with Eq. (4) shows that in order for p_i to reduce to the correct probability distribution q_i, the value of σ must be chosen so that $\mu = 1$. When $\mu \neq 1$, however, it is noteworthy that Eqs. (4) and (12) combine to imply that $p_i = q_i(\mu)$, where

$$q_i(\mu) \equiv \frac{q_i^{\mu}}{\sum_j q_j^{\mu}} \quad (15)$$

is just the escort distribution [31] of order μ associated with the correct probability distribution q_i. Maximization of the BGS entropy S subject to the constraint of Eq. (10) therefore constitutes a variational principle for generating those escort distributions.

It is straightforward to verify that the generalized constraint of Eq. (10) properly reduces to the linear and logarithmic constraints which are known to imply Eqs. (2) and (5), respectively. Inspection of Eqs. (2) and (4) shows that the approximation to W_i which implies Eq. (2) is $W_i \cong A\exp(-\beta E_i)$, where A is a constant. Thus $\log W_i \cong \log A - \beta E_i$, whereupon Eq. (10) reduces to a constraint on $\sum_i p_i E_i$. Similarly, the approximation to W_i which transforms Eq. (4) into Eq. (5) is $W_i \cong A(E_0 - E_i)^{\alpha}$. Thus $\log W_i \cong \log A + \alpha \log(E_0 - E_i)$, whereupon Eq. (10) reduces to a constraint on $\sum_i p_i \log(E_0 - E_i)$ as previously observed [16]. Both those constraints are therefore explained and uniquely implied by Eq. (10), and no longer appear arbitrary or inscrutable. Note that in these two special cases, violating the condition $\mu = 1$ would simply rescale the value of β by a factor of μ, while preserving the mathematical form of Eqs. (2) and (5). Of course, $1/\beta$ would then no longer represent the temperature of the heat bath.

The preceding analysis shows that the correct noncanonical probability distribution $p_i = q_i$ for a thermodynamic system in equilibrium with a finite heat bath can indeed be obtained by the constrained maximization of the BGS entropy S of the system alone. This result confirms the validity of such a procedure, but it should not be misinterpreted as an endorsement thereof, because the required constraint of Eq. (10) seems artificial and unnatural, and obviously lacks the simplicity and intuitive appeal of Eq. (3). Nor in retrospect should such simplicity have been expected, since Eq. (10) is in essence a mere mathematical artifice designed to reproduce the correct result by maximizing only a portion of the total entropy. The essential point is that in spite of its validity, there is no need or reason to actually make use of Eq. (10) [or Eq. (3) for that matter], as the unconstrained maximization of the total entropy S_0 is evidently simpler, more general, and more fundamental.

In the present context, the conventional linear constraint of Eq. (3), and the more general constraint of Eq. (10), essentially represent the influence of the heat bath on the system of interest. Thus it is hardly surprising that such constraints are no longer needed when the system and heat bath together are regarded, and consistently treated, as a single composite system. The total BGS entropy of that composite system then automatically accounts for the interactions between its constituent subsystems. However, these features are not specific to thermodynamic systems. This observation suggests that the preceding development may provide a useful paradigm for a more general treatment of composite systems in which constraints are eliminated by maximizing the total BGS entropy. It is hoped that the resulting description may sometimes be simpler than the use of complicated constraints such as Eq. (10) or those of Ref. [25]. We therefore proceed to formulate such an approach in more general terms.

Consider a composite system consisting of two interacting subsystems A and B, the states of which are labeled by indices i and j, respectively. Let p_{ij} denote the joint probability that system A occupies state i and system B simultaneously occupies state j. The marginal probability that system A occupies state i regardless of the state j of system B is simply $p_i = \sum_j p_{ij}$, and the conditional probability that system B occupies state j when system A is known or constrained to occupy state i is $p(j|i) \equiv p_{ij}/p_i$. The total BGS entropy of the composite system AB is therefore given by

$$S_{AB} = -\sum_{ij} p_{ij} \log p_{ij} = S_A + \sum_i p_i S_i^B, \quad (16)$$

where $S_A \equiv -\sum_i p_i \log p_i$, and

$$S_i^B \equiv -\sum_j p(j|i) \log p(j|i) \quad (17)$$

is the conditional entropy of system B when system A is in state i [30]. Let

$$q_i \equiv \frac{1}{Z_B} \exp S_i^B, \quad (18)$$

where $Z_B \equiv \sum_i \exp S_i^B$. Equations (16) and (18) combine to imply

$$S_{AB} = \sum_i p_i \log \frac{q_i}{p_i} + \log Z_B. \quad (19)$$

As before, the inequality $\log x \leqslant x - 1$ then implies that $S_{AB} \leqslant \log Z_B$, so that S_{AB} attains its unconstrained maximum value $\log Z_B$ when

$$p_i = q_i = \frac{1}{Z_B} \exp S_i^B. \quad (20)$$

The unconstrained maximization of S_{AB} may be regarded as an alternative PME for composite systems. In contrast to the Jaynes PME, however, that maximization need not be repeated in each particular case of interest, because it has already been performed in general, with Eq. (20) as the general result. In particular, Eq. (4) is merely the special case of Eq. (20) that results when $S_i^B = \log W_b(E_0 - E_i)$. Equation (20) bears a strong family resemblance and relationship to the well-known Einstein fluctuation formulas [7,12,32,33], which are not restricted to thermodynamic fluctuations [7]. Once again, the same p_i could evidently be obtained by maximizing S_A alone subject to a contrived constraint on the quantity $\sum_i p_i S_i^B$, but such a procedure seems ad hoc and pointless. Of course, Eq. (20) is useful only if S_i^B is either already known or can be plausibly approximated or postulated. However, the peculiarity of Eq. (10) suggests that devising a useful approximation to S_i^B may well be simpler, easier, or more intuitive than attempting to devise an appropriate constraint on the maximization of S_A alone. On the other hand, if the external agents or interactions that influence system A are hidden, unknown, or so poorly understood that one has no rational basis for determining or postulating S_i^B, then one can simply revert back to maximizing S_A alone subject to postulated or conjectured constraints. Nevertheless, the essential point remains that in its most fundamental form the PME strictly applies only to isolated systems, so the proper entropy to maximize is S_{AB} rather than S_A.

The preceding development easily generalizes to composite systems of three or more subsystems. Consider a ternary composite system ABC consisting of a subsystem A of primary interest and two additional subsystems B and C. The states of these subsystems are respectively labeled by i, j, and k; the joint probability that subsystems A, B, and C simultaneously occupy states i, j, and k is p_{ijk}; and the marginal probability distribution of subsystem A is $p_i = \sum_{jk} p_{ijk}$. The total BGS entropy of the composite system ABC is simply

$$S_{ABC} = -\sum_{ijk} p_{ijk} \log p_{ijk}. \tag{21}$$

But subsystems B and C can be conceptually regarded as a single composite subsystem BC with states labeled by jk, to which the preceding development can be immediately applied by means of the replacements $B \to BC$ and $j \to jk$. The unconstrained maximization of S_{ABC} therefore yields

$$p_i = \frac{1}{Z_{BC}} \exp S_i^{BC}, \tag{22}$$

where $Z_{BC} \equiv \sum_i \exp S_i^{BC}$,

$$S_i^{BC} \equiv -\sum_{jk} p(jk|i) \log p(jk|i) \tag{23}$$

and $p(jk|i) = p_{ijk}/p_i$.

However, the composite system ABC can equally well be regarded as the result of combining a binary composite system AB with a third subsystem C. The marginal probability that subsystem AB occupies state ij is then $p_{ij} = \sum_k p_{ijk}$, and the conditional probability that subsystem C occupies state k when subsystem AB occupies state ij is $p(k|ij) = p_{ijk}/p_{ij}$. Thus

$$S_{ABC} = S_{AB} + \sum_{ij} p_{ij} S_{ij}^C, \tag{24}$$

where $S_{AB} \equiv -\sum_{ij} p_{ij} \log p_{ij}$ and

$$S_{ij}^C \equiv -\sum_k p(k|ij) \log p(k|ij). \tag{25}$$

Of course, S_{AB} is no longer the total entropy, but Eqs. (16)–(19) nevertheless remain valid for an arbitrary p_{ij}, regardless of its origin. We may therefore combine Eqs. (19) and (24) to obtain

$$S_{ABC} = \sum_i p_i \log \frac{q_i}{p_i} + \sum_{ij} p_{ij} S_{ij}^C + \log Z_B, \tag{26}$$

where q_i is still given by Eqs. (17) and (18). Equation (26) is algebraically equivalent to Eq. (21), so its unconstrained maximization must still imply Eq. (22). However, inspection of Eq. (26) shows that the same p_i would be obtained by maximizing the *relative* entropy $\sum_i p_i \log(q_i/p_i)$ alone subject to a constraint on the quantity $\sum_i p_i Q_i$, where $Q_i \equiv \sum_j p(j|i) S_{ij}^C$. This observation provides an interesting new interpretation of the "method of maximum relative entropy" [34] as it applies to composite systems.

The inability of the Jaynes PME with linear constraints to produce noncanonical probability distributions has been interpreted by some authors as justification for defining generalized entropies. For example, it is well known that the power-law distribution of Eq. (5) can be derived by maximizing the Rényi or Tsallis entropy subject to the linear constraint of Eq. (3). However, the BGS entropy S is clearly more fundamental than the constraints, so it seems incongruous to modify S rather than the constraints. Moreover, the BGS form of S is uniquely determined by certain intuitively desirable properties and consistency conditions [5,30,31,35], at least one of which must be sacrificed if S is replaced by any of its proposed generalizations. Fortunately, the present analysis confirms the many previous observations (e.g., Refs. [16,18–24]) that no such replacement is necessary to obtain power-law or other noncanonical probability distributions. For this purpose it suffices to retain the BGS entropy S and either (a) replace Eq. (3) by an appropriate nonlinear constraint, or preferably (b) maximize S_0 or S_{AB} rather than S or S_A. The limitations of the Jaynes PME in its usual form are therefore simply artifacts of the linear constraints, and do not imply any deficiency in the BGS entropy itself.

Generalized entropies have become quite fashionable in recent decades [36,37] to the point where their proliferation has been likened to an infestation [23]. One is reminded of the old adage that a man with one watch knows what time it is. It has often been suggested that such generalizations are needed to describe self-similarity and scaling laws, which are ubiquitous in nonlinear and complex systems [31,38,39]. The popular Rényi and Tsallis entropies [40,41] are simple functions of $Q(\mu) \equiv \sum_j p_j^{\mu}$, and the analysis of scaling behavior is indeed facilitated by defining the escort distributions [31] $p_i(\mu) \equiv p_i^{\mu}/Q(\mu)$. However, the utility of escort distributions does not imply a deficiency in the BGS entropy or a need to introduce alternative entropies. The BGS entropy of the escort distribution of order μ is simply

$$S(\mu) = -\sum_i p_i(\mu)\log p_i(\mu) \tag{27}$$

and one easily verifies that

$$\begin{aligned} S(\mu) &= -\mu\,\frac{\partial}{\partial\mu}\log Q(\mu) + \log Q(\mu) \\ &= -\mu^2\,\frac{\partial}{\partial\mu}\left[\frac{1}{\mu}\log Q(\mu)\right]. \end{aligned} \tag{28}$$

Thus $S(\mu)$ is determined by $Q(\mu)$, and vice versa, so it evidently contains and represents precisely the same information as the Rényi and Tsallis entropies.

I am grateful to the anonymous referees for thoughtful and constructive comments and suggestions which resulted in significant improvements.

[1] *The Maximum Entropy Formalism*, edited by R. D. Levine and M. Tribus (MIT Press, Cambridge, MA, 1979).

[2] *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz (Reidel, Dordrecht, 1983).

[3] S. Guiasu and A. Shenitzer, Math. Intell. **7**, 42 (1985).

[4] *Maximum Entropy in Action*, edited by B. Buck and V. A. Macaulay (Oxford University Press, Oxford, 1991).

[5] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, UK, 2003).

[6] J. N. Kapur, *Maximum-Entropy Models in Science and Engineering* (Wiley, New York, 1993; New Age, New Delhi, 2009).

[7] J. D. Ramshaw, *The Statistical Foundations of Entropy* (World Scientific, Singapore, 2018).

[8] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

[9] E. T. Jaynes, Phys. Rev. **108**, 171 (1957).

[10] J. Uffink, Stud. Hist. Philos. Mod. Phys. **27**, 47 (1996).

[11] I. J. Ford, Phys. Rev. E **92**, 052142 (2015).

[12] R. Becker, *Theory of Heat*, 2nd ed. (Springer-Verlag, New York, 1967).

[13] R. Kubo, *Statistical Mechanics: An Advanced Course with Problems and Solutions* (North-Holland, Amsterdam, 1971).

[14] R. P. Feynman, *Statistical Mechanics: A Set of Lectures* (Benjamin/Cummings, Reading, MA, 1972).

[15] A. R. Plastino and A. Plastino, Phys. Lett. A **193**, 140 (1994).

[16] J. D. Ramshaw, Phys. Rev. E **98**, 020103(R) (2018).

[17] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).

[18] E. W. Montroll and M. F. Shlesinger, J. Stat. Phys. **32**, 209 (1983).

[19] J. T. Bendler, J. J. Fontanella, and M. F. Shlesinger, Physica D **193**, 67 (2004).

[20] D. H. Zanette and M. A. Montemurro, Phys. Lett. A **324**, 383 (2004).

[21] V. Nieves, J. Wang, R. L. Bras, and E. Wood., Phys. Rev. Lett. **105**, 118701 (2010).

[22] J. Peterson, P. D. Dixit, and K. A. Dill, Proc. Natl. Acad. Sci. USA **110**, 20380 (2013).

[23] M. Visser, New J. Phys. **15**, 043021 (2013).

[24] Á. Corral and M. García del Muro, Entropy **22**, 179 (2020).

[25] J.-F. Bercher, Phys. Lett. A **372**, 5657 (2008).

[26] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).

[27] K. Huang, *Statistical Mechanics*, 2nd ed. (Wiley, New York, 1987).

[28] J. Lee and S. Pressé, Phys. Rev. E **86**, 041126 (2012).

[29] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, Rev. Mod. Phys. **85**, 1115 (2013).

[30] A. I. Khinchin, *Mathematical Foundations of Information Theory* (Dover, New York, 1957).

[31] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems: An Introduction* (Cambridge University Press, Cambridge, UK, 1993).

[32] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed., Part 1 (Pergamon, Oxford, 1980).

[33] Y. Mishin, Ann. Phys. (NY) **363**, 48 (2015).

[34] A. Caticha, Entropy **23**, 895 (2021).

[35] J. E. Shore and R. W. Johnson, IEEE Trans. Inf. Theory **26**, 26 (1980).

[36] J. M. Amigo, S. G. Balogh, and S. Hernández, Entropy **20**, 813 (2018).

[37] S. G. Balogh, G. Palla, P. Pollner, and D. Czégel, Sci. Rep. **10**, 15516 (2020).

[38] M. Schroeder, *Fractals, Chaos, Power Laws* (Freeman, New York, 1991).

[39] G. I. Barenblatt, *Scaling* (Cambridge University Press, Cambridge, UK, 2003).

[40] A. Rényi, *Probability Theory* (North-Holland, Amsterdam, 1970; *Probability Theory* Dover, New York, 2007).

[41] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).