Breaking of ensemble equivalence for complex networks
Roccaverde, A.

Citation
Roccaverde, A. (2018, December 5). *Breaking of ensemble equivalence for complex networks*. Retrieved from https://hdl.handle.net/1887/67095

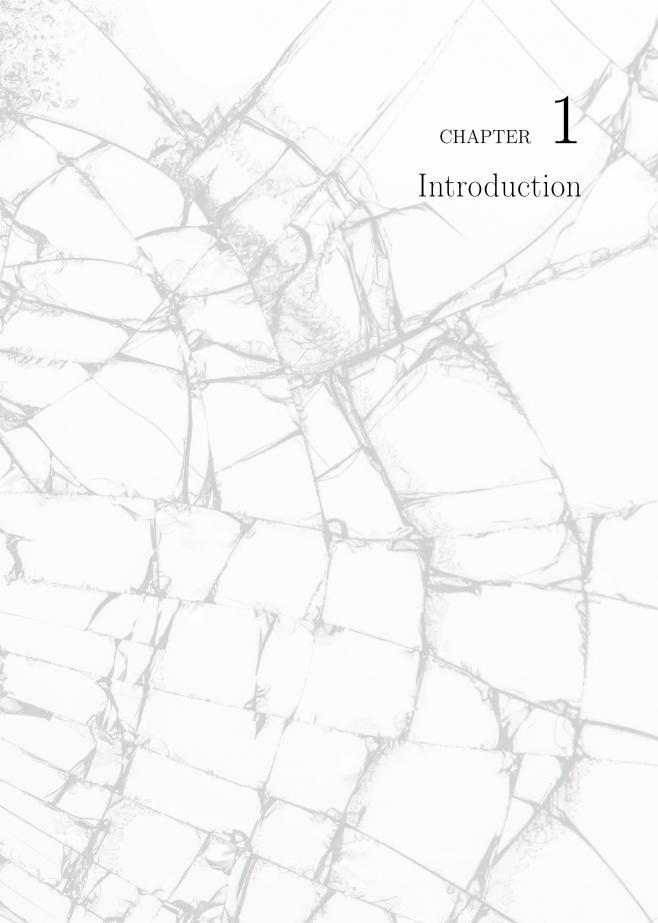Cover Page

Universiteit Leiden

Leiden University
Repository

The following handle holds various files of this Leiden University dissertation:
http://hdl.handle.net/1887/67095

Author: Roccaverde, A.
Title: Breaking of ensemble equivalence for complex networks
Issue Date: 2018-12-05

Introduction

§1.1 Gibbs Ensembles

In order to provide an introduction to Gibbs Ensembles we borrow from Garlaschelli, den Hollander, Roccaverde [49].

Statistical physics aims at describing collective behavior in systems consisting of a very large number of interacting particles (= atoms or molecules). This is a daunting task: a glass of water or a piece of iron can easily contain 10^{23} particles. Still, the hope is that the macroscopic properties of these particles as a whole can be explained from the microscopic interactions between them. For instance, we want to explain *why* water turns into ice (or vapor) at an appropriate temperature and *how* this transition exactly takes place. Similarly, we want to explain why a piece of iron at low temperature becomes magnetized when it is moved close to a magnet and remains magnetized after it is moved away from the magnet. We also want to understand why this does not happen when the temperature is high. For most physical systems the trajectories of the particles are so chaotic that they cannot be captured by explicit formulas. A full description would require knowledge of the positions and the speeds of all the particles at all times, which clearly is hopeless. Yet, we need large numbers of particles to explain collective phenomena: a single water molecule cannot transit from water to ice (or vapor).

The way out of this dilemma is offered by statistical physics: for most purposes a full microscopic description is not necessary: it suffices to have a macroscopic description in terms of a small number of relevant quantities, such as pressure, density and temperature. In statistical physics, the system is assumed to be a random sample, drawn from a set of allowed microscopic configurations that are consistent with a set of given macroscopic constraints. These constraints determine the pressure, density and temperature of the system. Collective phenomena, such as whether the system is a solid (ice), a liquid (water) or a gas (vapor), should follow from a combination of the microscopic interactions and the macroscopic constraints.

A physical system is rarely isolated. Typically, it is part of a larger system, that in turn is part of an even larger system, etc. For example, the molecules in a glass of water depend on what is outside the glass. The molecules at the top of the glass interact with the air above it. This air also contains water molecules, and a lively exchange takes place close to the surface of the water. We are thus tempted to believe that, in order to understand what happens inside the glass of water, we need to model all the molecules around the glass as well, and perhaps even all the molecules in the room the glass finds itself in. Fortunately, this is not the case, since the water molecules can only interact over short distances.

Statistical physics deals with the definition of the appropriate probability distribution over the set of allowed microscopic configuration, such as the locations and the speeds of the particles and how they bounce off each other and the wall of the container. These distributions need to take the macroscopic constraints into account. For instance, when the temperature is high the particles move quickly, which should be reflected in the choice of the probability distribution for the positions and the speeds. These probability distributions are called, in statistical physics, *ensembles*. They were introduced for the first time by Boltzmann [19] and then reformulated

in their modern probabilistic form by Gibbs [54]. Each ensemble describes how the system interacts with its surroundings and therefore represents a particular physical situation.

1 The *microcanonical ensemble*, where hard constraints are placed on both the energy and the number of particles: both are set to fixed values and are not allowed to vary.

2 The *canonical ensemble*, where a soft constraint is placed on the energy of the particles (in the sense that it may vary but with a fixed average), while a hard constraint is placed on the number of particles.

3 The *grandcanonical ensemble*, where both the energy and the number of particles are soft.

For systems of finite size, the three ensembles lead to different behavior. Therefore, in practical situations, the choice of ensemble is important and must be based on the physical situation that is described. In particular, an experimental physicist would use the microcanonical ensemble to model an isolated system (= a system that exchanges neither heat nor particles with its surroundings), the canonical ensemble to model a closed system (= a system that exchanges heat with an "external reservoir", with which it is in thermal equilibrium, but no particles), and the grandcanonical ensemble to model an open system (= a system that exchanges both heat and particles with the external reservoir, with which it is in thermal and chemical equilibrium). Choosing the wrong ensemble amounts to choosing the wrong microscopic probability distribution on which the computation of macroscopic quantities is based. For instance, if the experimental physicist is certain that the system under study does not exchange particles with its surroundings, then the grandcanonical ensemble is clearly not the right choice, and it would make the microscopic description of the system more noisy than is necessary.

§1.2 Equivalence of Ensembles

Statistical physics also deals with the problem of determining whether the ensembles give the same predictions when the system is very large. Traditionally, in physics books the three ensembles are assumed to be *thermodynamically equivalent*: for large systems fluctuations of macroscopic quantities around their average value are expected to be small and to be asymptotically vanishing as the number of particles tends to infinity. In the latter limit, called the thermodynamic limit, the soft constraints effectively become hard constraints. The assumption of ensemble equivalence dates back to Gibbs [54] and has been verified for traditional models of physical systems with short-range microscopic interactions and subject to a small number of macroscopic constraints. However, ensemble equivalence is not a simple concept. It has to be defined and studied carefully: depending on the level of description considered it can take different forms. We will talk about this in more detail in Section 1.3.

The general idea is that ensemble equivalence is convenient because it allows us to choose any of the three ensembles to work with. Soft constraints often are computationally easier to work with than hard constraints, which makes the choice of the canonical ensemble and the grandcanonical ensemble more convenient than that of the microcanonical ensemble. If ensemble equivalence holds and the system is large enough, then all three ensembles lead to the same macroscopic outcome for most of the relevant quantities. However, ensemble equivalence does not hold in general. This fact is important because, in such a situation, an experimental physicist must make a careful choice what ensemble to use for modeling the system, even when the system is large. A wrong choice means a wrong answer to macroscopic questions. Despite the fact that many textbooks still convey the message that ensemble equivalence holds for all physical systems, over the last decades various examples of physical systems have been found for which it breaks down ([85], [86], [87] and [29]).

Thus, breaking of ensemble equivalence means that different choices of ensemble lead to asymptotically different behavior. Consequently, while for applications based on ensemble equivalent models the choice of the working ensemble can be arbitrary and can be based on mathematical convenience, for those based on nonequivalent models the choice must be dictated by a criterion indicating which ensemble is the appropriate one to use. This criterion must be based on the *a priori* knowledge that is available about the network, i.e., which form of constraint (hard or soft) applies in practice.

§1.3 Definition of Ensemble Equivalence

In this section we give a brief introduction to the problem of ensemble equivalence. In his treatise [54], Gibbs argued that, in the so-called thermodynamic limit (when the number of particles goes to infinity), the microcanonical and the canonical ensemble become equivalent. Gibbs's argument was that, when the system is large, the fluctuations of the energy, in the canonical ensemble, become negligible with respect to the total energy. The canonical ensemble therefore essentially chooses a unique value of the energy, equal to the energy used to define the microcanonical ensemble. In this sense the use of the canonical ensemble, instead of the more complicated microcanonical ensemble, is justified and the ensembles are said to be equivalent. In other words, the equilibrium properties of the system can be described by using either the energy or the temperature as parameters. This equivalence can be proved in simple cases, for example, an ideal gas or non-interacting systems.

Many other complex systems have been studied with the help of statistical ensembles, which brought physicists to assume that the two ensembles are always equivalent ([89, 59, 6, 68, 91]). However more recently systems have been found where breaking of ensemble equivalence occurs. These examples include models of fluid turbulence [41], star formation [72] and networks [7, 85, 95].

The problem of ensemble (non)equivalence has been formulated in a more rigorous manner. Ellis, Haven and Turkington [40] studied two types of (non)equivalence. The first type is equivalence at the *thermodynamic level*, that has been studied the most so far. The second type is equivalence at the *macrostate level*, introduced in [40].

Another type of equivalence is the equivalence the *measure level*. Following [97], we present the problem of ensemble equivalence in these three different forms:

- *Thermodynamic equivalence*: the microcanonical and the canonical ensemble are said to be thermodynamically equivalent when the entropy (as a function of the energy) and the free energy (as a function of the temperature) are one-to-one related by a Legendre transform.

- *Macrostate equivalence*: the microcanonical and the canonical ensemble are said to be macrostate equivalent when the equilibrium values of the macrostate predicted by the microcanonical ensemble and the equilibrium values of the macrostate predicted by the canonical ensemble are the same.

- *Measure equivalence*: the microcanonical and the canonical ensemble are said to be measure equivalent when the Gibbs distribution defining the canonical ensemble at the microstate level converges to the distribution defined by Boltzmann's equiprobability postulate defining the microcanonical ensemble.

Touchette [97] proves that thermodynamic nonequivalence occurs whenever the microcanonical entropy function has one or more points of non-concavity and that macrostate and the thermodynamic equivalence are essentially equivalent. Measure equivalence is also proved to be equivalent to the other two types of equivalence. The main conclusion is that the three 'different' levels of ensemble equivalence are equivalent, whenever the setting is a general particle system, under the assumption that thermodynamic functions and equilibrium macrostates exists and are defined through large deviation principles (see [97] for more details).

The physical reason behind ensemble nonequivalence still remains to be clarified and part of this thesis is to understand some of the hidden mechanisms behind it. The type of equivalence considered throughout the thesis is that at the measure level. In the following we give the precise definition of the latter.

§1.3.1 Measure equivalence

Equivalence at the measure level concerns convergence of the canonical ensemble to the microcanonical ensemble at the microscopic level. In this section we give the mathematical definition of this type of equivalence and explain the idea behind it.

Relative entropy

P and Q are two discrete probability measures defined on the same space $\mathcal{X}$, with P absolutely continuous with respect to Q ($P \ll Q$). The *relative entropy* of P with respect to Q is defined as

$$S(P|Q) = \sum_{i \in \mathcal{X}} P(i) \ln \frac{P(i)}{Q(i)}. \tag{1.1}$$

The relative entropy $S(P|Q)$ is not a distance (it is not symmetric and does not satisfy the triangle inequality). However, $S(P|Q)$ is non-negative and equals zero if and only

if $P = Q$ almost everywhere. Moreover, Pinsker's inequality shows that $S(P|Q)$ is an upper bound on the total variation distance, namely,

$$d_{TV}(P, Q) = \frac{1}{2} \sum_{i \in \mathcal{X}} |P(i) - Q(i)| \le \sqrt{S(P|Q)}. \tag{1.2}$$

In the case of the microcanonical and the canonical ensemble, defined for an N-particle system (i.e., see equations (1.6) and (1.9)), we get $P_{\text{mic}}^N \ll P_{\text{can}}^N$ (but not vice versa). Therefore, the relative entropy of the microcanonical ensemble with respect to the canonical ensemble can be computed and takes the form

$$S(P_{\text{mic}}^N | P_{\text{can}}^N) = \sum_{i \in \mathcal{X}} P_{\text{mic}}^N(i) \ln \frac{P_{\text{mic}}^N(i)}{P_{\text{can}}^N(i)}. \tag{1.3}$$

The *specific relative entropy* is defined as the limit

$$s_\infty = \lim_{N \to \infty} \frac{1}{N} S(P_{\text{mic}}^N | P_{\text{can}}^N). \tag{1.4}$$

1.3.1 Definition (Measure equivalence). The microcanonical and the canonical ensemble are said to be equivalent at the measure level if

$$s_\infty = 0.$$

The immediate implication of ensemble equivalence (at the measure level) combined with Pinsker's inequality in (1.2) is that the total variation $d_{TV}(P_{\text{mic}}^N, P_{\text{can}}^N)$ grows slower than $\sqrt{N}$ as $N \to \infty$.

§1.4 Statistical Ensembles for Complex Networks

Ensemble (non)equivalence is usually studied for systems in which the Boltzmann distribution describes a certain physical interaction that is encapsulated in the energy. However, as already shown by Jaynes [61], the Boltzmann distribution describes much more general ensembles of systems with given constraints, namely, *all solutions to the maximum-entropy problem of inference from partial information*. In what follows 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. We define microcanonical and canonical ensembles for *complex networks*, and provide examples of networks that exhibit equivalence and nonequivalence of the ensembles at the measure level, introduced in Definition 1.3.1. The statistical mechanics approach turns out to be very powerful in the study of real-world networks, for which a detailed knowledge of the architecture is typically not available. The way proposed here is to study the complex networks through a probabilistic description, i.e., statistical ensembles. To that end the network is assumed to be a random sample drawn from a set of allowed configurations that are consistent with a set of known *topological constraints* [95]. In the following we give a rigorous definition of these statistical ensembles for complex networks.

§1.4.1 Microcanonical and Canonical Ensemble for Complex Networks

In Section 1.1 we explained why statistical physics deals with the definition of the appropriate probability distribution and what are the possible effects this has on the experiments. Here we consider two of the key choices of probability distribution, namely:

(1) The *microcanonical ensemble*, where the constraints are *hard* (i.e., are satisfied by each individual configuration).

(2) The *canonical ensemble*, where the constraints are *soft* (i.e., hold as ensemble averages, while individual configurations may violate the constraints).

(In both ensembles, the entropy is *maximal* subject to the given constraints.) We start by giving the rigorous definitions of the microcanonical and the canonical ensemble for complex networks.

For $n \in \mathbb{N}$, let $\mathcal{G}_n$ denote the set of all simple undirected graphs with n nodes. Any graph $G \in \mathcal{G}_n$ can be represented as an $n \times n$ matrix with elements

$$g_{ij}(G) = \begin{cases} 1 & \text{if there is a link between node } i \text{ and node } j, \\ 0 & \text{otherwise.} \end{cases} \tag{1.5}$$

Let $\vec{C}$ denote a vector-valued function on $\mathcal{G}_n$. Given is a specific value $\vec{C}^*$, which we assume to be *graphical*, i.e., realisable by at least one graph in $\mathcal{G}_n$.
The *microcanonical probability distribution* on $\mathcal{G}_n$ with *hard constraint* $\vec{C}^*$ is defined as

$$P_{\text{mic}}(G) = \begin{cases} \Omega_{\vec{C}^*}^{-1}, & \text{if } \vec{C}(G) = \vec{C}^*, \\ 0, & \text{else,} \end{cases} \tag{1.6}$$

where

$$\Omega_{\vec{C}^*} = |\{G \in \mathcal{G}_n \colon \vec{C}(G) = \vec{C}^*\}| \tag{1.7}$$

is the number of graphs that realise $\vec{C}^*$. The *canonical probability distribution* $P_{\text{can}}(G)$ on $\mathcal{G}_n$ is defined as the solution of the maximisation of the *entropy*

$$S_n(P_{\text{can}}) = - \sum_{G \in \mathcal{G}_n} P_{\text{can}}(G) \ln P_{\text{can}}(G) \tag{1.8}$$

subject to the normalisation condition $\sum_{G \in \mathcal{G}_n} P_{\text{can}}(G) = 1$ and to the *soft constraint* $\langle \vec{C} \rangle = \vec{C}^*$, where $\langle \cdot \rangle$ denotes the average w.r.t. P_{can}. This gives the formula (see [61])

$$P_{\text{can}}(G) = \frac{\exp[-H(G, \vec{\theta}^*)]}{Z(\vec{\theta}^*)}, \tag{1.9}$$

where

$$H(G, \vec{\theta}) = \vec{\theta} \cdot \vec{C}(G) \tag{1.10}$$

is the *Hamiltonian* and

$$Z(\vec{\theta}) = \sum_{G \in \mathcal{G}_n} \exp[-H(G, \vec{\theta})] \tag{1.11}$$

is the *partition function*. In (1.9) the parameter $\vec{\theta}$ must be set equal to the particular value $\vec{\theta}^*$ that realises $\langle \vec{C} \rangle = \vec{C}^*$. This value is unique and maximises the likelihood of the model given the data (see [51]). We next proceed with the equivalence definition of the ensembles.

§1.4.2 α_n-Equivalence of Ensembles

In order to define the equivalence at the measure level for complex networks, we follow [92, 48, 50] and define the *relative entropy* of P_{mic} w.r.t. P_{can} as

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = \sum_{G \in \mathcal{G}_n} P_{\mathrm{mic}}(G) \log \frac{P_{\mathrm{mic}}(G)}{P_{\mathrm{can}}(G)}, \tag{1.12}$$

and the α_n-*relative entropy* as [50]

$$s_{\alpha_n} = {\alpha_n}^{-1} S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}), \tag{1.13}$$

where α_n is a *scale parameter*. The limit of the relative entropy α_n-density is defined as

$$s_{\alpha_\infty} \equiv \lim_{n \to \infty} s_{\alpha_n} = \lim_{n \to \infty} {\alpha_n}^{-1} S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) \in [0, \infty]. \tag{1.14}$$

We say that the microcanonical and the canonical ensemble are equivalent on scale α_n if and only if

$$s_{\alpha_\infty} = 0. \tag{1.15}$$

This is a generalization of the standard measure equivalence definition given in Section 1.3.1. In fact, for complex networks, a specific parameter corresponding to the number of particles (the volume of the system) does not exist. For example, we could decide to use the number of nodes n as the parameter representing the volume of the system. On the other hand, we could also use the number of edges $\binom{n}{2}$, or even other n dependent factors instead. It becomes clear that computing the *specific relative entropy*, i.e., dividing by the number n of particles, does not have a precise meaning in the context of complex networks. This is the main reason why we choose to be general and use a parameter α_n. The choice of the scale α_n at which we check for (non)equivalence in this thesis is flexible and depends on the number of nodes n, on the constraint at hand and on its value as well. Indeed, we consider different choices of α_n for different models. In certain cases, we in fact prefer to reverse the point of view and look for the 'natural' or 'critical' scale α_n at which s_{α_∞} is positive and finite. This second approach allows us to immediately conclude that the ensembles are β_n-equivalent for all $\beta_n = \omega(\alpha_n)$ and nonequivalent when $\beta_n = \Omega(\alpha_n)$. For instance, if the constraint is on the *degree sequence*, then in the *sparse regime* the critical scale turns out to be $\alpha_n = n$ [92], [48] (in which case s_{α_∞} is the specific relative entropy 'per vertex'), while

in the *dense regime* it turns out to be $\alpha_n = n \log n$ [50]. For more details, see Section 1.5 for the sparse regime and Section 1.6 for the dense regime. On the other hand, if the constraint is on the *total numbers of edges and triangles*, with values different from what is typical for the Erdős-Renyi random graph in the dense regime, then the critical scale turns out to be $\alpha_n = n^2$ [38] (in which case s_{α_∞} is the specific relative entropy 'per edge'). This is discussed in more detail in Section 1.8.

Before considering specific cases, we recall an important observation made in [92]. The definition of $H(G, \vec{\theta})$ ensures that, for any $G_1, G_2 \in \mathcal{G}_n$, $P_{\mathrm{can}}(G_1) = P_{\mathrm{can}}(G_2)$ whenever $\vec{C}(G_1) = \vec{C}(G_2)$ (i.e., the canonical probability is the same for all graphs having the same value of the constraint). We may therefore rewrite (1.12) as

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = \log \frac{P_{\mathrm{mic}}(G^*)}{P_{\mathrm{can}}(G^*)}, \qquad (1.16)$$

where G^* is *any* graph in $\mathcal{G}_n$ such that $\vec{C}(G^*) = \vec{C}^*$ (recall that we have assumed that $\vec{C}^*$ is realisable by at least one graph in $\mathcal{G}_n$). The definition in (1.14) then becomes

$$s_{\alpha_\infty} = \lim_{n \to \infty} {\alpha_n}^{-1} \left[\log P_{\mathrm{mic}}(G^*) - \log P_{\mathrm{can}}(G^*) \right], \qquad (1.17)$$

which shows that breaking of ensemble equivalence coincides with $P_{\mathrm{mic}}(G^*)$ and $P_{\mathrm{can}}(G^*)$ having different large deviation behavior on scale α_n. (This is perfectly in line with what was discussed in Section 1.3). Note that (1.17) involves the microcanonical and canonical probabilities of a *single* configuration G^* realising the hard constraint. Apart from its theoretical importance, this fact greatly simplifies mathematical calculations.

To analyse breaking of ensemble equivalence, ideally we would like to be able to identify an underlying *large deviation principle* on a natural scale α_n. This is generally difficult, and so far has only been achieved in the dense regime with the help of *graphons*. See [38] and Section 1.8 to understand why.

§1.5 Summary of Chapter 2

While there is consensus that nonequivalence occurs when the microcanonical specific entropy is non-concave as a function of the energy density in the thermodynamic limit, the classification of the physical mechanisms at the *origin* of nonequivalence is still open. A possible and natural mechanism is the presence of *long-range interactions*. Similarly, phase transitions are naturally associated with long-range order. These "standard mechanisms" for ensemble nonequivalence have been documented also in the study of random graphs.

In Chapter 2 we study certain classes of unipartite networks [92], and show that ensemble nonequivalence can manifest itself via an additional, novel mechanism, unrelated to non-additivity or phase transitions: namely, the presence of an *extensive* number of local topological constraints, i.e., the degrees and/or the strengths (for weighted graphs) of all nodes.[1] This finding explains previously documented signatures of nonequivalence in random graphs with local constraints, such as a finite

---

[1]While in binary (i.e., simple) graphs the *degree* of a node is defined as the number of edges

difference between the microcanonical and the canonical entropy densities [1] and the non-vanishing of the relative fluctuations of the constraints [95]. How generally this result holds beyond the specific uni-partite and bi-partite cases considered so far remains an open question. By considering a much more general class of random graphs with a variable number of constraints, we confirm that the presence of an extensive number of local topological constraints breaks ensemble equivalence, even in the absence of phase transitions or non-additivity.

We start from the characterization of nonequivalence in the simple cases of uni-partite and bi-partite graphs already explored in [92], and subsequently move on to a very general class of graphs with an arbitrary multilayer structure and tunable intra-layer and inter-layer connectivity. The main theorems proved, which (mostly) concern the *sparse regime*, not only characterize nonequivalence *qualitatively*, they also provide a *quantitative* formula for the specific relative entropy. We discuss various important implications of our results, describing properties that are fully general, but also focusing on several special cases of empirical relevance. In addition, we provide an interpretation of the specific relative entropy formula in terms of Poissonisation of the degrees. We also discuss the implications of our results for the study of several empirically relevant classes of "modular" networks that have recently attracted interest in the literature, such as networks with a so-called multi-partite, multiplex [16], time-varying [58], block-model [57], [62] or community structure [43], [84].

§1.6   Summary of Chapter 3

In Chapter 3 we take a fresh look at breaking of ensemble equivalence by analyzing a formula for the relative entropy, based on the *covariance structure* of the canonical ensemble, recently put forward by Squartini and Garlaschelli [93]. We consider the case of a random graph with a given degree sequence (configuration model) and show that the formula correctly predicts that the specific relative entropy is determined by the scaling of the determinant of the covariance matrix of the constraints in the so called δ-tame dense regime, while it requires an extra correction term in the sparse regime and the ultra-dense regime. We also show that the different behaviors found in the different regimes correspond to the degrees being asymptotically Gauss in the dense regime and asymptotically Poisson in the sparse regime, and the dual degrees being asymptotically Poisson in the ultra-dense regime. We also show that, in general, in the canonical ensemble the degrees are distributed according to a multivariate version of the *Poisson-Binomial* distribution [100], which admits the Gauss distribution and the Poisson distribution as limits in appropriate regimes.

---

incident to that node, in weighted graphs (i.e., graphs where edges can carry weights) the *strength* of a node is defined as the total weight of all the edges incident to that node. In Chapter 2, we focus on binary graphs only.

§1.7 Summary of Chapter 4

In Chapters 2 and 3 breaking of ensemble equivalence between the microcanonical ensemble and the canonical ensemble is shown to occur when the constraint is put on the degree sequence (*configuration model*). In this case the constraint becomes a function of the number n of nodes and we can therefore ask an interesting question: How is the relative entropy affected when the number of constraints is reduced, possibly in a way that depends on n?

In Chapter 4 we answer this question by analyzing the effect on the relative entropy when the number of constraints is reduced, i.e., when only *part* of the nodes are constrained degree (and the remaining nodes are left unconstrained). Intuitively, the relative entropy is expected to decrease as the number of constraints decreases. However, this is not a trivial issue, because when the number of constraints is reduced both the microcanonical ensemble and the canonical ensemble change. We consider random graphs with a prescribed partial degree sequence (reduced constraint). The breaking of ensemble equivalence is studied by analyzing how the relative entropy changes as a function of the number of constraints. In particular it is shown that the relative entropy is a monotone function in the number of constraints at the *macroscopic level*, i.e., when a positive fraction of the constraints is removed. More precisely, when only m nodes are constrained and the remaining $n-m$ nodes are unconstrained, the relative entropy turns out to grow like $m \log n$ as $n \to \infty$.

Our analysis is based on a recent formula put forward by Squartini and Garlaschelli [93]. This formula predicts that the relative entropy is determined by the covariance matrix of the constraints in the canonical ensemble, in the regime where the graph is dense. Our result implies that ensemble equivalence breaks down whenever the dense regime is δ-tame, irrespective of the number of degrees m that are constrained, provided m is not too close to n. It is further shown that the expression of the relative entropy corresponds, in the dense regime, to the degrees in the microcanonical ensemble being asymptotically *multivariate Dirac* and in the canonical ensemble being asymptotically *Gauss*.

§1.8 Summary of Chapter 5

In Chapter 5 we analyze breaking of ensemble equivalence for the case in which topological constraints are imposed not only on the total number of edges but also on the total number of wedges, triangles, etc. We work in the dense regime, in which the number of edges per vertex scales proportionally with the number of vertices n. We compute the relative entropy of the two ensembles in the limit as n goes to ∞, where the two ensembles are said to be *equivalent* if this relative entropy divided by n^2 tends to zero (which, up to a constant, can be interpreted as the relative entropy per edge). In particular, we show that the relative entropy divided by n^2 tends to $s_\infty > 0$ when the constraints are *frustrated*. We base our analysis on a *large deviation principle for graphons* and we provide results for three different choices of constraints.

§1.9   Summary of Chapter 6

In Chapter 5 we considered a random graph subject to constraints on the total number of edges and the total number of triangles, in the dense regime. With the help of *large deviation theory for graphons*, we derived a variational formula for $s_\infty = \lim_{n\to\infty} n^{-2} s_n$, where n is the number of vertices and s_n is the *relative entropy* of the microcanonical ensemble with respect to the canonical ensemble. In Chapter 6 we analyze the behavior of s_∞ when the constraints are close to but different from those of the Erdős-Rényi random graph. It turns out that the behavior changes when the total number of triangles is larger, respectively, smaller than that of the Erdős-Rényi random graph with a given total number of edges. In particular, we find that $s_\infty > 0$ when the constraints are *frustrated*, i.e., $T_2^* \neq T_1^{*3}$ with T_1^* the edge density and T_2^* the triangle density. The Erdős-Rényi random graph corresponds to $T_2^* = T_1^{*3}$, for which $s_\infty = 0$. We identify the scaling behavior of s_∞ for fixed T_1^* and $T_2^* \downarrow T_1^{*3}$, respectively, $T_2^* \uparrow T_1^{*3}$, and prove that the way in which s_∞ tends to zero is different for the two limits. We also identify what the constrained random graph asymptotically looks like in the microcanonical ensemble.

§1.10   Development of the chapters

This thesis presents new results about breaking of ensemble equivalence for complex networks. Chapter 2 investigates the role of the number of constraints in the breaking of ensemble equivalence phenomenon. Chapter 2 continues and generalizes the work in [92] and shows that nonequivalence occurs in the presence of an extensive number of topological constraints. Chapter 2 first considers the class of unipartite graphs with the constraint on the degree sequence, in the sparse regime. After that, results are extended to the class of bipartite graphs, and to more complicated classes of graphs with a modular structure. The dense regime is investigated in Chapter 3, where a formula of the relative entropy based on the covariance structure of the canonical ensemble, recently put forward by Garlaschelli and Squartini [93], is confirmed. The study of the configuration model is continued in Chapter 4, where a different question is answered. While extensivity of the number of constraints in the number of nodes was shown to play a crucial role in the phenomenon of breaking of ensemble equivalence, it remains an open question how reduction of the number of constraints affects this phenomenon. Chapter 4 analyzes the effect on the relative entropy when the number of constraint is reduced. It shows that, under certain hypothesis, breaking of ensemble equivalence is monotone in the number of constraints.

Chapters 5 and 6 conclude this thesis with a study of dense graphs with constraints on subgraph structures. In Chapter 5 breaking of ensemble equivalence is analyzed for the case of topological constraints on the number of edges and different subgraphs (wedges, triangles, etc.) at the same time. Here a large deviation principle for graphons is used to prove that breaking of ensemble equivalence occurs whenever the constraints are *frustrated*. Chapter 6 is a continuation of Chapter 5, for the case where the constraints are on the number of edges and triangles at the same time. In particular, constraints are chosen to be close to, but different from, the

so called Erdős-Rényi line. It turns out that when the total number of triangles is larger or smaller then the total number of edges, the behavior of the relative entropy is completely different.

§1.11   Conclusions and Open Problems

In this thesis we analyze breaking of ensemble equivalence for Complex Networks with different types of constraints and in different regimes. The main conclusion of Chapter 2 and 3 is that the physical mechanism behind breaking of ensemble equivalence seems to be the *extensivity* of the number of constraints. In fact, both in the *sparse* and in the *dense* regime, the ensembles are shown to be non-equivalent whenever the the number of constraints grows extensively with the number of nodes. Moreover, Chapter 4 shows how breaking of ensemble equivalence reduces as the number of constrained nodes is reduced. On the other hand, Chapter 5 and 6 show a completely different mechanism behind breaking of ensemble equivalence, namely, *frustration* of the constraints. In the specific case where the constraint is on the number of edges and the number of triangles, the canonical ensemble scales like an Erdős-Rényi random graph with an appropriate edge density, but the microcanonical ensemble does not.

We conclude this introductory chapter with a number of open problems that can serve as a starting point for a future study of breaking of ensemble equivalence phenomenon in complex networks.

1 **Meaning of (non)equivalence**

   In this thesis we analyze breaking of ensemble equivalence at the measure level, i.e., we study the limit of the α_n-relative entropy (1.17) for different constraints, for different regimes and for different values of α_n. One consequence of α_n-equivalence can be derived through (1.2), also known as Pinsker's inequality. This relates a pseudo distance (the relative entropy) to a distance (the total variation distance) and implies that, whenever the ensembles are α_n-equivalent, the total variation distance between the microcanonical and the canonical ensemble does not grow faster than $\sqrt{\alpha_n}$. On the other hand, Pinsker's inequality does not provide full information about what nonequivalence means for typical quantities characterizing the network. It would be interesting to understand what nonequivalence translates into for simulations of real-world networks.

2 **Monotonicity of the relative entropy in the number of constraints**

   In Chapter 4 we analyze the effect on the breaking of ensemble equivalence when the number of constraints is reduced, i.e., when only part of the nodes are constrained in their degree (and the remaining nodes are left unconstrained). We find that the relative entropy is a monotone function in the number of constraints when a *positive fraction* of the constraints is removed.

   a The result of Chapter 4 is based on a formula recently put forward by Squartini and Garlaschelli (see [93], which provides compelling evidence but not a rigorous proof). It would be interesting to prove the monotonicity

property for the relative entropy in a way that does not depend on this formula and possibly for different regimes and other types of constraints as well.

b   Chapter 4 analyzes the relative entropy at a macroscopic level, but nothing is said about the microscopic level. More precisely, it would be interesting to understand how the relative entropy changes when a *single* constraint is removed, rather than a positive fraction of constraints. For example, what is the effect when the longest degree is removed? Is the effect the same or not when we decide to remove the smallest degree, or any other degree for that matter?

3   Functions of the constraints

In this thesis we analyze breaking of ensemble equivalence for a few specific types of constraint. The constraint is put on the number of edges in Chapter 2 and on the degree sequence in Chapters 3 and 4. In Chapter 5 and 6 the constraint is put on the number of edges and the number of triangles. It would be interesting to have a theorem proving the (non)equivalence of ensembles for general types of constraint, and possibly for general functions of the constraints as well.