Breaking of ensemble equivalence for complex networks

Roccaverde, A.

Citation

Version:	Not Applicable (or Unknown)
License:	[Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)
Downloaded from:	https://hdl.handle.net/1887/67095

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

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

Is Breaking of Ensemble Equivalence Monotone in the Number of Constraints?

This chapter is based on:

A. Roccaverde. Is breaking of ensemble equivalence monotone in the number of constraints? *Indagationes Mathematicae*, 2018

Abstract

Breaking of ensemble equivalence between the microcanonical ensemble and the canonical ensemble may occur for random graphs whose size tends to infinity, and is signaled by a non-zero specific relative entropy of the two ensembles. In [48] and [50] it was shown that breaking occurs when the constraint is put on the degree sequence (configuration model). It is not known what is the effect on the relative entropy 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). Intuitively, the relative entropy is expected to decrease. However, this is not a trivial issue because when constraints are removed both the microcanonical ensemble and the canonical ensemble change. In this paper a formula for the relative entropy valid for generic discrete random structures, recently formulated by Squartini and Garlaschelli, is used to prove that the relative entropy is monotone in the number of constraints when the constraint is on the degrees of the nodes. It is further shown that the expression for 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 *Gaussian*.

§4.1 Introduction and main results

§4.1.1 Background

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.)

Breaking of ensemble equivalence means that different choices of the ensemble lead to asymptotically different behaviors. Consequently, while for applications based on ensemble-equivalent models the choice of ensemble can be based on mathematical convenience, for those based on ensemble-nonequivalent models the choice should be determined by the system one wants to apply to, i.e., dictated by a theoretical criterion that indicates *a priori* which ensemble is the appropriate one to be used. It is known that ensemble equivalence may be broken, signaled by a non-zero specific relative entropy between the two ensembles. It is expected that when the number of constraints grows extensively in the number of nodes, then typically there is breaking of ensemble equivalence. This has been shown to be the case when the setting is simple or bipartite graphs and the constraint is on the number of links (1 constraint and ensemble equivalence) or on the full degree sequence (n constraints and non-equivalence) [48]. Later, in [50] and [38], also the dense regime was investigated and it was shown that the relative entropy between the two ensembles grows even faster. In general, the constraint is a multidimensional vector and its components represent the single quantities that are constrained. From now on, with the word 'constraint' we mean the 'vector constraint' and with the plural 'constraints' we mean the 'components' of the vector. This means when we talk about the number of constraints we actually mean the dimension of the vector constraint. In some cases this number can be very large, for example, when the constraint is on the degree sequence (a large number of nodes which need all to have the right degree).

Once the constraint becomes a function of the number n of nodes (for example, the degree sequence), we can 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? Intuitively, the relative entropy should decrease, but this is not a trivial issue because both the microcanonical and the canonical ensemble change when the constraints are changed. Of particular interest for the present paper is the main

result of [50]. There it was proven that, when a δ-tame degree sequence is put as a constraint on the set of simple graphs, than the relative entropy between the two ensembles grows as $n \log n$. 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 of the number of constraints. More precisely, when only m nodes are constrained and the remaining $n - m$ nodes are left unconstrained, the relative entropy is shown to grow like $m \log n$. 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 under the canonical ensemble, in the regime where the graph is dense. Our result implies that ensemble equivalence breaks down whenever the regime is δ-tame, irrespective of the number of degrees m that are constrained, provided m is not of order n.

Outline

Our paper is organized as follows. In Section 4.1 the background, the model and the main theorem are discussed. In Section 4.2 the main theorem is proved, together with a few basic lemmas that are needed along the way. Appendix A derives an expression for the canonical ensemble when a partial degree sequence is put as constraint. Appendix B discusses the δ-tame condition for a partial degree sequence.

The remainder of Section 4.1 is organized as follows. In Section 4.1.1 we discussed the background of the problem. In Section 4.1.2 we describe the model when the constraint is put on the full degree sequence, in Section 4.1.3 when the constraint is put on the partial degree sequence. Here we also define the δ-tame regime when the constraint is on the partial degree sequence. In Section 4.1.4 we state a formula for the relative entropy presented in [93] and state the main theorem. In section 4.1.5 we interpret the main theorem by stating how the degrees are distributed in the two ensembles.

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.

§4.1.2 Constraint on the full degree sequence

The model of this section comes from [48] and [50]. The full degree sequence of a graph $G \in \mathcal{G}_n$ is defined as the vector $\vec{k}(G) = (k_i(G))_{i=1}^n$ with $k_i(G) = \sum_{j \neq i} g_{ij}(G)$. The degree sequence is set 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

$$\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^n \in \{1, 2, \ldots, n-2\}^n. \tag{4.1}$$

This constraint was studied in various regimes: in [48] in the sparse regime, and in [50] in the ultra-dense and the δ-tame regime. The microcanonical ensemble, when the constraint is put on the degree sequence, is known as the *configuration model* and

has been studied in detail (see [95, 92, 99]). In the sparse (and in the ultra-dense) regime, the microcanonical ensemble cannot be computed exactly, but there are good approximations with an error that is vanishing when the relative entropy is computed in the limit as $n \to \infty$ [48], [9]. In the δ-tame regime, this approximation does not hold, but the relative entropy can still be investigated with other tools [50]. The canonical ensemble can be computed in every regime and takes the form

$$P_{\mathrm{can}}(G) = \prod_{1 \leq i < j \leq n} \left(p_{ij}^* \right)^{g_{ij}(G)} \left(1 - p_{ij}^* \right)^{1 - g_{ij}(G)}, \qquad (4.2)$$

with

$$p_{ij}^* = \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}}, \qquad (4.3)$$

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

$$\langle k_i \rangle = \sum_{\substack{1 \leq j \leq n \\ j \neq i}} p_{ij}^* = k_i^*, \qquad \forall \, 1 \leq i \leq n. \qquad (4.4)$$

The results in [48] show that there is breaking of ensemble equivalence with $\alpha_n = n$ when the regime is sparse and ultra-dense. The results in [50] show that the relative entropy grows like $\alpha_n = \frac{1}{2} n \log n$. The purpose of the paper is to investigate what happens when part of the n constraints degrees are removed and how the relative entropy is affected by this. In the next section the partial constraint is presented and the main theorem is stated.

§4.1.3   Constraint on the partial degree sequence

In this section we look at a different model. The constraint is put on the partial degree sequence instead of on the full degree sequence, more precisely, only the first $m < n$ nodes are constrained while the remaining nodes are left unconstrained. The partial degree sequence of a graph $G \in \mathcal{G}_n$ is defined as the vector $\vec{k}(G) = (k_i(G))_{i=1}^m$ where $k_i(G) = \sum_{1 \leq j \neq j \leq n} g_{ij}(G)$. The constraint is set to be a *specific m-dimensional* vector $\vec{k}^*$, which we assume to be *graphical*, i.e., there exist at least one graph $G^* \in \mathcal{G}_n$ with partial degree sequence $\vec{k}^*$. The constraint is therefore

$$\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^m \in \{1, 2, \ldots, n-2\}^m, \qquad (4.5)$$

As mentioned above, the microcanonical ensemble can be computed approximately when the constraint is put on the full degree sequence. However, when the constraint is put on the partial degree sequence, no good approximation is available. The situation is different for the canonical ensemble, which can still be computed. Appendix A is dedicated to the study of the canonical ensemble when a partial degree sequence is put as a constraint. This leads to

$$P_{\mathrm{can}}(G) = 2^{-\binom{n-m}{2}} \prod_{1 \leq i < j \leq m} \left(p_{ij}^* \right)^{g_{ij}(G)} \left(1 - p_{ij}^* \right)^{1 - g_{ij}(G)} \prod_{i=1}^m \left(p_i^* \right)^{s_i(G)} \left(1 - p_i^* \right)^{n - m - s_i(G)}$$

$$(4.6)$$

with

$$p_{ij}^* = \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}}, \qquad p_i^* = \frac{e^{-\theta_i^*}}{1 + e^{-\theta_i^*}}, \qquad s_i(G) = \sum_{j=m+1}^{n} g_{ij}(G), \qquad (4.7)$$

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

$$\langle k_i \rangle = \sum_{\substack{1 \le j \le m \\ j \ne i}} p_{ij}^* + (n-m) p_i^* = k_i^*, \qquad 1 \le i \le m. \qquad (4.8)$$

The canonical ensemble has an interesting dual structure, consisting of the product of two canonical probabilities, which we call unipartite probability and bipartite probability, and an overall factor $2^{-\binom{n-m}{2}}$. The unipartite probability,

$$\prod_{1 \le i < j \le m} \left(p_{ij}^* \right)^{g_{ij}(G)} \left(1 - p_{ij}^* \right)^{1 - g_{ij}(G)},$$

is precisely the canonical ensemble obtained when the constraint is put on the full degree sequence $\vec{u}^* = (u_i^*)_{i=1}^m$, with $u_i^* = \sum_{\substack{1 \le j \le m \\ j \ne i}} p_{ij}^*$, on the subset of graphs with m nodes $\mathcal{G}_m$. The bipartite probability,

$$\prod_{i=1}^{m} (p_i^*)^{s_i(G)} (1 - p_i^*)^{n-m-s_i(G)},$$

is precisely the canonical bipartite probability obtained when the constraint is put on the top layer of a bipartite graph. More precisely, the configuration space is the set of bipartite graphs $\mathcal{G}_{m,n-m}$ with m nodes in the top layer and $n-m$ nodes in the bottom layer. The constraint is put on the degree sequence in the top layer only and corresponds to the vector $\vec{b}^* = (b_i^*)_{i=1}^m$ with $b_i^* = (n-m)p_i^*$. Moreover, the average i-th degree $\langle k_i \rangle$ with respect to the canonical ensemble (4.6) equals k_i^* and is given by the balance equation (4.8). This equation shows that the i-th unipartite constraint u_i^* and the i-th bipartite constraint b_i^* sum up to the i-th original constraint k_i^*.

4.1.1 Definition. A partial degree sequence $\vec{k}^* = (k_i^*)_{i=1}^m$, put as a constraint on the set of configurations $\mathcal{G}_n$ with $m < n$, is said to be δ-tame if and only if there exists a $\delta \in \left(0, \frac{1}{2} \right]$ such that

$$\delta \le p_{ij}^* \le 1 - \delta, \qquad 1 \le i \ne j \le m, \qquad (4.9)$$

where p_{ij}^* are the canonical probabilities in (4.6)–(4.8).

It is easy to prove that, given a δ-tame partial degree sequence $\vec{k}^* = (k_i^*)_{i=1}^m$, the bipartite probabilities $(p_i^*)_{i=1}^m$ are also δ-tame, namely, satisfy

$$\delta' \le p_i^* \le 1 - \delta', \qquad \forall \, 1 \le i \le m, \qquad (4.10)$$

for some $\delta' \in \left(0, \frac{1}{2} \right]$. This is discussed in more detail in Appendix B. Condition (4.9) has a trivial implication for the degree sequence:

$$(m-1)\delta \le u_i^* \le (m-1)(1-\delta), \qquad 1 \le i \le m, \qquad (4.11)$$

CHAPTER 4

$$(n-m)\delta' \le b_i^* \le (n-m)(1-\delta'), \qquad 1 \le i \le m. \tag{4.12}$$

Since $\delta' = \frac{1}{1+(\frac{1-\delta}{\delta})^{3/2}} < \delta$ for all $\delta \in [0, 1/2)$ and $u_i^* + b_i^* = k_i^*$, it follows that

$$(n-1)\delta' \le k_i^* \le (n-1)(1-\delta'), \qquad 1 \le i \le m. \tag{4.13}$$

This means that δ-tame graphs are neither too thin (sparse regime) nor too dense (ultra-dense regime). It is natural to ask whether, conversely, condition (4.13), or a similar condition involving only the original degrees $\vec{k}^* = (k_i^*)_{i=1}^m$, is sufficient to prove that the partial degree sequence is δ-tame for some $\delta = \delta(\delta')$, in the sense of Definition 4.1.1. Unfortunately, this question is not easy to settle, but the following lemma provides a partial answer.

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

$$(n-1)\delta' + (n-m) \le k_i^* \le (n-1)(1-\delta'), \qquad 1 \le i \le m, \tag{4.14}$$

for some $\delta' \in (\frac{1}{4}, \frac{1}{2}]$. Then there exist $\delta = \delta(\delta') > 0$ and $n_0 = n_0(\delta') \in \mathbb{N}$ such that $\vec{k}^* = (k_i^*)_{i=1}^m$ is a δ-tame partial degree sequence, in the sense of Definition 4.1.1, for all $n \ge n_0$.

Proof. Condition (4.14), with $u_i^* = k_i^* - b_i^*$ and $b_i^* \in [0, n-m]$, gives

$$(n-1)\delta' \le u_i^* \le (n-1)(1-\delta'), \qquad 1 \le i \le m. \tag{4.15}$$

The proof follows from (4.15) and [9, Theorem 2.1]. In fact, applying that theorem with $\alpha = \delta'$, $\beta = 1 - \delta'$ and with $\delta' > \frac{1}{4}$, we get

$$\delta \le p_{ij}^* \le 1 - \delta, \qquad 1 \le i \ne j \le m. \tag{4.16}$$

Moreover, [9, Theorem 2.1] also gives information about the values of $\delta = \delta(\delta')$ and $n_0 = n_0(\delta')$. $\quad\blacksquare$

§4.1.4 Linking ensemble nonequivalence to the canonical covariances

In this section we describe 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

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim \log \frac{\sqrt{\det(2\pi Q)}}{T}, \qquad n \to \infty, \tag{4.17}$$

where

$$Q = (q_{ij})_{1 \leq i,j \leq K} \tag{4.18}$$

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

$$q_{ij} = \mathrm{Cov}_{P_{\mathrm{can}}}(C_i, C_j) = \langle C_i \, C_j \rangle - \langle C_i \rangle \langle C_j \rangle, \tag{4.19}$$

and

$$T = \prod_{i=1}^{K} \left[1 + O \left(1/\lambda_i^{(K)}(Q) \right) \right], \tag{4.20}$$

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

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

$$s_{\alpha_\infty} = \lim_{n \to \infty} \frac{\log \sqrt{\det(2\pi Q)}}{\alpha_n} + \tau_{\alpha_\infty} \tag{4.21}$$

with

$$\tau_{\alpha_\infty} = - \lim_{n \to \infty} \frac{\log T}{\alpha_n}. \tag{4.22}$$

The latter is zero when

$$\lim_{n \to \infty} \frac{|I_{K_n,R}|}{\alpha_n} = 0 \quad \forall \, R < \infty, \tag{4.23}$$

where $I_{K,R} = \{i = 1, \dots, K \colon \lambda_i^{(K)}(Q) \leq R\}$ 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, (4.23) 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.

4.1.3 Remark ([93]). Formula 4.1.1, for which [93] offers compelling evidence but not a mathematical proof, can be rephrased by saying that the natural choice of α_n is

$$\tilde{\alpha}_n = \log \sqrt{\det(2\pi Q)}. \tag{4.24}$$

Indeed, if all the constraints are linearly independent and (4.23) holds, then $\tau_{\tilde{\alpha}_n} = 0$ and

$$s_{\tilde{\alpha}_\infty} = 1, \tag{4.25}$$
$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = [1 + o(1)] \, \tilde{\alpha}_n. \tag{4.26}$$

Formula 4.1.1 has been verified in several examples, namely, all the models in [48] and [50].

Next we present our main theorem, which considers the case where the constraint is on the partial degree sequence $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^m$ in the δ-*tame regime* defined in Definition 4.1.1.

4.1.4 Theorem. *Suppose that:*

- *The constraint is put on the partial degree sequence $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^m$ on the space of simple graphs $\mathcal{G}_n$ with $0 \leq m \leq n$.*

- *$\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^m$ is a δ-tame partial degree sequence, namely, the canonical probabilities $(p_{ij}^*)_{1 \leq i \neq j \leq m}$ satisfy*

$$\delta \leq p_{ij}^* \leq 1 - \delta, \quad 1 \leq i \neq j \leq m. \tag{4.27}$$

- *Formula 4.1.1 is valid in the above framework.*

- *The scale parameter is $\alpha_n = \dfrac{m \log n}{2}$.*

- *$m = m(n)$ satisfies*

$$\lim_{n \to \infty} \frac{n - m}{m} \log n = \infty. \tag{4.28}$$

Then there is breaking of ensemble equivalence, and

$$s_{\alpha_\infty} = \lim_{n \to \infty} s_{\alpha_n} = 1. \tag{4.29}$$

Condition (4.28) fails when $n - m = O(\frac{m}{\log n})$, i.e., when the number of unconstrained nodes is sufficiently small. We expect that (4.29) continues to hold even in this case, but our proof breaks down.

§4.1.5 Discussion

Theorem 4.1.4 analyses the relative entropy at a macroscopic level, but says nothing about what happens at the microscopic level. More precisely, it does not identify how the relative entropy changes when a single constraint is removed, rather than a positive fraction of constraints. A microscopic analysis could reveal what is the effect when e.g. the longest degree is removed, or the smallest degree, or any other degree. The result in Theorem 4.1.4 is far from trivial. In fact, when the number of constraints is reduced, it can become either easier or more difficult to compute microcanonical and canonical ensembles. The case when the constraint is put on the degree sequence provides a clear example. If the constraint is put on the full degree sequence, then the microcanonical ensemble can be *asymptotically computed* [9]. As soon as one or more degrees are removed (meaning that some nodes are left unconstrained), the structure of the problem changes completely. The symmetry of the constraints is broken by the removal, and this makes it more difficult to compute the number of graphs with a prescribed partial degree sequence. On the other hand, the canonical problem can still be solved and has an interesting structure (Appendix A). This makes it possible to use the formula proposed by Garlaschelli and Squartini [93], which only makes use of the canonical ensemble to analyze the relative entropy between the two ensembles. Theorem 4.1.4 clearly exhibits the monotonicity property of the relative entropy in the case where the constraint is put on the degrees. Indeed, under the hypotheses

written above, the relative entropy $S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}})$ grows like $m \log n$, where m is the number of constrained nodes and n is the total number of nodes. This shows that the relative entropy is monotone in the number of constraints on scale n.

We next provide a heuristic explanation for Theorem 4.1.4 (in analogy with what was done in [48] and [50]).

Heuristic explanation of Theorem 4.1.4. Using (1.16), we can write the relative entropy between the ensembles as

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = \log \frac{P_{\mathrm{mic}}(G^*)}{P_{\mathrm{can}}(G^*)} = -\log[\Omega^n_{\vec{k}^*} P_{\mathrm{can}}(G^*)] = -\log Q^n[\vec{k}^*](\vec{k}^*), \quad (4.30)$$

where $\Omega^n_{\vec{k}^*}$ is the number of graphs with n nodes and partial degree sequence $\vec{k}^* = (k_i^*)_{i=1}^m$,

$$Q^n[\vec{k}^*](\vec{k}) = \Omega^n_{\vec{k}} P_{\mathrm{can}}\big(G^{\vec{k}}\big) \qquad (4.31)$$

is the probability that the partial degree sequence is equal to $\vec{k}$ under the canonical ensemble with constraint $\vec{k}^*$, $G^{\vec{k}}$ denotes an arbitrary graph with partial degree sequence $\vec{k}$, and $P_{\mathrm{can}}\big(G^{\vec{k}}\big)$ is the canonical probability rewritten for one such graph. Indeed, (1.9) shows that the canonical probability is constant for all graphs with the same constraint, in our case, for all graphs with the same partial degree sequence. Using (1.9) and (4.65) we can rewrite the canonical probability in the form

$$P_{\mathrm{can}}\big(G^{\vec{k}}\big) = 2^{-\binom{n-m}{2}} \prod_{i=1}^m \frac{x_i^{*k_i}}{(1+x_i^*)^{n-m}} \prod_{1 \le i < j \le m} (1 + x_i^* x_j^*)^{-1}, \qquad (4.32)$$

where $x_i^* = e^{-\theta_i^*}$, and $\vec{\theta}^* = (\theta_i^*)_{i=1}^m$ is the vector of Lagrange multipliers coming from (4.7). Equation (4.30) can be rewritten as

$$S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}}) = S\big(\,\delta[\vec{k}^*] \mid Q^n[\vec{k}^*]\,\big), \qquad (4.33)$$

where $\delta[\vec{k}^*] = \prod_{i=1}^m \delta[k_i^*]$ is the *multivariate Dirac distribution* with average $\vec{k}^*$. We can interpret the relative entropy between P_{mic} and P_{can} *on the set of graphs $\mathcal{G}_n$* as the relative entropy between $\delta[\vec{k}^*]$ and $Q^n[\vec{k}^*]$ *on the set of degree sequences*. More precisely, combining (4.31) and (4.32), we can rewrite $Q^n[\vec{k}^*](\vec{k})$ as

$$Q^n[\vec{k}^*](\vec{k}) = \Omega^n_{\vec{k}}\, 2^{-\binom{n-m}{2}} \prod_{i=1}^m \frac{x_i^{*k_i}}{(1+x_i^*)^{n-m}} \prod_{1 \le i < j \le m} (1 + x_i^* x_j^*)^{-1}. \qquad (4.34)$$

The distribution in (4.34) is a multivariate version of the *Poisson-Binomial distribution* [100]. The univariate Poisson-Binomial distribution describes the probability of a certain number of successes out of a total number of independent, possibly nonidentical, Bernoulli trials. In our case, the marginal probability that node i has degree k_i, under the canonical ensemble, irrespectively of the degree of any other node, is a univariate Poisson-Binomial distribution with a total number of $n-1$ Bernoulli trials: $m-1$ independent Bernoulli trials with success probabilities $\{p_{ij}^*\}_{1 \le j \ne i \le m}$ and $n-m$

independent Bernoulli trials with the same success probability p_i^*. The relation in (4.33) therefore becomes

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = S\big(\delta[\vec{k^*}] \mid \text{PoissonBinomial}[\vec{k^*}]\big), \qquad (4.35)$$

where PoissonBinomial$[\vec{k^*}]$ is the multivariate Poisson-Binomial distribution given by (4.34), i.e.,

$$Q^n[\vec{k^*}] = \text{PoissonBinomial}[\vec{k^*}]. \qquad (4.36)$$

The relative entropy between the microcanonical and the canonical ensemble can be seen as coming from the limiting 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 distributed as a multivariate Poisson-Binomial with average $\vec{k^*}$.

Two different regimes for the Poisson–Binomial distribution. As has already been said in Chapter 3, the univariate Poisson-Binomial distribution admits two asymptotic limits: Poisson or Gaussian [100]. A Poisson limit occurs whenever $\sum_{j \neq i} p_{ij}^* \to \lambda > 0$ and $\sum_{j \neq i} (p_{ij}^*)^2 \to 0$ as $n \to \infty$, while a Gaussian limit occurs whenever $p_{ij}^* \to \lambda_j > 0$ for all $j \neq i$ as $n \to \infty$. In the simple case of identical Bernoulli trials, i.e., all the probabilities $\{p_{ij}^*\}_{j \neq i}$ are equal, the univariate Poisson-Binomial distribution reduces to a Binomial distribution, which is known to admit Poisson and Gaussian limits. This implies that also the multivariate Poisson-Binomial distribution in (4.34) admits limits that should be consistent with the Poisson and Gaussian ones for its marginals. Below we present two different situations.

Gaussian constrained degrees in the δ-tame regime. Comparing (4.25) and (4.29), and using (4.24), we see that Theorem 4.1.4 shows that if the constraint is on the partial degree sequence, then

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim m \log n \sim \log \sqrt{\det(2\pi Q)} \qquad (4.37)$$

in the δ-tame regime and under the condition in (4.28). Equation (4.37) can be reinterpreted as

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) \sim S\big(\delta[\vec{k^*}] \mid \text{Normal}[\vec{k^*}, Q]\big), \qquad (4.38)$$

where Normal$[\vec{k^*}, Q]$ is the *multivariate Normal distribution* with mean $\vec{k^*}$ and covariance matrix Q. Basically, in the δ-tame regime,

$$Q^n[\vec{k^*}] \sim \text{Normal}[\vec{k^*}, Q]. \qquad (4.39)$$

The multivariate Poisson-Binomial distribution in (4.34) becomes asymptotically a multivariate Gaussian distribution which, in general, has a non-diagonal covariance matrix, i.e., there can be dependence between the degrees of the different nodes.

The right-hand side of (4.38) deserves clarification, because it has to be properly interpreted. In fact, it describes the relative entropy of a discrete distribution with respect to a continuous distribution. Technically, the Dirac distribution $\delta[\vec{k^*}]$ must be smoothed to a continuous distribution with support on a small ball around $\vec{k^*}$. Since the degrees are large, this does not affect the asymptotics.

Poisson-Binomial unconstrained degrees in the δ-tame regime. It is interesting to study the distribution of the degrees of the unconstrained nodes in the canonical ensemble. The canonical probability of the $(m+1)$-th degree (the first unconstrained node) can be computed and the same steps can be used to compute the canonical probabilities of the other unconstrained nodes, which follow the same probability law. The canonical probability that the $(m+1)$-th node is equal to some value $x \in \{0, 1, \ldots, n-1\}$ can be written as:

$$
\sum_{\substack{G \in \mathcal{G}_n \\ k_{m+1}(G)=x}} P_{\text{can}}(G)
$$

$$
= 2^{-\binom{n-m}{2}} \sum_{\substack{G \in \mathcal{G}_n \\ k_{m+1}(G)=x}} \prod_{1 \leq i < j \leq m} \left(p_{ij}^*\right)^{g_{ij}(G)} \left(1 - p_{ij}^*\right)^{1-g_{ij}(G)}
$$

$$
\prod_{i=1}^{m} \left(p_i^*\right)^{s_i(G)} \left(1 - p_i^*\right)^{n-m-s_i(G)}
$$

$$
= 2^{-\binom{n-m}{2}} \sum_{G \in A} \prod_{i=1}^{m} \left(p_i^*\right)^{g_{im+1}(G)} \left(1 - p_i^*\right)^{1-g_{im+1}(G)} \tag{4.40}
$$

$$
= 2^{-\binom{n-m}{2}} 2^{\binom{n-m-1}{2}} \sum_{G \in A \cap B} \prod_{i=1}^{m} p_i^{* g_{im+1}(G)} \left(1 - p_i^*\right)^{1-g_{im+1}(G)}
$$

$$
= \sum_{G \in A \cap B} \left(\tfrac{1}{2}\right)^{(n-m-1)} \prod_{i=1}^{m} p_i^{* g_{im+1}(G)} \left(1 - p_i^*\right)^{1-g_{im+1}(G)}
$$

$$
= P(Po - Bi[p_1^*, \ldots, p_m^*, \tfrac{1}{2}, \ldots, \tfrac{1}{2}] = x),
$$

where

$$
A = \left\{ G \in \mathcal{G}_n : k_{m+1}(G) = x, \ g_{ij}(G) = 0 \ \forall i \in [1, m], \ j = \in [1, n] \setminus \{m+1\}, \ i \neq j \right\},
$$
$$
B = \left\{ G \in \mathcal{G}_n : g_{ij}(G) = 0 \ \forall i = m+2, \ldots, n, \ j = m+2, \ldots, n, \ i \neq j \right\}, \tag{4.41}
$$

and $Po - Bi[p_1^*, \ldots, p_m^*, \tfrac{1}{2}, \ldots, \tfrac{1}{2}]$ is the Poisson-Binomial distribution given by the m independent trials p_i^*, $i = 1, \ldots, m$, and the $n - m - 1$ independent Bernoulli trials with the same success probability $\tfrac{1}{2}$. This means that, for each $j = m+1, \ldots, n$, the canonical probability of the degree of the j-th node is distributed as a Poisson-Binomial random variable with $n - 1$ entries: $p_1^*, \ldots, p_1^*, \tfrac{1}{2}, \ldots, \tfrac{1}{2}$.

§4.2 Proof of the Main Theorem

The proof is based on two lemmas, which are stated and proved in Section 4.2.1. In Section 4.2.2 Theorem 4.1.4 is proved.

§4.2.1 Preparatory lemmas

The following lemma gives an expression for the relative entropy.

4.2.1 Lemma. *If the constraint is on the partial degree sequence $(k_i^*)_{i=1}^m$, then the relative entropy in (1.16) equals*

$$S_n(P_{\text{mic}} \mid P_{\text{can}}) = \tfrac{1}{2} \log[\det(2\pi Q)] - \log T^*, \qquad (4.42)$$

where Q is the covariance matrix in (4.18) and T^* is the error in (4.20). The matrix $Q = (q_{ij})$ takes the form

$$\begin{cases} q_{ii} = \sum_{1 \leq j \leq m, j \neq i} p_{ij}^*(1 - p_{ij}^*) + (n - m)p_i^*(1 - p_i^*), & 1 \leq i \leq m, \\ q_{ij} = p_{ij}^*(1 - p_{ij}^*), & 1 \leq i \neq j \leq m. \end{cases} \qquad (4.43)$$

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

$$\mathcal{L}(\vec{\theta}) = \log P_{\text{can}}(G^* \mid \vec{\theta})$$

$$= \log \left[2^{-\binom{n-m}{2}} \prod_{1 \leq i < j \leq n} p_{ij}^{g_{ij}(G^*)}(1 - p_{ij})^{(1 - g_{ij}(G^*))} \prod_{i=1}^m p_i^{s_i(G^*)} (1 - p_i)^{n - m - s_i(G^*)} \right], \qquad (4.44)$$

with

$$p_{ij} = \frac{e^{-\theta_i - \theta_j}}{1 + e^{-\theta_i - \theta_j}}, \quad p_i = \frac{e^{-\theta_i}}{1 + e^{-\theta_i}}, \qquad (4.45)$$

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

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\vec{\theta}) = \langle k_i \rangle - k_i^*, \quad \frac{\partial}{\partial \theta_i} \mathcal{L}(\vec{\theta}) \bigg|_{\vec{\theta} = \vec{\theta}^*} = k_i^* - k_i^* = 0 \qquad (4.46)$$

and the second-order derivatives are

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} \mathcal{L}(\vec{\theta}) \bigg|_{\vec{\theta} = \vec{\theta}^*} = \langle k_i \rangle \langle k_j \rangle - \langle k_i \, k_j \rangle = -\text{Cov}_{P_{\text{can}}}(k_i, k_j). \qquad (4.47)$$

Taking the second-order derivatives of the log-likelihood function, we get (4.43). The proof of (4.42) uses [93, Formula 25]. $\blacksquare$

The following lemma shows that a diagonal approximation of the matrix Q is good for a δ-tame partial degree sequence and $\alpha_n = m \log n$.

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

$$\log(\det Q_D) + o(m \log n) \leq \log(\det Q) \leq \log(\det Q_D) \qquad (4.48)$$

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

Proof. 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) $\lambda_i(A) > -1$, $1 \le i \le m$,

then

$$e^{-\frac{m\rho^2(A)}{1+\lambda_{\min}(A)}} \det Q_D \le \det Q \le \det Q_D. \tag{4.49}$$

Here, $A = Q_D^{-1}Q_{\text{off}}$, 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), $\lambda_{\min}(A) = \min_{1 \le i \le n} \lambda_i(A)$, and $\rho(A) = \max_{1 \le i \le n} |\lambda_i(A)|$.
We verify (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 and Lemma B.1. In fact

$$0 < \delta^2 \le q_{ij} \le (1-\delta)^2 < 1, \tag{4.50}$$

and

$$(m-1)\delta^2 + (n-m)\delta'^2 \le q_{ii} =\le (m-1)(1-\delta)^2 + (n-m)(1-\delta')^2. \tag{4.51}$$

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

$$a_{ij} = \begin{cases} \frac{q_{ij}}{q_{ii}} = \frac{p_{ij}^*(1-p_{ij}^*)}{\sum_{1 \le k \le m, k \ne i} p_{ik}^*(1-p_{ik}^*)+(n-m)p_i^*(1-p_i^*)}, & 1 \le i \ne j \le m \\ 0 & 1 \le i = j \le m, \end{cases} \tag{4.52}$$

where q_{ij} is given by (4.43). The Gershgorin circle theorem says the eigenvalues of the matrix A satisfy

$$| \lambda_i(A) | \le R_i = \sum_{j \ne i} a_{ij} = \frac{\sum_{1 \le k \le m, k \ne i} p_{ik}^*(1-p_{ik}^*)}{\sum_{1 \le k \le m, k \ne i} p_{ik}^*(1-p_{ik}^*) + (n-m)p_i^*(1-p_i^*)}, \quad 1 \le i \le m. \tag{4.53}$$

Using the δ-tame condition, we find the bound

$$| \lambda_i(A) | \le \max_{1 \le i \le m} R_i < 1 - A(\delta), \tag{4.54}$$

with $A(\delta) = \frac{(n-m)\delta'}{(m-1)(1-\delta)^2+(n-m)(1-\delta')}$. In principle, $A(\delta)$ also depends on δ', but δ' is itself function of δ. Equation (4.54) immediately gives $\rho(A) < 1$, namely

$$-\frac{m\rho^2(A)}{1+\lambda_{\min}(A)} > -\frac{m}{1+\lambda_{\min}(A)}. \tag{4.55}$$

Next we show that

$$-\frac{m}{1+\lambda_{\min}(A)} = o(m \log n). \tag{4.56}$$

Together with (4.49) this will settle the claim in (4.48). We must show that

$$\lim_{n \to \infty} (1 + \lambda_{\min}(A)) \log n = \infty. \tag{4.57}$$

CHAPTER 4

Using equation (4.54) again, it follows $1 + \lambda_{\min}(A) > A(\delta)$. Therefore it suffices to prove that

$$\lim_{n \to \infty} A(\delta) \log n = \infty. \tag{4.58}$$

The result is trivial when $A(\delta)$ is constant ($\frac{n-m}{m} \to$ constant) or $A(\delta) \to \infty$ ($\frac{n-m}{m} \to \infty$). On the other hand, when $A(\delta) \to 0$ ($\frac{n-m}{m} \to 0$), the condition $\frac{n-m}{m} \log n \to \infty$ is needed to conclude the proof.

§4.2.2 Proof (Theorem 4.1.4)

Proof. When $\alpha_n = \frac{m \log n}{2}$, Lemma 4.2.1 says

$$\lim_{n \to \infty} \frac{S_n(P_{\mathrm{mic}} \mid P_{\mathrm{can}})}{\alpha_n} = \lim_{n \to \infty} \frac{\log 2\pi}{m \log n} + \lim_{n \to \infty} \frac{\log(\det Q)}{m \log n} - \lim_{n \to \infty} \frac{\log T^*}{2m \log n}. \tag{4.59}$$

The last term (the error) tends to zero. In fact, in [93] it is proved that $\lim_{n \to \infty} \frac{\log T^*}{m \log n} = 0$ unless the number of eigenvalues of Q that have a finite limit as $n \to \infty$ which is indeed the case when a partial δ-tame degree sequence is put as a constraint and $\alpha_n = \frac{m \log n}{2}$.

Using the δ-tame condition, we get from Lemma 4.2.2 that

$$\lim_{n \to \infty} \frac{\log(\det Q)}{m \log n} = \lim_{n \to \infty} \frac{\log(\det Q_D)}{m \log n}. \tag{4.60}$$

To conclude the proof it therefore suffices to show that

$$\lim_{n \to \infty} \frac{\log(\det Q_D)}{m \log n} = 1. \tag{4.61}$$

Using (4.51), we have

$$\begin{aligned}
\frac{\log[(m-1)\delta^2 + (n-m)\delta']}{\log n} &\leq \frac{\sum_{i=1}^m \log q_{ii}}{m \log n} = \frac{\log(\det Q_D)}{m \log n} \\
&\leq \frac{\log[(m-1)(1-\delta)^2 + (n-m)(1-\delta')]}{\log n}.
\end{aligned} \tag{4.62}$$

Both sides tend to 1 as $n \to \infty$, and so (4.61) follows.

§A Appendix

In this appendix we identify the structure of the canonical ensemble when the constraint is put on the partial degree sequence for the first $m < n$ nodes. The partial degree sequence $\vec{k}(G) = (k_i(G))_{i=1}^m$ is set to a *specific m-dimensional* vector $\vec{k}^*$, which is assumed to be *graphical*, i.e., there is at least one graph $G^* \in \mathcal{G}_n$ with partial degree sequence $\vec{k}^*$. The constraint is therefore

$$\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^m \in \{1, 2, \dots, n-2\}^m. \tag{4.63}$$

The canonical ensemble has Hamiltonian $H(G, \vec{\theta}) = \sum_{i=1}^{m} \theta_i k_i(G)$, where G is a graph belonging to $\mathcal{G}_n$, and $k_i(G) = \sum_{j \neq i} g_{ij}(G)$ is the degree of node i. It is easy to transform the Hamiltonian into

$$H(G, \vec{\theta}) = \sum_{1 \leq i < j \leq m} (\theta_i + \theta_j) g_{ij}(G) + \sum_{i=1}^{m} \theta_i \sum_{j=m+1}^{n} g_{ij}(G) \qquad (4.64)$$

Using this form, we see that the partition function equals

$$Z(\vec{\theta}) = \sum_{G \in \mathcal{G}_n} e^{-H(G, \vec{\theta})} = \sum_{G \in \mathcal{G}_n} \prod_{1 \leq i < j \leq m} e^{-(\theta_i + \theta_j) g_{ij}(G)} \prod_{i=1}^{m} \prod_{j=m+1}^{n} e^{-\theta_i g_{ij}(G)}$$

$$= 2^{\binom{n-m}{2}} \prod_{1 \leq i < j \leq m} (1 + e^{-(\theta_i + \theta_j)}) \prod_{i=1}^{m} \prod_{j=m+1}^{n} (1 + e^{-\theta_i}) \qquad (4.65)$$

$$= 2^{\binom{n-m}{2}} \prod_{1 \leq i < j \leq m} (1 + e^{-(\theta_i + \theta_j)}) \prod_{i=1}^{m} (1 + e^{-\theta_i})^{(n-m)}.$$

Inserting the partition function into the canonical expression, we get

$$P_{\text{can}}(G \mid \theta) = 2^{-\binom{n-m}{2}} \prod_{1 \leq i < j \leq m} p_{ij}^{g_{ij}(G)} (1 - p_{ij})^{1 - g_{ij}(G)} \prod_{i=1}^{m} p_i^{s_i(G)} (1 - p_i)^{n - m - s_i(G)}$$

$$(4.66)$$

with

$$p_{ij} = \frac{e^{-\theta_i - \theta_j}}{1 + e^{-\theta_i - \theta_j}}, \qquad p_i = \frac{e^{-\theta_i}}{1 + e^{-\theta_i}}, \qquad s_i(G) = \sum_{j=m+1}^{n} g_{ij}(G). \qquad (4.67)$$

It remains to tune the Lagrange multipliers to the values such that the average constraint equals the vector $\vec{C}^* = \vec{k}^* = (k_i^*)_{i=1}^{m} \in \{1, 2, \ldots, n-2\}^m$. The average energy of the i-th degree with respect to the probability distribution $P_{\text{can}}(\cdot \mid \theta)$ corresponds to the derivative with respect to θ_i of the logarithm of the partition function (free energy). This means that the values $(\theta_i^*)_{i=1}^{m}$ must satisfy

$$\langle k_i \rangle = \sum_{\substack{1 \leq j \leq m \\ j \neq i}} p_{ij}^* + (n - m) p_i^* = k_i^*, \qquad 1 \leq i \leq m \qquad (4.68)$$

with

$$p_{ij}^* = \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}}, \qquad p_i^* = \frac{e^{-\theta_i^*}}{1 + e^{-\theta_i^*}}. \qquad (4.69)$$

The canonical ensemble therefore takes the form

$$P_{\text{can}}(G) = 2^{-\binom{n-m}{2}} \prod_{1 \leq i < j \leq m} {p_{ij}^*}^{g_{ij}(G)} \left(1 - p_{ij}^*\right)^{1 - g_{ij}(G)} \prod_{i=1}^{m} {p_i^*}^{s_i(G)} \left(1 - p_i^*\right)^{n - m - s_i(G)}.$$

$$(4.70)$$

The expression in (4.70) has an interpretation. Indeed, the canonical formula che be split into two parts:

$$P_{\text{can}}(G) = P_{\text{can}}^{U}(G) P_{\text{can}}^{B}(G) \, 2^{-\binom{n-m}{2}}, \qquad (4.71)$$

with

$$P_{\text{can}}^U(G) = \prod_{1 \leq i < j \leq m} p_{ij}^{*\,g_{ij}(G)} \left(1 - p_{ij}^*\right)^{1-g_{ij}(G)} \tag{4.72}$$

and

$$P_{\text{can}}^B(G) = \prod_{i=1}^m p_i^{*\,s_i(G)} \left(1 - p_i^*\right)^{n-m-s_i(G)}. \tag{4.73}$$

The unipartite probability, $P_{\text{can}}^U(G)$, is the canonical probability obtained when the constraint is put on the full degree sequence $\vec{u}^* = (u_i^*)_{i=1}^m$ on the set $\mathcal{G}_m$. The constrained degree sequence is precisely $u_i^* = \sum_{\substack{1 \leq j \leq m \\ j \neq i}} p_{ij}^*$. The bipartite probability $P_{\text{can}}^B(G)$ is the canonical bipartite probability obtained when the constraint is put only on the top layer of a bipartite graph. In this case the configuration space is the set of bipartite graphs $\mathcal{G}_{m,n-m}$ with m nodes on the top layer and $n-m$ nodes on the bottom layer. The constrained top layer degree sequence is $\vec{b}^* = (b_i^*)_{i=1}^m$, where $b_i^* = (n-m)p_i^*$. The third factor $2^{-\binom{n-m}{2}}$ is the inverse of the number of possible (unconstrained) graphs with $n-m$ nodes. In conclusion, the canonical probability in (4.70) can be interpreted as the product of two canonical probabilities, $P_{\text{can}}^U(G)$ and $P_{\text{can}}^B(G)$, and the number $2^{-\binom{n-m}{2}}$. Both canonical probabilities have an m-dimensional degree sequence as a constraint $\vec{u}^* = (u_i^*)_{i=1}^m$ and $\vec{b}^* = (b_i^*)_{i=1}^m$, put on the respective configuration spaces. Furthermore, two degree sequences sum up to the original degree sequence, namely,

$$u_i^* + b_i^* = k_i^* \qquad \forall i = 1, \dots, m. \tag{4.74}$$

For this reason $(p_{ij}^*)_{i,j=1}^m$ are called the unipartite probabilities and $(p_i^*)_{i=1}^m$ the bipartite probabilities.

§B Appendix

In this appendix we identify the structure of the δ-tame condition when a partial degree sequence $(k_i^*)_{i=1}^m$ is put as a constraint on $\mathcal{G}_n$. The definition comes from the situation where a full degree sequence $(k_i^*)_{i=1}^n$ is fixed on $\mathcal{G}_n$ [9]. In the full degree sequence situation the canonical probability takes the form

$$P_{\text{can}}(G) = \prod_{1 \leq i < j \leq n} \left(p_{ij}^*\right)^{g_{ij}(G)} \left(1 - p_{ij}^*\right)^{1-g_{ij}(G)} \tag{4.75}$$

with

$$p_{ij}^* = \frac{e^{-\theta_i^* - \theta_j^*}}{1 + e^{-\theta_i^* - \theta_j^*}} \quad \forall\, i \neq j, \tag{4.76}$$

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

$$\langle k_i \rangle = \sum_{\substack{1 \leq j \leq n \\ j \neq i}} p_{ij}^* = k_i^*, \qquad 1 \leq i \leq n. \tag{4.77}$$

The degree sequence $(k_i^*)_{i=1}^n$ is said to be δ-tame when there exists a $\delta \in (0, \frac{1}{2}]$ such that, for each $1 \leq i \neq j \leq n$, the canonical probabilites satisfy

$$\delta < p_{ij}^* < 1 - \delta. \tag{4.78}$$

B.1 Definition (δ-tame partial degree sequence). We say that such a sequence is δ-tame when there exists a $\delta \in (0, \frac{1}{2}]$ such that, for each $1 \leq i \neq j \leq m$, the canonical probabilites defined in (4.6)–(4.8) satisfy

$$\delta < p_{ij}^* < 1 - \delta \qquad \forall\, 1 \leq i \neq j \leq m. \tag{4.79}$$

B.2 Lemma. *If $(k_i^*)_{i=1}^m$ is a partial degree sequence on $\mathcal{G}_n$ and it is δ-tame in the sense of Definition B.1, then the canonical bipartite probabilities satisfy*

$$\delta' < p_i^* < 1 - \delta' \qquad \forall\, 1 \leq i \neq j \leq m, \tag{4.80}$$

for some $\delta' \in (0, \frac{1}{2}]$.

Proof. The canonical probabilities, tuned with the proper $(\theta_i^*)_{i=1}^m$, satisfy

$$p_{ij}^* = \frac{x_i x_j}{1 + x_i x_j}, \qquad p_i^* = \frac{x_i}{1 + x_i}, \quad x_i = e^{-\theta_i^*}. \tag{4.81}$$

Since $(k_i)_{i=1}^m$ is a partial δ-tame degree sequence, Definition B.1 says that

$$\delta < p_{ij}^* < (1 - \delta). \tag{4.82}$$

From this it follows that

$$\frac{\delta}{1 - \delta} < x_i x_j < \frac{1 - \delta}{\delta}. \tag{4.83}$$

Using (4.83) for different indices i, j, k, we get

$$\left(\frac{\delta}{1 - \delta}\right)^2 < x_i^2 x_j x_k < \left(\frac{1 - \delta}{\delta}\right)^2. \tag{4.84}$$

Using again (4.83) for the indices j and k, we get

$$\left(\frac{\delta}{1 - \delta}\right)^{3/2} < x_i < \left(\frac{1 - \delta}{\delta}\right)^{3/2}. \tag{4.85}$$

Using (4.85) and $p_i^* = \frac{x_i}{1 + x_i} = \frac{1}{1 + \frac{1}{x_i}}$, we obtain that

$$\delta' < p_i^* < 1 - \delta' \tag{4.86}$$

with $\delta' = \frac{1}{1 + (\frac{1 - \delta}{\delta})^{3/2}}$. Note that $0 < \delta \leq \frac{1}{2}$ implies $0 < \delta' \leq \frac{1}{2}$.