EURANDOM PREPRINT SERIES

2015-001

December 30, 2014

Breaking of ensemble equivalence in networks

Tiziano Squartini, Joey de Mol, Frank den Hollander, Diego Garlaschelli ISSN 1389-2355

Breaking of ensemble equivalence in networks

Tiziano Squartini,^{1,2} Joey de Mol,^{1,3} Frank den Hollander,³ and Diego Garlaschelli¹

Lorentz Institute for Theoretical Physics, University of Leiden (The Netherlands)

²Institute for Complex Systems, University of Rome "La Sapienza" (Italy)

³Mathematical Institute, University of Leiden (The Netherlands)

(Dated: December 30, 2014)

It is generally believed that, for physical systems in the thermodynamic limit, the microcanonical description as a function of energy coincides with the canonical description as a function of temperature. However, various examples have been identified for which the microcanonical and canonical ensembles are not equivalent. A complete theory of this intriguing phenomenon is still missing. Here we show that ensemble nonequivalence can manifest itself also in discrete enumeration problems. We show that, for discrete systems, ensemble equivalence reduces to equivalence of the large deviation properties of microcanonical and canonical probabilities of a single microstate. As specific examples, we consider ensembles of graphs with topological constraints. We find that, while graphs with a given number of links are ensemble-equivalent, graphs with a given degree sequence (including random regular graphs, sparse scale-free networks, and core-periphery networks) are not. This mathematical result provides a theoretical explanation for various 'anomalies' that have recently been observed in networks, namely, the non-vanishing of canonical fluctuations in the configuration model and of the difference between microcanonical and canonical entropies of random regular graphs. While it is generally believed that ensemble nonequivalence is associated with long-range interactions, our findings show that it may naturally arise in systems with local constraints as well.

PACS numbers: 05.20.Gg,02.10.Ox,89.75.Hc

Background. In statistical physics, calculating the equilibrium properties of a system with a given energy requires averaging over the so-called microcanonical ensemble [1, 2], i.e., the uniform distribution on the set of all particle configurations having a prescribed energy. Apart from trivial examples, this is a mathematically challenging task. Moreover, it is difficult to physically realize a situation where there is no uncertainty in the energy of the system. Therefore, it is often preferable to work with the so-called *canonical ensemble* [2], i.e., a probability distribution with maximal entropy on the set of all particle configurations that does allow for the 'wrong' energy, but in such a way that the average energy matches a prescribed value. This is achieved through the selection of an appropriate temperature, mathematically arising as the Lagrange multiplier enforcing the average energy.

Starting with the work of Gibbs [2], the microcanonical and canonical ensembles have been shown to be equivalent in the thermodynamic limit (i.e., when the number of particles in the system tends to infinity) for physical systems with short-range interactions. The original argument is that in the canonical ensemble at fixed temperature the energy fluctuations are negligible with respect to the average energy, so that in the thermodynamic limit the canonical ensemble is effectively microcanonical with a sharp value of the energy. Today, most textbooks in statistical physics still convey the message that the equivalence of ensembles holds universally for every physical system, justifying the use of energy and temperature as two different parameters giving an equivalent description.

However, in the past 15 years various studies have highlighted that ensemble equivalence breaks down for cer-

tain many-body systems encountered in models of fluid turbulence [3, 4], quantum phase separation [5–7], star formation [8, 9], nuclear fragmentation [10], and spin interaction [11]. Physically, it is believed that one of the main causes of ensemble nonequivalence is the presence of long-range interactions. However, a complete theoretical understanding of this intriguing problem is still missing.

Mathematically, ensemble non-equivalence has been approached in various ways [12, 13]. In particular, the microcanonical and canonical ensembles are said to be thermodynamically equivalent [7] when the entropy and the free energy of the system are one-to-one related via a Legendre transform. The ensembles are said to be macrostate equivalent [12] when the sets of equilibrium values of the macrostate (energy, magnetization, etc.) are the same. Finally, a recent and mathematically appealing definition is that of measure equivalence [13], according to which the ensembles are said to be equivalent when the canonical probability distribution converges to the microcanonical probability distribution in the thermodynamic limit. Under certain hypotheses, the three definitions have been shown to be equivalent [13]. Moreover, large deviation theory [14] shows that the ensembles are nonequivalent on all three levels when the microcanonical entropy function is nonconcave as a function of the energy density in the thermodynamic limit [13]. This provides important new insight, because for a long time physicists believed that entropy is always strictly concave.

Here we study ensemble nonequivalence for discrete enumeration problems, in particular, for *networks* with topological constraints [15–17]. Usually, ensemble nonequivalence is studied for systems in which the

Boltzmann distribution describes a certain physical interaction that is encapsulated in the energy (including e.g. spin interactions on random graphs [11]). ever, as shown by Javnes [18], the Boltzmann distribution describes much more general ensembles of systems with given constraints, namely, all solutions to the information-theoretic maximum-entropy problem of inference from partial information (where the available information plays the role of the constraint). Building on this fact, we argue that, for any discrete enumeration problem where we need to count microcanonical configurations compatible with a given constraint, there exists a 'dual' problem involving canonical configurations induced by the same constraint. In this article, we prove a general result relating measure equivalence to the equivalence of the large deviation properties of microcanonical and canonical probabilities, and provide examples of networks with constraints that exhibit nonequivalence.

Relative entropy and large deviations. Using the measure-based definition of ensemble equivalence [13], we say that the microcanonical and canonical ensembles are equivalent if and only if their specific relative entropy is zero. Since we are considering discrete systems, we need the following discrete version of the relative entropy:

$$S_N(P_{\text{mic}}||P_{\text{can}}) \equiv \sum_{\mathbf{G} \in \mathcal{G}_N} P_{\text{mic}}(\mathbf{G}) \ln \frac{P_{\text{mic}}(\mathbf{G})}{P_{\text{can}}(\mathbf{G})},$$
 (1)

where \mathcal{G}_N denotes the set of all configurations (in our later examples, graphs) with N particles (nodes), and P_{mic} , P_{can} denote the microcanonical and canonical probability distributions respectively. Following [13], we say that the ensembles are measure equivalent if and only if

$$s \equiv \lim_{N \to \infty} \frac{S_N(P_{\text{mic}}||P_{\text{can}})}{N} = 0.$$
 (2)

Before considering specific cases, we first prove an important general result. Any discrete enumeration problem naturally induces a microcanonical distribution

$$P_{\text{mic}}(\mathbf{G}) = \begin{cases} 1/\Omega_{\vec{C}^*}, & \text{if } \vec{C}(\mathbf{G}) = \vec{C}^*, \\ 0, & \text{else,} \end{cases}$$
 (3)

where $\Omega_{\vec{C}^*} \equiv |\{\mathbf{G} \in \mathcal{G}_N \colon \vec{C}(\mathbf{G}) = \vec{C}^*\}|$ is the number of configurations that realize a given hard constraint \vec{C}^* (which we take to be a vector because it may concern several topological constraints simultaneously). Following Jaynes [18], we introduce a 'dual' enumeration problem involving a canonical probability distribution $P_{\text{can}}(\mathbf{G})$ defined as the solution of the maximization of the Shannon entropy $S \equiv -\sum_{\mathbf{G} \in \mathcal{G}_N} P_{\text{can}}(\mathbf{G}) \ln P_{\text{can}}(\mathbf{G})$ subject to the soft constraint $\langle \vec{C} \rangle = \vec{C}^*$, where $\langle \cdot \rangle$ denotes the average w.r.t. P_{can} , and to the normalization condition $\sum_{\mathbf{G} \in \mathcal{G}_N} P_{\text{can}}(\mathbf{G}) = 1$ [15]. This gives

$$P_{\text{can}}(\mathbf{G}) = \frac{e^{-H(\mathbf{G}, \vec{\theta}^*)}}{Z(\vec{\theta}^*)},\tag{4}$$

where $H(\mathbf{G}, \vec{\theta}) \equiv \vec{\theta} \cdot \vec{C}(\mathbf{G})$ is the *Hamiltonian* and $Z(\vec{\theta}) \equiv \sum_{\mathbf{G} \in \mathcal{G}_N} e^{-H(\mathbf{G}, \vec{\theta})}$ is the *partition function*. Note that in eq.(4) the parameter $\vec{\theta}$ must be set to the particular value $\vec{\theta}^*$ that realizes $\langle \vec{C} \rangle = \vec{C}^*$ [17]. This value also maximizes the likelihood of the data [19].

We now make a simple but crucial observation. Noting from the form of $H(\mathbf{G}, \vec{\theta})$ that $P_{\mathrm{can}}(\mathbf{G}_1) = P_{\mathrm{can}}(\mathbf{G}_2)$ when $\vec{C}(\mathbf{G}_1) = \vec{C}(\mathbf{G}_2)$ (i.e., the canonical probability is the same for all configurations having the same value of the constraint), we may rewrite eq.(1) as

$$S_N(P_{\text{mic}}||P_{\text{can}}) = \ln \frac{P_{\text{mic}}(\mathbf{G}^*)}{P_{\text{can}}(\mathbf{G}^*)},\tag{5}$$

where \mathbf{G}^* is any configuration in \mathcal{G}_N such that $\vec{C}(\mathbf{G}^*) = \vec{C}^*$. The condition for equivalence in eq.(2) then becomes

$$\lim_{N \to \infty} \frac{1}{N} \ln P_{\text{mic}}(\mathbf{G}^*) = \lim_{N \to \infty} \frac{1}{N} \ln P_{\text{can}}(\mathbf{G}^*), \quad (6)$$

which shows that the breaking of ensemble equivalence coincides with $P_{\rm mic}(\mathbf{G}^*)$ and $P_{\rm can}(\mathbf{G}^*)$ having different large deviation properties [14]. Importantly, this local condition involves the microcanonical and canonical probabilities of a single configuration \mathbf{G}^* realizing the hard constraint. Apart from its theoretical importance, this result greatly simplifies mathematical calculations. Examples. To illustrate the above concepts, we consider specific classes of random graphs. We begin with an example of equivalence, namely, graphs with a fixed number of links, i.e., $\vec{C} \equiv L$. Writing $L = \lambda M$, where $M \equiv N(N-1)/2$ is the number of pairs of nodes and λ is the fraction of realized links, we have

$$\Omega_{L^*} = \binom{M}{L^*} = \binom{M}{\lambda^* M}, \qquad 0 < \lambda^* < 1.$$
 (7)

The canonical ensemble can be obtained from eq.(4) by setting $H(\mathbf{G}, \theta) = \theta L(\mathbf{G})$ and $p^* \equiv \frac{e^{-\theta^*}}{1+e^{-\theta^*}} = \lambda^*$ [17]. This produces the *Erdős-Rényi random graph* where each pair of nodes is connected with probability p^* :

$$P_{\text{can}}(\mathbf{G}) = (p^*)^{L(\mathbf{G})} (1 - p^*)^{M - L(\mathbf{G})}.$$
 (8)

We can now compute the relative entropy from eq.(5) as

$$S(P_{\text{mic}}||P_{\text{can}}) = -\lambda^* M \ln \lambda^* - (1 - \lambda^*) M \ln(1 - \lambda^*) + -\ln \binom{M}{\lambda^* M} = \ln \sqrt{2\pi \lambda^* (1 - \lambda^*) M} + O(1/M),$$
(9)

where we have used Stirling's formula $n! = (n/e)^n \sqrt{2\pi n}$ $[1 + O(1/n)], n \to \infty$ [20]. This gives

$$s = \lim_{N \to \infty} \frac{\ln \sqrt{2\pi\lambda^*(1-\lambda^*)M}}{N} = 0, \tag{10}$$

proving ensemble equivalence in this case.

We next consider three examples of networks with a given degree sequence, i.e., $\vec{C} \equiv \vec{k} = (k_1, \dots, k_N)$. This coincides with the *configuration model*, where every node has a given degree [15]. The microcanonical number $\Omega_{\vec{k}}$ is not known in general, but asymptotic results exist.

• We first treat the 'sparse case' where

$$k_{\max} \equiv \max_{1 \le i \le N} k_i = o(\sqrt{N}). \tag{11}$$

In this regime it is known that [21, 22]

$$\Omega_{\vec{k}^*} = \frac{\sqrt{2} \left(\frac{2L^*}{e}\right)^{L^*}}{\prod_{i=1}^{N} k_i^*!} e^{-f(\vec{k}^*) + \frac{1}{4} + o(N^{-1}\overline{k^*}^3)}, \qquad (12)$$

where $\overline{k} = N^{-1} \sum_{i=1}^{N} k_i$ (average degree), $L = N\overline{k}/2$ (number of links), and $f(\vec{k}) \equiv (\overline{k^2}/2\overline{k})^2$ with $\overline{k^2} = N^{-1} \sum_{i=1}^{N} k_i^2$. The canonical ensemble is described by eq.(4) where $H(\mathbf{G}, \vec{\theta}) = \vec{\theta} \cdot \vec{k}(\mathbf{G})$ and $\vec{\theta}^*$ is the solution of

$$\sum_{i \neq i} \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}} = k_i^* \quad \forall i$$
 (13)

as proved in [17]. Setting $p_{ij}^* \equiv e^{-\theta_i^*-\theta_j^*}/(1+e^{-\theta_i^*-\theta_j^*}),$

$$P_{\text{can}}(\mathbf{G}) = \prod_{i=1}^{N} \prod_{j < i} (p_{ij}^*)^{g_{ij}} (1 - p_{ij}^*)^{1 - g_{ij}}, \qquad (14)$$

where g_{ij} is the (i, j)-entry of the adjacency matrix of the graph **G**. Eq.(11) ensures that $k_{\text{max}} = o(\sqrt{L})$, a condition under which eq.(13) is solved as [17]

$$p_{ij}^* \sim e^{-\theta_i^* - \theta_j^*} = \frac{k_i^* k_j^*}{2L^*} = o(1),$$
 (15)

where \sim means that the quotient tends to 1. This implies $\theta_i^* \sim -\ln(k_i^*/\sqrt{2L^*})$ and $\ln(1-p_{ij}^*) \sim -k_i^*k_j^*/2L^*$. Thus

$$\ln P_{\text{can}}(\mathbf{G}^*) = \sum_{i=1}^{N} \sum_{j < i} [g_{ij}^* \ln p_{ij}^* + (1 - g_{ij}^*) \ln(1 - p_{ij}^*)]$$

$$= \sum_{i=1}^{N} \sum_{j < i} [-g_{ij}^* (\theta_i^* + \theta_j^*) + \ln(1 - p_{ij}^*)]$$

$$= -\sum_{i=1}^{N} k_i^* \theta_i^* + \sum_{i=1}^{N} \sum_{j < i} \ln(1 - p_{ij}^*)$$

$$\sim \sum_{i=1}^{N} k_i^* \ln k_i^* - L^* \ln(2L^*) - L^*. \tag{16}$$

Combining eqs.(5), (12) and (16), we obtain

$$S(P_{\text{mic}}||P_{\text{can}}) \sim -L^* \ln(2L^*) + L^* + \sum_{i=1}^{N} \ln(k_i^*!)$$

$$+ f(\vec{k}^*) - \frac{1}{4} + o(N^{-1}\overline{k^*}^3)$$

$$- \sum_{i=1}^{N} k_i^* \ln k_i^* + L^* \ln(2L^*) + L^*$$
 (17)

which leads to

$$S(P_{\text{mic}}||P_{\text{can}}) \sim \sum_{i=1}^{N} \left[\ln(k_i^*!) - k_i^* \ln k_i^* + k_i^* \right] + f(\vec{k}^*) - \frac{1}{4} + o(N^{-1} \overline{k^*}^3).$$
 (18)

Abbreviating $\xi \equiv \lim_{N\to\infty} f(\vec{k}^*)/N$, and using bars to denote limiting averages over nodes, we arrive at

$$s = \overline{\ln[k^*!/(k^*/e)^{k^*}]} + \xi \ge \overline{\ln\sqrt{2\pi k^*}} + \xi, \tag{19}$$

where the inequality holds for every $k^* \ge 1$ [20]. Since $\xi \ge 0$, we have s > 0, proving ensemble nonequivalence.

We consider two examples in the sparse regime. The first case is that of sparse regular networks, where every node has the same degree, i.e., $k_i \equiv k^*$ with $k^* = o(\sqrt{N})$. Then $\xi = 0$ and $\overline{\ln k^*} = \ln k^*$, and so eq.(19) becomes

$$s \ge \ln \sqrt{2\pi k^*}, \qquad k^* = o(\sqrt{N}). \tag{20}$$

The nonequivalence for regular graphs can be traced back to the fact that the canonical ensembles for the two examples considered so far, namely, $\vec{C} = L$ and $\vec{C} = \vec{k}$, coincide via the identification $p^* = L^*/M = k^*/(N-1)$. Since the microcanonical ensembles are different, at most one of the two examples can show ensemble equivalence. As we proved ensemble equivalence for $\vec{C} = \vec{k}$ as well. Incidentally, note that if k^* grows with N, then s diverges like $\ln k^*$, which signals an extreme violation of equivalence.

The second case is the class of sparse scale-free networks [23], defined by a truncated power-law degree distribution of the form $F_N(k) = N^{-1} \sum_{i=1}^N 1_{\{k_i = k\}} = A_{k_c} k^{-\gamma}$ for $1 \le k < k_c$ and $F_N(k) = 0$ for $k \ge k_c$, with $2 < \gamma < 3$ (a typical range for real-world networks) and $1/k_c = o(1)$. The cut-off $k_c = o(\sqrt{N})$ [23] ensures eq.(11), so that eq.(15) is valid. Approximating $F_N(k)$ by a continuous distribution, we see that the normalization of F_N implies that $A_{k_c} \approx (\gamma - 1)/(1 - k_c^{-(\gamma - 1)})$. Computing $\overline{k^a} \approx A_{k_c} \int_1^{k_c} k^a k^{-\gamma} dk$, a = 1, 2, we get

$$f(\vec{k}^*) \approx \left(\frac{\gamma - 2}{2(3 - \gamma)} \frac{k_c^{3 - \gamma} - 1}{1 - k_c^{-(\gamma - 2)}}\right)^2 = o(N^{3 - \gamma}), \quad (21)$$

which leads to $\xi = 0$. Similarly, computing $\overline{\ln k} \approx A_{k_c} \int_1^{k_c} (\ln k) k^{-\gamma} dk$, we get

$$\overline{\ln k^*} \approx \frac{1}{\gamma - 1} - \frac{\ln k_c}{k_c^{\gamma - 1} - 1}, \qquad k_c = o(\sqrt{N})$$
 (22)

and so eq.(19) becomes

$$s \ge \overline{\ln \sqrt{2\pi k^*}} \approx \frac{1}{2(\gamma - 1)} + \ln \sqrt{2\pi}.$$
 (23)

This shows that sparse scale-free networks display non-equivalence. Also note that as γ decreases, the heterogeneity of the degree distribution increases and the violation of non-equivalence gets more severe.

• We next treat a class of extremely heterogeneous networks, for which eq.(11) is violated and therefore eq.(12) cannot be used. Specifically, we consider h-star graphs with a fully connected core of h hubs and a periphery of N-h leaves, each with a single connection to one of the hubs. The periphery is divided into h sets of equal size (N-h)/h, such that every leaf in a set is connected to the same hub. The microcanonical number of such graphs is the multinomial coefficient

$$\Omega_{\vec{k}^*} = \binom{N-h}{\frac{N-h}{h} \dots \frac{N-h}{h}} = \frac{(N-h)!}{\left(\frac{N-h}{h}!\right)^h}.$$
 (24)

Applying Stirling's approximation [20], we get

$$P_{\text{mic}}(\mathbf{G}^*) = h^{h-N} \left[2\pi (N-h) \right]^{\frac{h-1}{2}} h^{-\frac{h}{2}} [1 + o(1)].$$
 (25)

The canonical distribution is still described by eqs.(13) and (14). Since the degree of a periphery node is $k_P = 1$ and that of a core node is $k_C = h - 1 + \frac{N-h}{h}$, eq.(13) reduces to only two independent equations

$$h p_{CP}^* + (N - h - 1) p_{PP}^* = 1,$$
 (26)

$$(h-1) p_{CC}^* + (N-h) p_{CP}^* = (h-1) + \frac{N-h}{h}, (27)$$

where p_{CC}^* , p_{PP}^* and p_{CP}^* denote the connection probability (evaluated at $\vec{\theta} = \vec{\theta}^*$) between two core nodes, two periphery nodes, and a core node and a periphery node, respectively. The implicit (for $\vec{\theta}^*$) solution is

$$p_{CC}^* = 1, p_{PP}^* = 0, p_{CP}^* = h^{-1}.$$
 (28)

Therefore the canonical distribution is

$$P_{\text{can}}(\mathbf{G}^*) = \left(\frac{1}{h}\right)^{N-h} \left(1 - \frac{1}{h}\right)^{(h-1)(N-h)},$$
 (29)

the relative entropy is

$$S(P_{\text{mic}}||P_{\text{can}}) = (N - h)(h - 1)\ln\left(\frac{h}{h - 1}\right)$$
(30)
+ $(h - 1)\ln\sqrt{2\pi(N - h)} - h\ln\sqrt{h}$,

and the specific relative entropy is

$$s = (h-1)\ln\left(\frac{h}{h-1}\right). \tag{31}$$

So we again find nonequivalence for all h > 1. (For h = 1, both ensembles admit only one star-like configuration and are therefore equivalent.)

Conclusion. Our results show that, while it is generally believed that ensemble nonequivalence is associated with long-range interactions, it may naturally arise in systems with local (topological) constraints as well, and more in general in discrete enumeration problems. Our findings

contribute to the theoretical understanding of nonequivalence in discrete systems, by linking it to the nonequivalence of the large deviation properties of microcanonical and canonical probabilities of a single microstate. Our proof of the breakdown of ensemble equivalence in graphs with given degree sequence provides a theoretical explanation for some otherwise anomalous recent observations, namely the fact that the canonical and microcanonical entropies of random regular graphs are different even in the thermodynamic limit [16] and the non-vanishing of canonical fluctuations in the configuration model [24].

FdH is supported through the NWO Gravitation Grant NETWORKS. DG is supported through the EU project MULTIPLEX (contract 317532) and the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands) with funds from beneficiaries of Duyfken Trading Knowledge BV, Amsterdam, the Netherlands.

- [1] L. Boltzmann, Wiener Berichte 2, 373-435 (1877).
- [2] J.W. Gibbs, Elementary Principles of Statistical Mechanics, Yale University Press, New Haven, Connecticut (1902).
- [3] R.S. Ellis, K. Haven, and B. Turkington, J. Stat. Phys. 101, 999 (2000).
- [4] R.S. Ellis, K. Haven, and B. Turkington, Nonlinearity 15, 239 (2002).
- [5] M. Blume, V.J. Emery, and R.B. Griffiths, *Phys. Rev. A* 4, 1071 (1971).
- [6] J. Barré, D. Mukamel, and S. Ruffo, Phys. Rev. Lett. 87, 030601 (2001).
- [7] R.S. Ellis, H. Touchette, and B. Turkington, *Physica A* 335, 518 (2004).
- [8] D. Lynden-Bell, *Physica A* **263**, 293 (1999).
- [9] P.-H. Chavanis, Astron. & Astrophys. **401**, 15 (2003).
- [10] M. D'Agostino, Phys. Lett. B 473, 219 (2000).
- [11] J. Barré and B. Goncalves, Physica A 386, 212-218 (2007).
- [12] H. Touchette, R.S. Ellis, and B. Turkington, *Physica A* 340, 138-146 (2004).
- [13] H. Touchette, arXiv:1403.6608.
- [14] F. den Hollander, Large Deviations, Fields Institute Monographs 14, American Mathematical Society, Providence RI (2000).
- [15] J. Park and M.E.J. Newman, Phys. Rev. E 70, 066117 (2004).
- [16] K. Anand and G. Bianconi, Phys. Rev. E 80, 045102 (2009).
- [17] T. Squartini and D. Garlaschelli, New Journal of Physics 13, 083001 (2011).
- [18] E.T. Jaynes, Phys. Rev. 106, 4 (1957).
- [19] D. Garlaschelli and M.I. Loffredo, Phys. Rev. E 78, 015101(R) (2008).
- [20] H. Robbins, Amer. Math. Monthly 62, 26-29 (1955).
- [21] E.A. Bender, Discrete Math. 10, 217-223 (1974).
- [22] B.D. McKay and N.C. Wormald, Combinatorica 11, 369-382 (1991).
- [23] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B 38, 205-9 (2004).
- [24] T. Squartini, R. Mastrandrea, and D. Garlaschelli, arXiv:1406.1197 (2014).