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

CHAPTER 3

Covariance structure behind breaking of ensemble equivalence in random graphs

This chapter is based on:
D. Garlaschelli, F. den Hollander, and A. Roccaverde. Covariance structure behind breaking of ensemble equivalence in random graphs. *J. Stat. Phys.*, Jul 2018

Abstract

For a random graph subject to a topological constraint, the *microcanonical ensemble* requires the constraint to be met by every realisation of the graph ('hard constraint'), while the *canonical ensemble* requires the constraint to be met only on average ('soft constraint'). It is known that *breaking of ensemble equivalence* may occur when the size of the graph tends to infinity, signalled by a non-zero specific relative entropy of the two ensembles. In this paper we analyse a formula for the relative entropy of generic discrete random structures recently put forward by Squartini and Garlaschelli. We consider the case of a random graph with a given degree sequence (configuration model), and show that in the dense regime this formula correctly predicts that the specific relative entropy is determined by the scaling of the determinant of the matrix of canonical covariances of the constraints. The formula also correctly predicts that an extra correction term is required in the sparse regime and in the ultra-dense regime. We further show that the different expressions correspond to the degrees in the canonical ensemble being asymptotically *Gaussian* in the dense regime and asymptotically *Poisson* in the sparse regime (the latter confirms what we found in earlier work), and the dual degrees in the canonical ensemble being asymptotically *Poisson* in the ultra-dense regime. In general, we show that the degrees follow a multivariate version of the *Poisson-Binomial* distribution in the canonical ensemble.

§3.1 Introduction and main results

§3.1.1 Background and outline

For most real-world networks, a detailed knowledge of the architecture of the network is not available and one must work with a probabilistic description, where 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]. Statistical physics deals with the definition of the appropriate probability distribution over the set of configurations and with the calculation of the resulting properties of the system. Two key choices of probability distribution are:

(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.)

In the limit as the size of the network diverges, the two ensembles are traditionally *assumed* to become equivalent, as a result of the expected vanishing of the fluctuations of the soft constraints (i.e., the soft constraints are expected to become asymptotically hard). However, it is known that this equivalence may be broken, as signalled by a non-zero specific relative entropy of the two ensembles (= on an appropriate scale). In earlier work various scenarios were identified for this phenomenon (see [92], [48], [38] and references therein). In the present paper we take a fresh look at breaking of ensemble equivalence by analysing 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 this 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 dense regime, while it requires an extra correction term in the sparse regime and the ultra-dense regime. We also show that the different behaviours found in the different regimes correspond to the degrees being asymptotically Gaussian in the dense regime and asymptotically Poisson in the sparse regime, and the dual degrees being asymptotically Poisson in the ultra-dense regime. We further note 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 Gaussian distribution and the Poisson distribution as limits in appropriate regimes.

Our results imply that, in all three regimes, ensemble equivalence breaks down in the presence of an extensive number of constraints. This confirms the need for a *principled choice* of the ensemble used in practical applications. Three examples serve as an illustration:

(a) *Pattern detection* is the identification of nontrivial structural properties in a real-world network through comparison with a suitable *null model*, i.e., a random graph model that preserves certain local topological properties of the network (like the degree sequence) but is otherwise completely random.

(b) *Community detection* is the identification of groups of nodes that are more densely connected with each other than expected under a null model, which is a popular special case of pattern detection.

(c) *Network reconstruction* employs purely local topological information to infer higher-order structural properties of a real-world network. This problem arises whenever the global properties of the network are not known, for instance, due to confidentiality or privacy issues, but local properties are. In such cases, optimal inference about the network can be achieved by maximising the entropy subject to the known local constraints, which again leads to the two ensembles considered here.

Breaking of ensemble equivalence means that different choices of the ensemble lead to asymptotically different behaviours. 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 ensemble-nonequivalent models the choice should 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 the constraint (hard or soft) applies in practice.

The remainder of this section is organised as follows. In Section 3.1.2 we introduce the constraints to be considered, which are on the *degree sequence*. In Section 3.1.3 we introduce the various regimes we will be interested in and state a formula for the relative entropy when the constraint is on the degree sequence. In Section 3.1.4 we state the formula for the relative entropy proposed in [93] and present our main theorem. In Section 3.1.5 we close with a discussion of the interpretation of this theorem and an outline of the remainder of the paper.

The microcanonical and the canonical ensemble, as well as the relative entropy density have been defined in Section 1.4.1 and 1.4.2.

§3.1.2 Constraint on the degree sequence

The degree sequence of a graph $G \in \mathcal{G}_n$ is defined as $\vec{k}(G) = (k_i(G))_{i=1}^n$ with $k_i(G) = \sum_{j\neq i} g_{ij}(G)$. In what follows we constrain the degree sequence to a *specific value* $\vec{k}^*$, which we assume to be *graphical*, i.e., there is at least one graph with degree sequence $\vec{k}^*$. The constraint is therefore

$$\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^n \in \{1, 2, \dots, n-2\}^n, \tag{3.1}$$

The microcanonical ensemble, when the constraint is on the degree sequence, is known as the *configuration model* and has been studied intensively (see [95, 92, 99]). For later use we recall the form of the canonical probability in the configuration model, namely,

$$P_{\mathrm{can}}(G) = \prod_{1\le i<j\le n} \left(p_{ij}^*\right)^{g_{ij}(G)} \left(1 - p_{ij}^*\right)^{1-g_{ij}(G)} \tag{3.2}$$

with

$$p_{ij}^* = \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}} \tag{3.3}$$

and with the vector of Lagrange multipliers tuned to the value $\vec{\theta}^* = (\theta_i^*)_{i=1}^n$ such that

$$\langle k_i \rangle = \sum_{j \neq i} p_{ij}^* = k_i^*, \qquad 1 \leq i \leq n. \tag{3.4}$$

Using (1.16), we can write

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = \log \frac{P_{\text{mic}}(G^*)}{P_{\text{can}}(G^*)} = -\log[\Omega_{\vec{k}^*} P_{\text{can}}(G^*)] = -\log Q[\vec{k}^*](\vec{k}^*), \tag{3.5}$$

where $\Omega_{\vec{k}}$ is the number of graphs with degree sequence $\vec{k}$,

$$Q[\vec{k}^*](\vec{k}\,) = \Omega_{\vec{k}}\, P_{\text{can}}\big(G^{\vec{k}}\big) \tag{3.6}$$

is the probability that the degree sequence is equal to $\vec{k}$ under the canonical ensemble with constraint $\vec{k}^*$, $G^{\vec{k}}$ denotes an arbitrary graph with degree sequence $\vec{k}$, and $P_{\text{can}}\big(G^{\vec{k}}\big)$ is the canonical probability in (3.2) rewritten for one such graph:

$$P_{\text{can}}\big(G^{\vec{k}}\big) = \prod_{1 \leq i < j \leq n} \big(p_{ij}^*\big)^{g_{ij}(G^{\vec{k}})} \big(1 - p_{ij}^*\big)^{1 - g_{ij}(G^{\vec{k}})} = \prod_{i=1}^n (x_i^*)^{k_i} \prod_{1 \leq i < j \leq n} (1 + x_i^* x_j^*)^{-1}. \tag{3.7}$$

In the last expression, $x_i^* = e^{-\theta_i^*}$, and $\vec{\theta} = (\theta_i^*)_{i=1}^n$ is the vector of Lagrange multipliers coming from (3.3).

§3.1.3 Relevant regimes

The breaking of ensemble equivalence was analysed in [48] in the so-called *sparse regime*, defined by the condition

$$\max_{1 \leq i \leq n} k_i^* = o(\sqrt{n}\,). \tag{3.8}$$

It is natural to consider the opposite setting, namely, the *ultra-dense regime* in which the degrees are close to $n-1$,

$$\max_{1 \leq i \leq n} (n - 1 - k_i^*) = o(\sqrt{n}\,). \tag{3.9}$$

This can be seen as the *dual* of the sparse regime. We will see in Appendix B that under the map $k_i^* \mapsto n - 1 - k_i^*$ the microcanonical ensemble and the canonical ensemble preserve their relationship, in particular, their relative entropy is invariant.

It is a challenge to study breaking of ensemble equivalence *in between* the sparse regime and the ultra-dense regime, called the *dense regime*. In what follows we consider a subclass of the dense regime, called the *δ-tame regime*, in which the graphs are subject to a certain uniformity condition.

3.1.1 Definition. A degree sequence $\vec{k}^* = (k_i^*)_{i=1}^n$ is called δ-tame if and only if there exists a $\delta \in \left(0, \frac{1}{2}\right]$ such that

$$\delta \leq p_{ij}^* \leq 1 - \delta, \qquad 1 \leq i \neq j \leq n, \tag{3.10}$$

where p_{ij}^* are the canonical probabilities in (3.2)–(3.4).

3.1.2 Remark. The name δ-tame is taken from [9], which studies the number of graphs with a δ-tame degree sequence. Definition 3.1.1 is actually a reformulation of the definition given in [9]. See Appendix A for details.

The condition in (3.10) implies that

$$(n-1)\delta \leq k_i^* \leq (n-1)(1-\delta), \qquad 1 \leq i \leq n, \tag{3.11}$$

i.e., δ-tame graphs are nowhere too thin (sparse regime) nor too dense (ultra-dense regime).

It is natural to ask whether, conversely, condition (3.11) implies that the degree sequence is δ'-tame for some $\delta' = \delta'(\delta)$. Unfortunately, this question is not easy to settle, but the following lemma provides a partial answer.

3.1.3 Lemma. *Suppose that $\vec{k}^* = (k_i^*)_{i=1}^n$ satisfies*

$$(n-1)\alpha \leq k_i^* \leq (n-1)(1-\alpha), \qquad 1 \leq i \leq n, \tag{3.12}$$

for some $\alpha \in (\frac{1}{4}, \frac{1}{2}]$. Then there exist $\delta = \delta(\alpha) > 0$ and $n_0 = n_0(\alpha) \in \mathbb{N}$ such that $\vec{k}^* = (k_i^*)_{i=1}^n$ is δ-tame for all $n \geq n_0$.

Proof. The proof follows from [9, Theorem 2.1]. In fact, by picking $\beta = 1 - \alpha$ in that theorem, we find that we need $\alpha > \frac{1}{4}$. The theorem also gives information about the values of $\delta = \delta(\alpha)$ and $n_0 = n_0(\alpha)$.

§3.1.4 Linking ensemble nonequivalence to the canonical covariances

In this section we investigate an important formula, recently put forward in [93], for the scaling of the relative entropy under a general constraint. The analysis in [93] allows for the possibility that not all the constraints (i.e., not all the components of the vector $\vec{C}$) are linearly independent. For instance, $\vec{C}$ may contain redundant replicas of the same constraint(s), or linear combinations of them. Since in the present paper we only consider the case where $\vec{C}$ is the degree sequence, the different components of $\vec{C}$ (i.e., the different degrees) are linearly independent.

When a K-dimensional constraint $\vec{C}^* = (C_i^*)_{i=1}^K$ with independent components is imposed, then a key result in [93] is the formula

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim \log \frac{\sqrt{\det(2\pi Q)}}{T}, \qquad n \to \infty, \tag{3.13}$$

where

$$Q = (q_{ij})_{1 \leq i,j \leq K} \tag{3.14}$$

is the $K \times K$ covariance matrix of the constraints under the canonical ensemble, whose entries are defined as

$$q_{ij} = \mathrm{Cov}_{P_{\mathrm{can}}}(C_i, C_j) = \langle C_i \, C_j \rangle - \langle C_i \rangle \langle C_j \rangle, \tag{3.15}$$

and

$$T = \prod_{i=1}^{K} \left[1 + O\left(1/\lambda_i^{(K)}(Q) \right) \right], \tag{3.16}$$

with $\lambda_i^{(K)}(Q) > 0$ the i-th eigenvalue of the $K \times K$ covariance matrix Q. This result can be formulated rigorously as

3.1.1 Formula ([93]). If all the constraints are linearly independent, then the limiting relative entropy α_n-density equals

$$s_{\alpha_\infty} = \lim_{n\to\infty} \frac{\log \sqrt{\det(2\pi Q)}}{\alpha_n} + \tau_{\alpha_\infty} \tag{3.17}$$

with α_n the 'natural' speed and

$$\tau_{\alpha_\infty} = - \lim_{n\to\infty} \frac{\log T}{\alpha_n}. \tag{3.18}$$

The latter is zero when

$$\lim_{n\to\infty} \frac{|I_{K_n,R}|}{\alpha_n} = 0 \quad \forall\, R < \infty, \tag{3.19}$$

where $I_{K,R} = \{i = 1, \ldots, K \colon \ \lambda_i^{(K)}(Q) \leq R\}$ with $\lambda_i^{(K)}(Q)$ the i-th eigenvalue of the K-dimensional covariance matrix Q (the notation K_n indicates that K may depend on n). Note that $0 \leq I_{K,R} \leq K$. Consequently, (3.19) is satisfied (and hence $\tau_{\alpha_\infty} = 0$) when $\lim_{n\to\infty} K_n/\alpha_n = 0$, i.e., when the number K_n of constraints grows slower than α_n.

3.1.4 Remark ([93]). Formula 3.1.1, for which [93] offers compelling evidence but not a mathematical proof, can be rephrased by saying that the natural choice of α_n is

$$\tilde{\alpha}_n = \log \sqrt{\det(2\pi Q)}. \tag{3.20}$$

Indeed, if all the constraints are linearly independent and (3.19) holds, then $\tau_{\tilde{\alpha}_n} = 0$ and

$$s_{\tilde{\alpha}_\infty} = 1, \tag{3.21}$$

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = [1 + o(1)]\, \tilde{\alpha}_n. \tag{3.22}$$

We now present our main theorem, which considers the case where the constraint is on the degree sequence: $K_n = n$ and $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^n$. This case was studied in [48], for which $\alpha_n = n$ in the *sparse regime with finite degrees*. Our results here focus on three new regimes, for which we need to increase α_n: the *sparse regime with growing degrees*, the *δ-tame regime*, and the *ultra-dense regime with growing dual degrees*. In all these cases, since $\lim_{n\to\infty} K_n/\alpha_n = \lim_{n\to\infty} n/\alpha_n = 0$, Formula 3.1.1 states that (3.17) holds with $\tau_{\tilde{\alpha}_n} = 0$. Our theorem provides a rigorous and independent mathematical proof of this result.

3.1.5 Theorem. *Formula* 3.1.1 *is true with $\tau_{\alpha_\infty} = 0$ when the constraint is on the degree sequence $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^n$, the scale parameter is $\alpha_n = n\,\overline{f_n}$ with*

$$\overline{f_n} = n^{-1} \sum_{i=1}^n f_n(k_i^*) \quad \textit{with} \quad f_n(k) = \frac{1}{2} \log\left[\frac{k(n-1-k)}{n}\right], \tag{3.23}$$

and the degree sequence belongs to one of the following three regimes:

- *The sparse regime with growing degrees:*

$$\max_{1\le i\le n} k_i^* = o(\sqrt{n}\,), \qquad \lim_{n\to\infty} \min_{1\le i\le n} k_i^* = \infty. \tag{3.24}$$

- *The δ-tame regime (see* (3.2) *and Lemma* 3.1.3*):*

$$\delta \le p_{ij}^* \le 1-\delta, \quad 1 \le i \ne j \le n. \tag{3.25}$$

- *The ultra-dense regime with growing dual degrees:*

$$\max_{1\le i\le n} (n-1-k_i^*) = o(\sqrt{n}\,), \qquad \lim_{n\to\infty} \min_{1\le i\le n} (n-1-k_i^*) = \infty. \tag{3.26}$$

In all three regimes there is breaking of ensemble equivalence, and

$$s_{\alpha_\infty} = \lim_{n\to\infty} s_{\alpha_n} = 1. \tag{3.27}$$

§3.1.5 Discussion and outline

Comparing (3.21) and (3.27), and using (3.20), we see that Theorem 3.1.5 shows that if the constraint is on the degree sequence, then

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim n\overline{f_n} \sim \log\sqrt{\det(2\pi Q)} \tag{3.28}$$

in each of the three regimes considered. Below we provide a heuristic explanation for this result (as well as for our previous results in [48]) that links back to (3.5). In Section 3.2 we prove Theorem 3.1.5.

Chapter 3

Poisson-Binomial degrees in the general case. Note that (3.5) can be rewritten as

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = S\big(\,\delta[\vec{k}^*] \mid Q[\vec{k}^*]\,\big), \tag{3.29}$$

where $\delta[\vec{k}^*] = \prod_{i=1}^n \delta[k_i^*]$ is the *multivariate Dirac distribution* with average $\vec{k}^*$. This has the interesting interpretation that the relative entropy between the distributions P_{mic} and P_{can} *on the set of graphs* coincides with the relative entropy between $\delta[\vec{k}^*]$ and $Q[\vec{k}^*]$ *on the set of degree sequences.*

To be explicit, using (3.6) and (3.7), we can rewrite $Q[\vec{k}^*](\vec{k})$ as

$$Q[\vec{k}^*](\vec{k}) = \Omega_{\vec{k}} \prod_{i=1}^{n} (x_i^*)^{k_i} \prod_{1\le i<j\le n} (1 + x_i^* x_j^*)^{-1}. \tag{3.30}$$

We note that the above distribution is a multivariate version of the *Poisson-Binomial distribution* (or Poisson's Binomial distribution; see Wang [100]). In the univariate case, the Poisson-Binomial distribution describes the probability of a certain number of successes out of a total number of independent and (in general) not identical Bernoulli trials [100]. In our case, the marginal probability that node i has degree k_i in the canonical ensemble, irrespectively of the degree of any other node, is indeed a univariate Poisson-Binomial given by $n-1$ independent Bernoulli trials with success probabilities $\{p_{ij}^*\}_{j\neq i}$. The relation in (3.29) can therefore be restated as

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = S\big(\,\delta[\vec{k}^*] \mid \text{PoissonBinomial}[\vec{k}^*]\,\big), \tag{3.31}$$

where PoissonBinomial$[\vec{k}^*]$ is the multivariate Poisson-Binomial distribution given by (3.30), i.e.,

$$Q[\vec{k}^*] = \text{PoissonBinomial}[\vec{k}^*]. \tag{3.32}$$

The relative entropy can therefore be seen as coming from a situation in which the microcanonical ensemble forces the degree sequence to be exactly $\vec{k}^*$, while the canonical ensemble forces the degree sequence to be Poisson-Binomial distributed with average $\vec{k}^*$.

It is known that the univariate Poisson-Binomial distribution admits two asymptotic limits: (1) a Poisson limit (if and only if, in our notation, $\sum_{j\neq i} p_{ij}^* \to \lambda > 0$ and $\sum_{j\neq i} (p_{ij}^*)^2 \to 0$ as $n \to \infty$ [100]); (2) a Gaussian limit (if and only if $p_{ij}^* \to \lambda_j > 0$ for all $j \neq i$ as $n \to \infty$, as follows from a central limit theorem type of argument). If all the Bernoulli trials are identical, i.e., if all the probabilities $\{p_{ij}^*\}_{j\neq i}$ are equal, then the univariate Poisson-Binomial distribution reduces to the ordinary Binomial distribution, which also exhibits the well-known Poisson and Gaussian limits. These results imply that also the general multivariate Poisson-Binomial distribution in (3.30) admits limiting behaviours that should be consistent with the Poisson and Gaussian limits discussed above for its marginals. This is precisely what we confirm below.

Poisson degrees in the sparse regime. In [48] it was shown that, for a sparse degree sequence,

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim \sum_{i=1}^{n} S\big(\,\delta[k_i^*] \mid \text{Poisson}[k_i^*]\,\big). \tag{3.33}$$

The right-hand side is the sum over all nodes i of the relative entropy of the *Dirac distribution* with average k_i^* w.r.t. the *Poisson distribution* with average k_i^*. We see that, under the sparseness condition, the constraints act on the nodes essentially independently. We can therefore reinterpret (3.33) as the statement

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim S\big(\,\delta[\vec{k}^*] \mid \text{Poisson}[\vec{k}^*]\,\big), \tag{3.34}$$

where $\text{Poisson}[\vec{k}^*] = \prod_{i=1}^n \text{Poisson}[k_i^*]$ is the *multivariate Poisson distribution* with average $\vec{k}^*$. In other words, in this regime

$$Q[\vec{k}^*] \sim \text{Poisson}[\vec{k}^*], \tag{3.35}$$

i.e. the joint multivariate Poisson-Binomial distribution (3.30) essentially decouples into the product of marginal univariate Poisson-Binomial distributions describing the degrees of all nodes, and each of these Poisson-Binomial distributions is asymptotically a Poisson distribution.

Note that the Poisson regime was obtained in [48] under the condition in (3.8), which is less restrictive than the aforementioned condition $k_i^* = \sum_{j\neq i} p_{ij}^* \to \lambda > 0$, $\sum_{j\neq i}(p_{ij}^*)^2 \to 0$ under which the Poisson distribution is retrieved from the Poisson-Binomial distribution [100]. In particular, the condition in (3.8) includes both the case with growing degrees included in Theorem 3.1.5 (and consistent with Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$) and the case with finite degrees, which *cannot* be retrieved from Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$, because it corresponds to the case where all the $n = \alpha_n$ eigenvalues of Q remain finite as n diverges (as the entries of Q themselves do not diverge), and indeed (3.19) does not hold.

Poisson degrees in the ultra-dense regime. Since the ultra-dense regime is the dual of the sparse regime, we immediately get the heuristic interpretation of the relative entropy when the constraint is on an ultra-dense degree sequence $\vec{k}^*$. Using (3.34) and the observations in Appendix B (see, in particular (B.2)), we get

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim S\big(\,\delta[\vec{\ell}^*] \mid \text{Poisson}[\vec{\ell}^*]\,\big), \tag{3.36}$$

where $\vec{\ell}^* = (\ell_i^*)_{i=1}^n$ is the dual degree sequence given by $\ell_i^* = n - 1 - k_i^*$. In other words, under the microcanonical ensemble the dual degrees follow the distribution $\delta[\vec{\ell}^*]$, while under the canonical ensemble the dual degrees follow the distribution $Q[\vec{\ell}^*]$, where in analogy with (3.35),

$$Q[\vec{\ell}^*] \sim \text{Poisson}[\vec{\ell}^*]. \tag{3.37}$$

Similar to the sparse case, the multivariate Poisson-Binomial distribution (3.30) reduces to a product of marginal, and asymptotically Poisson, distributions governing the different degrees.

Again, the case with finite dual degrees cannot be retrieved from Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$, because it corresponds to the case where Q has a diverging (like $n = \alpha_n$) number of eigenvalues whose value remains finite as $n \to \infty$, and (3.19) does not hold. By contrast, the case with growing dual degrees can be retrieved from Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$ because (3.19) holds, as confirmed in Theorem 3.1.5.

Gaussian degrees in the dense regime. We can reinterpet (3.28) as the statement

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim S\big(\,\delta[\vec{k^*}] \mid \text{Normal}[\vec{k^*}, Q]\,\big), \tag{3.38}$$

where Normal$[\vec{k^*}, Q]$ is the *multivariate Normal distribution* with mean $\vec{k^*}$ and covariance matrix Q. In other words, in this regime

$$Q[\vec{k^*}] \sim \text{Normal}[\vec{k^*}, Q], \tag{3.39}$$

i.e., the multivariate Poisson-Binomial distribution (3.30) is asymptotically a multivariate Gaussian distribution whose covariance matrix is in general not diagonal, i.e., the dependencies between degrees of different nodes do *not* vanish, unlike in the other two regimes. Since all the degrees are growing in this regime, so are all the eigenvalues of Q, implying (3.19) and consistently with Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$, as proven in Theorem 3.1.5.

Note that the right-hand side of (3.38), being the relative entropy of a discrete distribution with respect to a continuous distribution, needs to be properly interpreted: the Dirac distribution $\delta[\vec{k^*}]$ needs to be smoothened to a continuous distribution with support in a small ball around $\vec{k^*}$. Since the degrees are large, this does not affect the asymptotics.

Crossover between the regimes. An easy computation gives

$$S\big(\,\delta[k_i^*] \mid \text{Poisson}[k_i^*]\,\big) = g(k_i^*) \quad \text{with} \quad g(k) = \log\left(\frac{k!}{e^{-k}k^k}\right). \tag{3.40}$$

Since $g(k) = [1 + o(1)]\frac{1}{2}\log(2\pi k)$, $k \to \infty$, we see that, as we move from the sparse regime with finite degrees to the sparse regime with growing degrees, the scaling of the relative entropy in (3.33) nicely links up with that of the dense regime in (3.38) via the common expression in (3.28). Note, however, that since the sparse regime with growing degrees is in general incompatible with the dense δ-tame regime, in Theorem 3.1.5 we have to obtain the two scalings of the relative entropy under disjoint assumptions. By contrast, Formula 3.1.1 with $\tau_{\alpha_\infty} = 0$, and hence (3.22), unifies the two cases under the simpler and more general requirement that all the eigenvalues of Q, and hence all the degrees, diverge. Actually, (3.22) is expected to hold in the even more general hybrid case where there are both finite and growing degrees, provided the number of finite-valued eigenvalues of Q grows slower than α_n [93].

Other constraints. It would be interesting to investigate Formula 3.1.1 for constraints other than on the degrees. Such constraints are typically much harder to analyse. In [38] constraints are considered on the total number of edges and the total number of triangles *simultaneously* ($K = 2$) in the dense regime. It was found that, with $\alpha_n = n^2$, breaking of ensemble equivalence occurs for some 'frustrated' choices of these numbers. Clearly, this type of breaking of ensemble equivalence does not arise from the recently proposed [93] mechanism associated with a diverging number of constraints as in the cases considered in this paper, but from the more traditional [97] mechanism of a phase transition associated with the frustration phenomenon.

Outline. Theorem 3.1.5 is proved in Section 3.2. In Appendix A we show that the canonical probabilities in (3.2) are the same as the probabilities used in [9] to define a δ-tame degree sequence. In Appendix B we explain the duality between the sparse regime and the ultra-dense regime.

§3.2 Proof of the Main Theorem

In Section 3.2.2 we prove Theorem 3.1.5. The proof is based on two lemmas, which we state and prove in Section 3.2.1.

§3.2.1 Preparatory lemmas

The following lemma gives an expression for the relative entropy.

3.2.1 Lemma. *If the constraint is a δ-tame degree sequence, then the relative entropy in* (1.16) *scales as*

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = [1 + o(1)]\, \tfrac{1}{2} \log[\det(2\pi Q)], \tag{3.41}$$

where Q is the covariance matrix in (3.14)*. This matrix $Q = (q_{ij})$ takes the form*

$$\begin{cases} q_{ii} = k_i^* - \sum_{j\neq i} (p_{ij}^*)^2 = \sum_{j\neq i} p_{ij}^*(1 - p_{ij}^*), & 1 \le i \le n, \\ q_{ij} = p_{ij}^*(1 - p_{ij}^*), \quad 1 \le i \ne j \le n. \end{cases} \tag{3.42}$$

Proof. To compute $q_{ij} = \mathrm{Cov}_{P_{\text{can}}}(k_i, k_j)$ we take the second order derivatives of the log-likelihood function

$$\begin{aligned} \mathcal{L}(\vec{\theta}) &= \log P_{\text{can}}(G^* \mid \vec{\theta}) \\ &= \log \left[\prod_{1\le i<j\le n} p_{ij}^{g_{ij}(G^*)} (1 - p_{ij})^{(1 - g_{ij}(G^*))} \right], \quad p_{ij} = \frac{e^{-\theta_i - \theta_j}}{1 + e^{-\theta_i - \theta_j}} \end{aligned} \tag{3.43}$$

in the point $\vec{\theta} = \vec{\theta}^*$ [93]. Indeed, it is easy to show that the first-order derivatives are [51]

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\vec{\theta}) = \langle k_i \rangle - k_i^*, \quad \frac{\partial}{\partial \theta_i} \mathcal{L}(\vec{\theta}) \bigg|_{\vec{\theta} = \vec{\theta}^*} = k_i^* - k_i^* = 0 \tag{3.44}$$

and the second-order derivatives are

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} \mathcal{L}(\vec{\theta}) \bigg|_{\vec{\theta} = \vec{\theta}^*} = \langle k_i\, k_j \rangle - \langle k_i \rangle \langle k_j \rangle = \mathrm{Cov}_{P_{\text{can}}}(k_i, k_j). \tag{3.45}$$

This readily gives (3.42).

The proof of (3.41) uses [9, Eq. (1.4.1)], which says that if a δ-tame degree sequence is used as constraint, then

$$P_{\text{mic}}^{-1}(G^*) = \Omega_{\vec{C}^*} = \frac{e^{H(p^*)}}{(2\pi)^{n/2} \sqrt{\det(Q)}}\, e^C, \tag{3.46}$$

where Q and p^* are defined in (3.42) and (3.70) below, while e^C is sandwiched between two constants that depend on δ:

$$\gamma_1(\delta) \leq e^C \leq \gamma_2(\delta). \tag{3.47}$$

From (3.46) and the relation $H(p^*) = -\log P_{\mathrm{can}}(G^*)$, proved in Lemma A.1 below, we get the claim.

The following lemma shows that the diagonal approximation of $\log(\det Q)/n\overline{f}_n$ is good when the degree sequence is δ-tame.

3.2.2 Lemma. *Under the δ-tame condition,*

$$\log(\det Q_D) + o(n\,\overline{f}_n) \leq \log(\det Q) \leq \log(\det Q_D) \tag{3.48}$$

with $Q_D = \mathrm{diag}(Q)$ the matrix that coincides with Q on the diagonal and is zero off the diagonal.

Proof. We use [60, Theorem 2.3], which says that if

(1) $\det(Q)$ is real,

(2) Q_D is non-singular with $\det(Q_D)$ real,

(3) $\lambda_i(A) > -1$, $1 \leq i \leq n$,

then

$$e^{-\frac{n\rho^2(A)}{1+\lambda_{\min}(A)}} \det Q_D \leq \det Q \leq \det Q_D. \tag{3.49}$$

Here, $A = Q_D^{-1} Q_{\mathrm{off}}$, with Q_{off} the matrix that coincides with Q off the diagonal and is zero on the diagonal, $\lambda_i(A)$ is the i-th eigenvalue of A (arranged in decreasing order), $\lambda_{\min}(A) = \min_{1\leq i\leq n} \lambda_i(A)$, and $\rho(A) = \max_{1\leq i\leq n} |\lambda_i(A)|$.

We begin by verifying (1)–(3).

(1) Since Q is a symmetric matrix with real entries, $\det Q$ exists and is real.

(2) This property holds thanks to the δ-tame condition. Indeed, since $q_{ij} = p^*_{i,j}(1 - p^*_{i,j})$, we have

$$0 < \delta^2 \leq q_{ij} \leq (1-\delta)^2 < 1, \tag{3.50}$$

which implies that

$$0 < (n-1)\delta^2 \leq q_{ii} = \sum_{j\neq i} q_{ij} \leq (n-1)(1-\delta)^2. \tag{3.51}$$

(3) It is easy to show that $A = (a_{ij})$ is given by

$$a_{ij} = \begin{cases} \frac{q_{ij}}{q_{ii}}, & 1 \leq i \neq j \leq n, \\ 0, & 1 \leq i \leq n, \end{cases} \tag{3.52}$$

where q_{ij} is given by (3.42). Since $q_{ij} = q_{ji}$, the matrix A is symmetric. Moreover, since $q_{ii} = \sum_{j \neq i} q_{ij}$, the matrix A is also Markov. We therefore have

$$1 = \lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A) \geq -1. \tag{3.53}$$

From (3.50) and (3.52) we get

$$0 < \frac{1}{n-1}\left(\frac{\delta}{1-\delta}\right)^2 \leq a_{ij} \leq \frac{1}{n-1}\left(\frac{1-\delta}{\delta}\right)^2. \tag{3.54}$$

This implies that the Markov chain on $\{1, \ldots, n\}$ with transition matrix A starting from i can return to i with a positive probability after an arbitrary number of steps ≥ 2. Consequently, the last inequality in (3.53) is strict.

We next show that

$$\frac{n\rho^2(A)}{1 + \lambda_{\min}(A)} = o(n\,\overline{f}_n). \tag{3.55}$$

Together with (3.49) this will settle the claim in (3.48). From (3.53) it follows $\rho(A) = 1$, so we must show that

$$\lim_{n\to\infty} [1 + \lambda_{\min}(A)]\,\overline{f}_n = \infty. \tag{3.56}$$

Using [104, Theorem 4.3], we get

$$\lambda_{\min}(A) \geq -1 + \frac{\min_{1\leq i\neq j\leq n} \pi_i a_{ij}}{\min_{1\leq i\leq n} \pi_i}\,\mu_{\min}(L) + 2\gamma. \tag{3.57}$$

Here, $\pi = (\pi_i)_{i=1}^n$ is the invariant distribution of the reversible Markov chain with transition matrix A, while $\mu_{\min}(L) = \min_{1\leq i\leq n} \lambda_i(L)$ and $\gamma = \min_{1\leq i\leq n} a_{ii}$, with $L = (L_{ij})$ the matrix such that, for $i \neq j$, $L_{ij} = 1$ if and only if $a_{ij} > 0$, while $L_{ii} = \sum_{j\neq i} L_{ij}$.

We find that $\pi_i = \frac{1}{n}$ for $1 \leq i \leq n$, $L_{ij} = 1$ for $1 \leq i \neq j \leq n$, $L_{ii} = n - 1$ for $1 \leq i \leq n$, and $\gamma = 0$. Hence (3.57) becomes

$$\lambda_{\min}(A) \geq -1 + (n-2) \min_{1\leq i\neq j\leq n} a_{ij} \geq -1 + \frac{n-2}{n-1}\left(\frac{\delta}{1-\delta}\right)^2, \tag{3.58}$$

where the last inequality comes from (3.54). To get (3.56) it therefore suffices to show that $\overline{f}_\infty = \lim_{n\to\infty} \overline{f}_n = \infty$. But, using the δ-tame condition, we can estimate

$$\begin{aligned} \frac{1}{2}\log\left[\frac{(n-1)\delta(1-\delta+n\delta)}{n}\right] \leq \overline{f}_n &= \frac{1}{2n}\sum_{i=1}^n \log\left[\frac{k_i^*(n-1-k_i^*)}{n}\right] \\ &\leq \frac{1}{2}\log\left[\frac{(n-1)(1-\delta)(\delta+n(1-\delta))}{n}\right], \end{aligned} \tag{3.59}$$

and both bounds scale like $\frac{1}{2}\log n$ as $n \to \infty$.

§3.2.2 Proof (Theorem 3.1.5)

Proof. We deal with each of the three regimes in Theorem 3.1.5 separatetely.

The sparse regime with growing degrees. Since $\vec{k}^* = (k_i^*)_{i=1}^n$ is a sparse degree sequence, we can use [48, Eq. (3.12)], which says that

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = \sum_{i=1}^n g(k_i^*) + o(n), \qquad n \to \infty, \tag{3.60}$$

where $g(k) = \log\left(\frac{k!}{k^k e^{-k}}\right)$ is defined in (3.40). Since the degrees are growing, we can use Stirling's approximation $g(k) = \frac{1}{2}\log(2\pi k) + o(1)$, $k \to \infty$, to obtain

$$\sum_{i=1}^n g(k_i^*) = \tfrac{1}{2}\sum_{i=1}^n \log\left(2\pi k_i^*\right) + o(n) = \tfrac{1}{2}\left[n\log 2\pi + \sum_{i=1}^n \log k_i^*\right] + o(n). \tag{3.61}$$

Combining (3.60)–(3.61), we get

$$\frac{S_n(P_{\text{mic}} \mid P_{\text{can}})}{n\overline{f}_n} = \tfrac{1}{2}\left[\frac{\log 2\pi}{\overline{f}_n} + \frac{\sum_{i=1}^n \log k_i^*}{n\overline{f}_n}\right] + o(1). \tag{3.62}$$

Recall (3.23). Because the degrees are sparse, we have

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n \log k_i^*}{n\overline{f}_n} = 2. \tag{3.63}$$

Because the degrees are growing, we also have

$$\overline{f}_\infty = \lim_{n\to\infty} \overline{f}_n = \infty. \tag{3.64}$$

Combining (3.62)–(3.64) we find that $\lim_{n\to\infty} S_n(P_{\text{mic}} \mid P_{\text{can}})/n\overline{f}_n = 1$.

The ultra-dense regime with growing dual degrees. If $\vec{k}^* = (k_i^*)_{i=1}^n$ is an ultra-dense degree sequence, then the dual $\vec{\ell}^* = (\ell_i^*)_{i=1}^n = (n-1-k_i^*)_{i=1}^n$ is a sparse degree sequence. By Lemma B.2, the relative entropy is invariant under the map $k_i^* \to \ell_i^* = n-1-k_i^*$. So is $\bar{f}_n$, and hence the claim follows from the proof in the sparse regime.

The δ-tame regime. It follows from Lemma 3.2.1 that

$$\lim_{n\to\infty} \frac{S_n(P_{\text{mic}} \mid P_{\text{can}})}{n\overline{f}_n} = \tfrac{1}{2}\left[\lim_{n\to\infty} \frac{\log 2\pi}{\overline{f}_n} + \lim_{n\to\infty} \frac{\log(\det Q)}{n\overline{f}_n}\right]. \tag{3.65}$$

From (3.59) we know that $\overline{f}_\infty = \lim_{n\to\infty} \overline{f}_n = \infty$ in the *δ-tame regime*. It follows from Lemma 3.2.2 that

$$\lim_{n\to\infty} \frac{\log(\det Q)}{n\overline{f}_n} = \lim_{n\to\infty} \frac{\log(\det Q_D)}{n\overline{f}_n}. \tag{3.66}$$

To conclude the proof it therefore suffices to show that

$$\lim_{n\to\infty} \frac{\log(\det Q_D)}{n\overline{f}_n} = 2. \tag{3.67}$$

Using (3.51) and (3.59), we may estimate

$$\frac{2\log[(n-1)\delta^2]}{\log\frac{(n-1)(1-\delta)(\delta+n(1-\delta))}{n}} \leq \frac{\sum_{i=1}^{n}\log(q_{ii})}{n\,\overline{f}_n} = \frac{\log(\det Q_D)}{n\,\overline{f}_n} \leq \frac{2\log[(n-1)(1-\delta)^2]}{\log\frac{(n-1)\delta(1-\delta+n\delta)}{n}}. \tag{3.68}$$

Both sides tend to 2 as $n\to\infty$, and so (3.67) follows.

§A Appendix

Here we show that the canonical probabilities in (3.2) are the same as the probabilities used in [9] to define a δ-tame degree sequence.

For $q=(q_{ij})_{1\leq i,j\leq n}$, let

$$E(q) = -\sum_{1\leq i\neq j\leq n} q_{ij}\log q_{ij} + (1-q_{ij})\log(1-q_{ij}). \tag{3.69}$$

be the entropy of q. For a given degree sequence $(k_i^*)_{i=1}^n$, consider the following maximisation problem:

$$\begin{cases} \max E(q), \\ \sum_{j\neq i} q_{ij} = k_i^*,\ 1\leq i\leq n, \\ 0\leq q_{ij}\leq 1,\ 1\leq i\neq j\leq n. \end{cases} \tag{3.70}$$

Since $q\mapsto E(q)$ is strictly concave, it attains its maximum at a unique point.

A.1 Lemma. *The canonical probability takes the form*

$$P_{\mathrm{can}}(G) = \prod_{1\leq i<j\leq n} \left(p_{ij}^*\right)^{g_{ij}(G)}\left(1-p_{ij}^*\right)^{1-g_{ij}(G)}, \tag{3.71}$$

where $p^*=(p_{ij}^*)$ *solves* (3.70)*. In addition,*

$$\log P_{\mathrm{can}}(G^*) = -H(p^*). \tag{3.72}$$

Proof. It was shown in [48] that, for a degree sequence constraint,

$$P_{\mathrm{can}}(G) = \prod_{1\leq i<j\leq n} \left(p_{ij}^*\right)^{g_{ij}(G)}\left(1-p_{ij}^*\right)^{1-g_{ij}(G)} \tag{3.73}$$

with $p_{ij}^* = \frac{e^{-\theta_i^*-\theta_j^*}}{1+e^{-\theta_i^*-\theta_j^*}}$, where $\vec{\theta}^*$ has to be tuned such that

$$\sum_{j\neq i} p_{ij}^* = k_i^*, \qquad 1\leq i\leq n. \tag{3.74}$$

On the other hand, the solution of (3.70) via the Lagrange multiplier method gives that

$$q_{ij}^* = \frac{e^{-\phi_i^*-\phi_j^*}}{1+e^{-\phi_i^*-\phi_j^*}}, \tag{3.75}$$

where $\vec{\phi}^*$ has to be tuned such that

$$\sum_{j \neq i} q^*_{ij} = k^*_i, \qquad 1 \leq i \leq n. \tag{3.76}$$

This implies that $q^*_{ij} = p^*_{ij}$ for all $1 \leq i \neq j \leq n$. Moreover,

$$\begin{aligned} \log P_{\mathrm{can}}(G^*) + H(p^*) &= \sum_{1 \leq i < j \leq n} g_{ij}(G^*) \log\left(\frac{p^*_{ij}}{1-p^*_{ij}}\right) - \sum_{1 \leq i < j \leq n} p^*_{ij} \log\left(\frac{p^*_{ij}}{1-p^*_{ij}}\right) \\ = -\sum_{1 \leq i < j \leq n} g_{ij}(G^*)(\theta^*_i + \theta^*_j) &+ \sum_{1 \leq i < j \leq n} p^*_{ij}(\theta^*_i + \theta^*_j) = \sum_{i=1}^{n} \theta^*_i \sum_{j \neq i}(p^*_{ij} - g_{ij}(G^*)) = 0, \end{aligned} \tag{3.77}$$

where the last equation follows from the fact that

$$\sum_{j \neq i} g_{ij}(G^*) = \sum_{j \neq i} p^*_{ij} = k^*_i, \qquad 1 \leq i \leq n. \tag{3.78}$$

§B Appendix

We explain the duality between the sparse regime and the ultra-dense regime.

Let $\vec{k}^* = (k^*_i)_{i=1}^n$ be an ultra-dense degree sequence,

$$\max_{1 \leq i \leq n} (n - 1 - k^*_i) = o(\sqrt{n}), \tag{3.79}$$

and let $\vec{\ell}^* = (\ell^*_i)_{i=1}^n$ be the *dual* degree sequence defined by $\ell^*_i = n - 1 - k^*_i$. Clearly, $\vec{\ell}^* = (\ell^*_i)_{i=1}^n$ is a sparse degree sequence,

$$\max_{1 \leq i \leq n} \ell^*_i = o(\sqrt{n}). \tag{3.80}$$

B.1 Lemma. *Let P_{can} and $\widehat{P_{\mathrm{can}}}$ denote the canonical ensembles in* (1.9) *when $\vec{C}^* = \vec{k}^* = (k^*_i)_{i=1}^n$, respectively, $\vec{C}^* = \vec{\ell}^* = (\ell^*_i)_{i=1}^n$. Then*

$$P_{\mathrm{can}}(G) = \widehat{P_{\mathrm{can}}}(G_c), \qquad G \in \mathcal{G}_n, \tag{3.81}$$

where G and G_c are complementary graphs, i.e.,

$$g_{ij}(G_c) = 1 - g_{ij}(G), \qquad 1 \leq i \neq j \leq n. \tag{3.82}$$

Proof. From the definition of the canonical probabilities we have

$$P_{\mathrm{can}}(G) = P_{\mathrm{can}}(G \mid \vec{\theta}^*), \qquad \widehat{P_{\mathrm{can}}}(G) = P_{\mathrm{can}}(G \mid \vec{\phi}^*), \tag{3.83}$$

where

$$P_{\mathrm{can}}(G \mid \vec{\theta}) = \frac{\exp[-\vec{\theta} \cdot \vec{k}(G)]}{Z(\vec{\theta})}, \qquad \vec{k}(G) = \sum_{j \neq i} g_{ij}(G), \tag{3.84}$$

and the values $\vec{\theta}^*$ and $\vec{\phi}^*$ are such that

$$\left.\frac{\partial F(\vec{\theta})}{\partial \theta_i}\right|_{\vec{\theta}=\vec{\theta}^*} = -\langle k_i \rangle_{P_{\mathrm{can}}(\cdot \mid \vec{\theta}^*)} = -k_i^*, \tag{3.85}$$

$$\left.\frac{\partial F(\vec{\theta})}{\partial \theta_i}\right|_{\vec{\theta}=\vec{\phi}^*} = -\langle k_i \rangle_{P_{\mathrm{can}}(\cdot \mid \vec{\phi}^*)} = -\ell_i^*. \tag{3.86}$$

The free energy is $F(\vec{\theta}) = \log Z(\vec{\theta})$, and its i-th partial derivative in the Lagrange multiplier that fixes the average of the i-th constraint. We show that $\vec{\theta}^* = -\vec{\phi}^*$.

Write

$$Z(\vec{\theta}) = \sum_{G\in\mathcal{G}_n} e^{-\vec{\theta}\cdot\vec{k}(G)} = \sum_{G\in\mathcal{G}_n} e^{-\sum_{i=1}^n \theta_i (n-1-k(G_c))} = e^{-(n-1)\sum_{i=1}^n \theta_i}\, Z(-\vec{\theta}). \tag{3.87}$$

Using (3.85) and (3.87), we get

$$-k_i^* = \left.\frac{\partial F(\vec{\theta})}{\partial \theta_i}\right|_{\vec{\theta}=\vec{\theta}^*} = -(n-1) + \langle k_i \rangle_{P_{\mathrm{can}}(\cdot \mid -\vec{\theta}^*)}. \tag{3.88}$$

Since $k_i^* = n-1-\ell_i^*$, we obtain

$$\ell_i^* = \langle k_i \rangle_{P_{\mathrm{can}}(\cdot \mid -\vec{\theta}^*)}. \tag{3.89}$$

From (3.86), (3.89) and the uniqueness of the Lagrange multipliers, we get

$$\vec{\theta}^* = -\vec{\phi}^*. \tag{3.90}$$

Using (3.87) and (3.90), we obtain

$$\begin{aligned} \widehat{P_{\mathrm{can}}}(G_c) &= P_{\mathrm{can}}(G_c \mid \vec{\phi}^*) = P_{\mathrm{can}}(G_c \mid -\vec{\theta}^*) = \frac{\exp[\vec{\theta}^*\cdot\vec{k}(G_c)]}{Z(-\vec{\theta}^*)} \\ &= \frac{\exp[-\vec{\theta}^*\cdot\vec{k}(G)]}{Z(-\vec{\theta}^*)\, e^{-(n-1)\sum_{i=1}^n \theta_i^*}} = \frac{\exp[-\vec{\theta}^*\cdot\vec{k}(G)]}{Z(\vec{\theta}^*)} = P_{\mathrm{can}}(G), \end{aligned} \tag{3.91}$$

which settles (3.81).

B.2 Lemma. *Let*

- P_{mic} *and* P_{can} *denote the microcanonical ensemble in* (1.6)*, respectively, the canonical ensemble in* (1.9)*, when* $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^n$ *with* k_i^* *satisfying* (3.79)*.*
- $\widehat{P_{\mathrm{mic}}}$ *and* $\widehat{P_{\mathrm{can}}}$ *denote the microcanonical ensemble in* (1.6)*, respectively, the canonical ensemble in* (1.9)*, when* $\vec{C}^* = \vec{\ell}^* = (\ell_i^*)_{i=1}^n$ *with* $\ell_i^* = n-1-k_i^*$ *the dual degree satisfying* (3.80)*.*

Then the relative entropy in (1.16) *satisfies*

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = S_n(\widehat{P_{\mathrm{mic}}} \mid \widehat{P_{\mathrm{can}}}). \tag{3.92}$$

Proof. Consider a graph G^* with degree sequence $\vec{k}(G^*) = \vec{k}^*$. Then

$$P_{\mathrm{mic}}(G^*) = |\{G \in \mathcal{G}_n \colon \vec{k}(G) = \vec{k}^*\}|^{-1} = |\{G \in \mathcal{G}_n \colon \vec{k}(G) = \vec{\ell}^*\}|^{-1} = \widehat{P_{\mathrm{mic}}}(G_c^*), \tag{3.93}$$

where G_c^* and G^* are complementary graphs, so that $\vec{k}(G_c^*) = \vec{\ell}^*$. Using Lemma B.1, we have

$$P_{\mathrm{can}}(G^*) = \widehat{P_{\mathrm{can}}}(G_c^*). \tag{3.94}$$

Combine (1.16), (3.93) and (3.94), to get

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = \log \frac{P_{\mathrm{mic}}(G^*)}{P_{\mathrm{can}}(G^*)} = \log \frac{\widehat{P_{\mathrm{mic}}}(G_c^*)}{\widehat{P_{\mathrm{can}}}(G_c^*)} = S_n(\widehat{P_{\mathrm{mic}}} \mid \widehat{P_{\mathrm{can}}}). \tag{3.95}$$

CHAPTER 3