EURANDOM PREPRINT SERIES

2020-013

September 10, 2020

A spectral signature of breaking of ensemble equivalence for constrained random graphs

P. Dionigi, D. Garlaschelli, F. den Hollander, M. Mandjes
ISSN 1389-2355

A spectral signature of breaking of ensemble equivalence for constrained random graphs

P. Dionigi [1]
D. Garlaschelli [2]
F. den Hollander [1]
M. Mandjes [4]

September 10, 2020

Abstract

For random systems subject to a constraint, the *microcanonical ensemble* requires the constraint to be met by every realisation ('hard constraint'), while the *canonical ensemble* requires the constraint to be met only on average ('soft constraint'). It is known that for random graphs subject to topological constraints *breaking of ensemble equivalence* may occur when the size of the graph tends to infinity, signalled by a non-vanishing *specific relative entropy* of the two ensembles. We investigate to what extent breaking of ensemble equivalence is manifested through the *largest eigenvalue* of the adjacency matrix of the graph. We consider two examples of constraints in the dense regime: (1) fix the degrees of the vertices (= the degree sequence); (2) fix the sum of the degrees of the vertices (= twice the number of edges). Example (1) imposes an extensive number of local constraints and is known to lead to breaking of ensemble equivalence. Example (2) imposes a single global constraint and is known to lead to ensemble equivalence. Our working hypothesis is that breaking of ensemble equivalence corresponds to a *non-vanishing difference of the expected values of the largest eigenvalue* under the two ensembles. We verify that, in the limit as the size of the graph tends to infinity, the difference between the expected values of the largest eigenvalue in the two ensembles does not vanish for (1) and vanishes for (2). A key tool in our analysis is a transfer method that uses relative entropy to determine whether probabilistic estimates can be carried over from the canonical ensemble to the microcanonical ensemble, and illustrates how breaking of ensemble equivalence may prevent this from being possible.

Key words. Constrained random graphs; canonical and micro-canonical ensembles; ensemble equivalence; relative entropy; adjacency matrix; largest eigenvalue; Erdős-Rényi random graph; regular random graph.

MSC2020. 60B20, 60C05, 60K35,

Acknowledgment. The work of PD, FdH and MM was supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation-grant NETWORKS-024.002.003, the work of DG by the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, The Netherlands).

[1]Mathematical Institute, Leiden University,
P.O. Box 9512, 2300 RA Leiden, The Netherlands

[2]Lorentz Institute for Theoretical Physics, Leiden University,
P.O. Box 9504, 2300 RA Leiden, The Netherlands

[3]IMT School for Advanced Studies,
Piazza S. Francesco 19, 55100 Lucca, Italy

[4]Korteweg-de Vries Institute, University of Amsterdam,
P.O. Box 94248, 1090 GE Amsterdam, The Netherlands

[*]E-mails: p.dionigi@math.leidenuniv.nl; garlaschelli@lorentz.leidenuniv.nl; denholla@math.leidenuniv.nl; M.R.H.Mandjes@uva.nl

1 Introduction

Background. Spectral properties of random graphs have been studied intensively in past years. A non-exhaustive list of key contributions is [3, 4, 7, 9–11, 14, 23]. Both the adjacency matrix and the Laplacian matrix have been popular. Scaling properties have been derived for the spectral distribution and the largest eigenvalue, with focus on central limit and large deviation behaviour. Most papers deal with random graphs whose edges are drawn *independently*. Different types of behaviour show up in the *dense regime* (where the number of edges is of the order of the square of the number of vertices), in the *sparse regime* (where the number of edges is of the order of the number of vertices), and in various regimes in between.

In this paper we focus on the *largest eigenvalue* of the non-normalized and non-centred adjacency matrix for a class of *constrained* random graphs. The largest eigenvalue is a highly non-linear functional of the entries of the adjacency matrix and therefore carries global information about the structure of the graph. Constraints are natural in the framework of statistical mechanics and *Gibbs ensembles*. Typically, they introduce a dependence between the edges that makes the spectral analysis challenging.

Breaking of ensemble equivalence (BEE). One of the interesting phenomena exhibited by certain classes of constrained random graphs is *Breaking of Ensemble Equivalence* (BEE). To understand what this is, we recall that in statistical physics different microscopic descriptions are available for a system that is subjected to a constraint, referred to as *Gibbs ensembles*. In the *microcanonical ensemble* the constraint is *hard*, i.e., each microscopic realisation of the system matches the constraint *exactly*. In the *canonical ensemble* the constraint is *soft*, i.e., is met only *on average*. For finite systems the two ensembles are clearly different, since they represent different physical situations (energetic isolation, respectively, thermal equilibrium with a reservoir at an appropriate temperture). However, the general belief is that this discrepancy vanishes in the thermodynamic limit. This expectation, referred to as *Equivalence of Ensembles* (EE), permeates the theory of Gibbs ensembles. It turns out that for many physical systems EE holds, but not for all. We refer to [21] for more background.

For interacting particle systems, EE has been studied at three different levels: thermodynamic, macrostate and measure. It was shown in [21] that these levels are equivalent. The present paper uses the measure level, which is based on the vanishing of the specific relative entropy. In [8, 12, 18, 19], the phenomenon of BEE was studied for random graphs subject to different types of constraints. It was found that, interestingly, BEE is the rule rather than the exception for constraints that are either *extensive* in the number of vertices or *frustrated*. An overview can be found in [16].

Spectral signature of BEE. Let A be the adjacency matrix of a random graph on n vertices, i.e., $A = \{a_{ij}\}_{i,j\in[n]}$ with $a_{ij} = 1_{\{i\sim j\}}$. Let λ_1 denote its largest eigenvalue. For $i \in [n]$, let k_i be the degree of vertex i. Write $\mathbb{E}_{\text{can}}$ and $\mathbb{E}_{\text{mic}}$ to denote expectation with respect to the canonical, respectively, microcanonical ensemble. Our *working hypothesis* is that $\mathbb{E}_{\text{can}}[\lambda_1] - \mathbb{E}_{\text{mic}}[\lambda_1]$ vanishes as $n \to \infty$ if and only if EE holds. We will verify this hypothesis for two specific examples of constraints in the dense regime: (1) fix the degrees of the vertices (= the degree sequence); (2) fix the sum of the degrees of the vertices (= twice the number of edges). Example (1) corresponds to the so-called *configuration model*. We consider the particular case where all the degrees are fixed at a common value $d(n)$, in which case the microcanonical ensemble becomes the $d(n)$-regular random graph. For both examples the canonical ensemble coincides with the Erdős-Rényi random graph with an appropriate retention probability [18].

For Erdős-Rényi random graphs, λ_1 was studied for various different regimes in [11,14]. We will need the following result.

Proposition 1.1. [11] *Let $G(n,p)$ be the Erdős-Rényi random graph on n vertices with retention probability $p \in (0,1)$. Write $\mathbb{P}_{G(n,p)}$ to denote the probability distribution of $G(n,p)$. Let λ_1 be the largest eigenvalue of the adjacency matrix of $G(n,p)$. Then, for every $p \in (0,1)$,*

$$\lim_{n\to\infty} \mathbb{P}_{G(n,p)} \left(\frac{\lambda_1 - [(n-1)\,p - (1-p)]}{\sqrt{2p(1-p)}} > x \right) = \mathrm{erf}(x), \qquad x \in \mathbb{R}, \tag{1.1}$$

where $\mathrm{erf}(\cdot)$ *is the standard error function.*

Note that Proposition 1.1 implies that $\lim_{n\to\infty}\{\mathbb{E}_{\mathrm{can}}[\lambda_1] - [(n-1)p - (1-p)]\} = 0$.

Our main result is the following theorem.

Theorem 1.2.
(1) *Let the constraint be $k_i = d(n)$, $i \in [n]$, with $nd(n)$ even, such that $\lim_{n\to\infty} d(n)/n = p$ for some $p \in (0,1)$. Then*

$$\lim_{n\to\infty} \left(\mathbb{E}_{\mathrm{can}}\left[\lambda_1\right] - \mathbb{E}_{\mathrm{mic}}\left[\lambda_1\right]\right) = 1 - p > 0. \tag{1.2}$$

(2) *Let the constraint be $\frac{1}{2}\sum_{i\in[n]} k_i = L(n)$ such that $\lim_{n\to\infty} 2L(n)/n^2 = p$ for some $p \in (0,1)$. Then*

$$\lim_{n\to\infty} \left(\mathbb{E}_{\mathrm{can}}\left[\lambda_1\right] - \mathbb{E}_{\mathrm{mic}}\left[\lambda_1\right]\right) = 0. \tag{1.3}$$

The restriction that $nd(n)$ is even is needed to make the constraint *graphical*, i.e., there exist simple graphs that meet the constraint. Furthermore, the required scaling for both (1) and (2) implies that we are in the dense regime. Note the remarkable fact that both $\mathbb{E}_{\mathrm{mic}}[\lambda_1]$ and $\mathbb{E}_{\mathrm{can}}[\lambda_1]$ tend to infinity as $n \to \infty$ while their difference remains bounded.

As shown in [12,18], BEE occurs in example (1) and EE in example (2), and hence Theorem 1.2 supports our working hypothesis that BEE corresponds to a non-vanishing difference of the expected largest eigenvalues under the two ensembles. We are presently unable to deal with the non-dense regime, but simulations carried out for example (1) in the sparse regime indicate that the working hypothesis holds up there as well.

Outline. The remainder of this paper is organised as follows. In Section 2 we recall the definition of the microcanonical and the canonical ensemble in the setting of constrained random graphs. Section 3 describes our main tool: a transfer method based on relative entropy, which carries over estimates on rare events from the canonical ensemble to the microcanonical ensemble, and describe its role in the general framework of BEE. In Section 4 we prove Theorem 1.2(1), in Section 5 we prove Theorem 1.2(2).

2 Gibbs ensembles for constrained random graphs

Consider the discrete probability space $(\mathbb{G}_n, \mathcal{B}, \mathbb{P})$, with $\mathbb{G}_n$ the set of all simple graphs on n vertices, $\mathcal{B} = 2^{\mathbb{G}_n}$ the power set of $\mathbb{G}_n$ consisting of all the subsets of $\mathbb{G}_n$, and $\mathbb{P}$ a probability measure.

A *constraint* is defined to be a vector-valued function $\vec{C}\colon \mathbb{G}_n \to \mathbb{R}^d$. Fix a value $\vec{C}^\star$ that is *graphical*, i.e., $\vec{C}(g) = \vec{C}^\star$ for at least one $g \in \mathbb{G}_n$. Define

$$\Gamma_{\vec{C}^\star} = \left\{ g \in \mathbb{G}_n \colon\ \vec{C}(g) = \vec{C}^\star \right\}. \tag{2.1}$$

The *microcanonical ensemble* is the uniform probability distribution on $\Gamma_{\vec{C}^\star}$:

$$\mathbb{P}_{\mathrm{mic}}(g) = \begin{cases} 1/|\Gamma_{\vec{C}^\star}|, & \text{if } g \in \Gamma_{\vec{C}^\star}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

The *canonical ensemble* is defined via the Hamiltonian $H(g, \vec{\theta}) = \langle \vec{\theta}, \vec{C}(g) \rangle$ (where $\langle \cdot, \cdot \rangle$ denotes the scalar product), namely,

$$\mathbb{P}_{\mathrm{can}}(g) = \frac{1}{Z_{\vec{\theta}^\star}}\, \mathrm{e}^{-H(g, \vec{\theta}^\star)}, \qquad g \in \mathbb{G}_n, \tag{2.3}$$

with the normalising factor

$$Z_{\vec{\theta}^\star} = \sum_{g \in \mathbb{G}_n} \mathrm{e}^{-H(g, \vec{\theta}^\star)}, \tag{2.4}$$

called the *partition function*. Note that both $\mathbb{P}_{\mathrm{mic}}$ and $\mathbb{P}_{\mathrm{can}}$ depend on n, but we suppress this dependence. The parameter $\vec{\theta}$ is set to the particular value $\vec{\theta}^\star$ that realises the constraint:

$$\mathbb{E}_{\mathrm{can}}[\vec{C}]\Big|_{\theta = \theta^\star} = \vec{C}^\star. \tag{2.5}$$

The constraint $\vec{C}^\star$, apart from being graphical, must also be *irreducible*, i.e., no subset of the constraint is redundant [19]. The latter is needed to make sure that $\vec{\theta}^\star$ is unique.

The relative entropy of $\mathbb{P}_{\mathrm{mic}}$ w.r.t. $\mathbb{P}_{\mathrm{can}}$ is defined as

$$\begin{aligned} S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}}) &= \sum_{g \in \mathbb{G}_n} \mathbb{P}_{\mathrm{mic}}(g) \log \frac{\mathbb{P}_{\mathrm{mic}}(g)}{\mathbb{P}_{\mathrm{can}}(g)} = \frac{1}{|\Gamma_{\vec{C}^\star}|} \sum_{g \in \Gamma_{\vec{C}^\star}} \log \frac{\mathbb{P}_{\mathrm{mic}}(g)}{\mathbb{P}_{\mathrm{can}}(g)} \\ &= -\frac{1}{|\Gamma_{\vec{C}^\star}|} \log \left[|\Gamma_{\vec{C}^\star}|\, \mathbb{P}_{\mathrm{can}}(g^\star) \right] \sum_{g \in \Gamma_{\vec{C}^\star}} 1 = -\log \mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star}) \end{aligned} \tag{2.6}$$

where we use the convention $0 \log 0 = 0$ and $g^\star$ is any graph in $\Gamma_{\vec{C}^\star}$. EE in the measure sense is defined as the vanishing of the relative entropy density, i.e., $\lim_{n\to\infty} n^{-1} S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}}) = 0$ (see [21]).

3 Transfer method

Comparison of the two ensembles. The additional freedom in the canonical ensemble implies that there is less dependence between the constituent random variables. In our case these random variables are the edges of the graph. For example, if the constraint is on the degree sequence, then the microcanonical ensemble corresponds to the *hard configuration model* (which in the case of constant degrees becomes the regular random graph), while the canonical ensemble corresponds to the *soft configuration model* (which is a special case of the generalized random graph model). The former requires an algorithm that randomly pairs half-edges and creates dependencies, while the latter is constructed via a sequence of independent random trials (which results in a multivariate Poisson-Binomial distribution for the degrees of the vertices [12]). Consequently, in the canonical ensemble calculations are carried out more easily. For example, a lot is known about spectral properties of adjacency matrices of random graphs under the canonical ensemble: because the entries of the adjacency matrix are independent, powerful tools from random matrix theory can be used. The challenge is to transfer properties from the canonical ensemble to the microcanonical ensemble without performing elaborate combinatorial computations.

Transfer principle. We start by noting that

$$\mathbb{P}_{\mathrm{mic}}\,(B) = \frac{\mathbb{P}_{\mathrm{can}}(B)}{\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star})}, \qquad B \subseteq \Gamma_{\vec{C}^\star}. \tag{3.1}$$

The latter holds because $g \mapsto H(g, \vec{\theta}^\star)$ and $g \mapsto \mathbb{P}_{\mathrm{can}}(g)$ are constant on the support of $\mathbb{P}_{\mathrm{mic}}$, i.e., all microcanonical realisations have the same probability under the canonical ensemble. In particular,

$$\mathbb{P}_{\mathrm{can}}(B \mid \Gamma_{\vec{C}^\star}) = \mathbb{P}_{\mathrm{mic}}(B), \qquad B \in \mathcal{B}. \tag{3.2}$$

Consequently, we have the following *transfer principle*.

Lemma 3.1. *For every $B \in \mathcal{B}$, if $\lim_{n\to\infty} \mathbb{P}_{\mathrm{can}}(B \mid \Gamma_{\vec{C}^\star}) = 0$, then $\lim_{n\to\infty} \mathbb{P}_{\mathrm{mic}}\,(B) = 0$.*

Distinguishing sets. Let $\mathcal{E}_{\mathcal{P}} \in \mathcal{B}$ be the subset of $\mathbb{G}_n$ consisting of all graphs that possess a certain property $\mathcal{P}$ we are interested in, i.e.,

$$\mathcal{E}_{\mathcal{P}} = \{g \in \mathbb{G}_n \colon\ g \text{ has property } \mathcal{P}\}\,. \tag{3.3}$$

Write $[\mathcal{E}_{\mathcal{P}}]^c$ to denote the complementary event. The crucial step in the argument underlying the transfer method is to find the right event $[\mathcal{E}_{\mathcal{P}}]^c$ that asymptotically implies failure of the property $\mathcal{P}$ that we want to transfer from the canonical ensemble to the microcanonical ensemble.

For the remainder, two events are important: $\mathcal{E}_{\mathcal{P}} \cap \Gamma_{\vec{C}^\star}$ and $[\mathcal{E}_{\mathcal{P}}]^c \cap \Gamma_{\vec{C}^\star}$. These represent the sets that are in the support of $\mathbb{P}_{\mathrm{mic}}$ for which property $\mathcal{P}$ holds and fails, respectively. Our focus will be on replacing $\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c \cap \Gamma_{\vec{C}^\star})$ by $\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c)$. Since $\mathbb{P}_{\mathrm{mic}}([\mathcal{E}_{\mathcal{P}}]^c \cap \Gamma_{\vec{C}^\star}) \le \mathbb{P}_{\mathrm{mic}}([\mathcal{E}_{\mathcal{P}}]^c)$, if we are able to prove that $\lim_{n\to\infty} \mathbb{P}_{\mathrm{mic}}([\mathcal{E}_{\mathcal{P}}]^c) = 0$, then we also have $\lim_{n\to\infty} \mathbb{P}_{\mathrm{mic}}([\mathcal{E}_{\mathcal{P}}]^c \cap \Gamma_{\vec{C}^\star}) = 0$, and we say that the property defining the set $\mathcal{E}_{\mathcal{P}}$ holds with high probability as $n \to \infty$. As explained in Section 2,

$$\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c \mid \Gamma_{\vec{C}^\star}) = \frac{\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c \cap \Gamma_{\vec{C}^\star})}{\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star})} \le \frac{\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c)}{\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star})}, \tag{3.4}$$

and so if we manage to prove that $\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c) = o(\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star}))$, then we obtain $\lim_{n\to\infty} \mathbb{P}_{\mathrm{mic}}([\mathcal{E}_{\mathcal{P}}]^c) = 0$.

Role of elative entropy and BEE. Equation (3.4) sets the scale at which the transfer method is effective. This scale is given by the denominator $\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star})$. Indeed, if it happens that $\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c) \neq o(\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star}))$, then (3.4) is ineffective. Importantly, from (2.6) we have

$$\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star}) = \mathrm{e}^{-S_n(\mathbb{P}_{\mathrm{mic}} \,\|\, \mathbb{P}_{\mathrm{can}})}. \tag{3.5}$$

This leads to an interesting connection between BEE and the transferability of a property $\mathcal{P}$: if $\mathbb{P}_{\text{can}}([\mathcal{E}_{\mathcal{P}}]^c) = o(\mathrm{e}^{-S_n(\mathbb{P}_{\text{mic}} \,\|\, \mathbb{P}_{\text{can}})})$, then $\lim_{n\to\infty} \mathbb{P}_{\text{mic}}([\mathcal{E}_{\mathcal{P}}]^c) = 0$. Since EE coincides with $S_n(\mathbb{P}_{\text{mic}} \,\|\, \mathbb{P}_{\text{can}}) = o(n)$, when the ensembles are equivalent it is easier to transfer. Our proof of Theorem 1.2(2) makes use of precisely this fact, and $\mathcal{P}$ is a certain concentration inequality for the largest eigenvalue of the adjacency matrix. By contrast, BEE makes the transfer more difficult. Indeed, Theorem 1.2(1) can be seen as an example where the same concentration inequality $\mathcal{P}$ cannot be transferred because the relative entropy is of higher order, namely, $S_n(\mathbb{P}_{\text{mic}} \,\|\, \mathbb{P}_{\text{can}}) = \Theta(n \log n)$ [12, 18].

Largest eigenvalue. We know from the results in [21] that whenever BEE occurs, there must exist quantities whose macrostate expectation is different under the two ensembles. Clearly, not all macroscopic quantities are good candidates for this. For instance, any linear combination of the constraints necessarily has the same expected value under the two ensembles. What we propose as a candidate is the largest eigenvalue of the adjacency matrix of the graph, because this is a highly nonlinear function of the imposed constraints and is sensitive to the global structure of the graph. In Sections 4–5 we will consider two examples of constraints in the dense regime: (1) fix the degrees of all the vertices; (2) fix the total number of edges. For the former we focus on the special case where all the degrees are equal.

4 Proof of Theorem 1.2(1): constraint on the degree sequence

The d-regular random graph with n vertices, written $G_{n,d}$, coincides with the microcanonical ensemble with constraint $\vec{C}^\star = (d, \dots, d)$ on the degree sequence, where we allow $d = d(n)$. The largest eigenvalue of the adjacency matrix of $G_{n,d}$ equals d, irrespective of n. The Erdős-Rényi random graph with retention probability $p = d/(n-1)$ coincides with the canonical ensemble with the same constraint.

In order to understand the difference in behaviour of λ_1 under the two ensembles, we need Proposition 1.1. Indeed, the result in (1.1), which actually holds for a generic symmetric random matrix subject to specific regularity conditions, can be interpreted as follows. The adjacency matrix A associated with $G(n,p)$ consists of elements $\{a_{ij}\}_{i,j\in[n]}$ that are identically 0 when $i = j$ and *Bernoulli random trials* ($a_{ij} = 0, 1$) with success probability p when $i \neq j$. The largest eigenvalue of the deterministic matrix $\bar{A}$ whose entries are $\bar{a}_{ij} = \mathbb{E}_{\text{can}}[a_{ij}] = p$ when $i \neq j$ and $\bar{a}_{ij} = 0$ when $i = j$ is given by $\lambda_1(\bar{A}) = (n-1)p$. Hence, compared to $\lambda_1(\bar{A})$, λ_1 is shifted by a random variable whose distribution is $\mathcal{N}(1-p, 2p(1-p))$. It is important to note that the parameters of this shift depend on p only, i.e., do not vary with n. In [11] it is shown that the error term in (1.1) is of order $O(1/\sqrt{n})$, relying on the fact that in the canonical ensemble the eigenvector $\vec{v}_1$ corresponding to the largest eigenvalue λ_1 is very close to the vector $\vec{\mathbb{1}} = (1, \dots, 1)$ (i.e., the norm of the projection of $\vec{v}_1$ onto $\vec{\mathbb{1}}$ is much larger than the norm of the projection of $\vec{v}_1$ onto the perpendicular space $\vec{\mathbb{1}}^{\perp}$).

It was shown in [12] that BEE holds in the so-called δ-tame regime, which correponds to $\delta \leq p = d/(n-1) \leq 1 - \delta$ with $\delta \in (0, \frac{1}{2}]$ (see [12, Definition 1.1]). Since

$$\lim_{n\to\infty} \left(\mathbb{E}_{\text{can}}\left[\lambda_1\right] - \mathbb{E}_{\text{mic}}\left[\lambda_1\right]\right) = 1 - p > 0, \tag{4.1}$$

we see that BEE comes with a non-vanishing difference of the expected largest eigenvalues under the two ensembles. This settles Theorem 1.2(1).

5 Proof of Theorem 1.2(2): constraint on the total number of edges

Consider the case where the constraint is on the total number of edges: $\vec{C}(g) = \vec{C}^\star = \binom{n}{2}p$ for some $p \in (0,1)$. Then the canonical ensemble is still the Erdős-Rényi random graph with parameter p. It was proved in [18] that the two ensembles are asymptotically equivalent on scale n. In particular, it was shown that $S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}}) = \log n + \Theta(1)$. The canonical probability of drawing a microcanonical realisation is given by (3.5):

$$\mathbb{P}_{\mathrm{can}}(\Gamma_{\vec{C}^\star}) = \mathrm{e}^{-S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}})} = \mathrm{e}^{-\log n + \Theta(1)} = \Theta(n^{-1}). \tag{5.1}$$

Together with (3.4), this tells us that if we can find an event $[\mathcal{E}_{\mathcal{P}}]^c$ such that $\mathbb{P}_{\mathrm{can}}([\mathcal{E}_{\mathcal{P}}]^c) = o(n^{-1})$, then we know that $\lim_{n\to\infty} \mathbb{P}_{\mathrm{mic}}(\mathcal{E}_{\mathcal{P}}) = 1$. Our goal is to adapt the method in [11] in order to find a probabilistic bound that is good enough for (3.4) to be applied in combination with (5.1).

In Section 5.1 we prove a concentration inequality for the degrees under the canonical ensemble (Lemma 5.1). In Section 5.2 we use this to prove a concentration inequality for a functional of the degrees that approximates the largest eigenvalue well (Lemma 5.3). In Section 5.3 we transfer the latter to the microcanonical ensemble (Lemma 5.4), and show that this leads to a negligible shift of the largest eigenvalue.

5.1 Concentration for the degrees under the canonical ensemble

For $i \neq j$, $\mathbb{E}_{\mathrm{can}}[a_{ij}] = p$ and $\mathrm{Var}_{\mathrm{can}}[a_{ij}] = p(1-p)$. In what follows we abbreviate $\mu = p$ and $\sigma^2 = p(1-p)$. We write $\vec{\mathbb{1}} = \vec{v}_1 + \vec{r}$ with $\vec{r} \in \vec{\mathbb{1}}^\perp$, $\langle \vec{v}_1, \vec{r}\rangle = 0$ and $A\vec{v}_1 = \lambda_1 \vec{v}_1$. Following the power method in [15], we define

$$\vec{K} = A\vec{\mathbb{1}} = A(\vec{v}_1 + \vec{r}) = \lambda_1 \vec{v}_1 + A\vec{r}, \tag{5.2}$$

which is the vector of row sums of the matrix A, i.e., the vector of degrees of the vertices (the degree sequence). Centering $\vec{K}$ by $\Theta\vec{\mathbb{1}}$ with $\Theta = \mathbb{E}[K_i] = (n-1)p$, we get

$$\vec{K} - \Theta\vec{\mathbb{1}} = (\lambda_1 - \Theta)\vec{v}_1 + (A\vec{r} - L\vec{r}). \tag{5.3}$$

Our key step is the following lemma.

Lemma 5.1. *With σ^2 denoting $p(1-p)$, there exist two constants $c_1, c_2 \in (0,\infty)$ such that*

$$\mathbb{P}_{\mathrm{can}}\left(\left|\sum_{i=1}^n (K_i - \Theta)^2 - \sigma^2 n(n-1)\right| \geq t\right) \leq c_2 \mathrm{e}^{-c_1 t/n^{3/2}}. \tag{5.4}$$

Proof. The term $\sum_{i=1}^n (K_i - \Theta)^2$ can be written as

$$\sum_{i=1}^n \left(\sum_{j=1}^n (a_{ij} - \mathbb{E}_{\mathrm{can}}[a_{ij}])\right)^2 = \sum_{i=1}^n \left(\sum_{j=1}^n b_{ij}\right)^2 = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n b_{ij} b_{ik}, \tag{5.5}$$

where

$$b_{ij} = a_{ij} - \mathbb{E}_{\mathrm{can}}[a_{ij}] = \begin{cases} a_{ij} - p, & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \tag{5.6}$$

are the centred entries of the adjacency matrix. Note that

$$\mathbb{E}_{\mathrm{can}}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} b_{ij}b_{ik}\right] = \sigma^2 n(n-1). \tag{5.7}$$

Straightforward counting shows that the sum in (5.5) contains $O(n^3)$ different terms. Let us represent $b_{ij} = b_{ji}$ by a variable X_α, $\alpha \in \left[\binom{n}{2}\right]$. Then (5.5) can be rewritten in the form

$$\sum_{\alpha,\beta\in\left[\binom{n}{2}\right]} h_{\alpha\beta} X_\alpha X_\beta, \tag{5.8}$$

which is the quadratic form of the matrix

$$H = \{h_{\alpha\beta}\}_{\alpha,\beta\in\left[\binom{n}{2}\right]}. \tag{5.9}$$

Because there is a one-to-one correspondence between the terms in (5.8) and (5.5), we can conclude that H has $O(n^3)$ entries, whose values are either 1 (off-diagonal) or 2 (diagonal). We can apply to (5.8) the *Hanson-Wright* inequality (see [13] or [1, Theorem 1.4, item 6]).

Theorem 5.2. *Let $X = (X_1, \dots X_N)$ be mean-zero square-integrable random variables taking values in $\mathbb{R}$, and let $\xi > 0$ be such that*

$$\|X\|_{\psi_2} = \inf\left\{t > 0 : \mathbb{E}\left[\exp\left(\|X\|_2^2/t^2\right)\right] \leq 2\right\} \leq \xi. \tag{5.10}$$

Let $H = (h_{\alpha\beta})_{\alpha\beta\in[N]}$ be a real symmetric matrix. Then the random variable

$$Y = \sum_{\alpha,\beta\in[N]} h_{\alpha\beta} X_\alpha X_\beta \tag{5.11}$$

satisfies

$$\mathbb{P}\left(|Y - \mathbb{E}[Y]| \geq t\right) \leq 2\exp\left(-\frac{1}{C}\min\left\{\frac{t^2}{\xi^4\|H\|_{\mathrm{HS}}^2}, \frac{t}{\xi^2\|H\|_{\ell_2^N\to\ell_2^N}}\right\}\right), \qquad t > 0, \tag{5.12}$$

where C is a suitable constant,

$$\|H\|_{\mathrm{HS}}^2 = \sum_{\alpha,\beta\in[N]} h_{\alpha\beta}^2 \tag{5.13}$$

is the Hilbert-Schmidt norm of H, and

$$\|H\|_{\ell_2^N\to\ell_2^N}^2 = \sup\left\{\sum_{\alpha,\beta\in[N]} h_{\alpha\beta}x_\alpha y_\beta : \sum_{\alpha\in[N]} x_\alpha^2 \leq 1, \sum_{\alpha\in[N]} y_\alpha^2 \leq 1\right\} \tag{5.14}$$

is the $\ell_2^N \to \ell_2^N$ norm of H.

In our setting, $N = \binom{n}{2}$. Since $|X_\alpha| < 1$, we have $\|X\|_{\psi_2} \leq 1/\log 2$, so that (5.10) applies with $\xi = 1/\log 2$. Since H has bounded entries, we have $\|H\|_{\mathrm{HS}}^2 = O(n^3)$. Moreover, by the Cauchy-Schwarz inequality we have

$$\|H\|_{\ell_2^N\to\ell_2^N}^2 = \sup\left\{\|Hx\|_2 \colon \|x\|_2 \leq 1\right\} = \|H\|_{\mathrm{op}}, \tag{5.15}$$

where the latter is the operator norm of H. But

$$\|H\|_{\mathrm{HS}}^2 = \mathrm{Tr}(H^\dagger H) \geq \lambda_{\max}(H^\dagger H) = \|H\|_{\mathrm{op}}^2, \tag{5.16}$$

and so the exponent in the right-hand side of (5.12) is bounded below by

$$\min\left\{\frac{t^2}{\xi^4 n^3}, \frac{t}{\xi^2 n^{3/2}}\right\} \geq \frac{c_3 t}{n^{3/2}}, \tag{5.17}$$

where c_3 is a suitable constant. Taking $c_1 \leq c_3/C$, with C the constant appearing in (5.12), we obtain (5.4). □

We end this section with an immediate consequence of Lemma 5.1. Picking $t = \sigma^2 n^2$ and using that, for appropriately chosen constants C_1, C_2, C_3, C_4,

$$\frac{\sigma^4 n^4}{\|H\|_{\mathrm{HS}}^2} \geq \frac{\sigma^4 n^4}{C_1 n^3} \geq C_2 n, \qquad \frac{\sigma^2 n^2}{L^2 \|H\|_{\mathrm{op}}} \geq \frac{\sigma^2 n^2}{C_3 \|H\|_{\mathrm{HS}}} \geq C_4 \sqrt{n}, \tag{5.18}$$

we find that there are constants $\tilde{c} \leq C_4/C$ and $\tilde{C}$ such that

$$\begin{aligned} &\mathbb{P}_{\mathrm{can}}\left(\left|\sum_{i=1}^n (K_i - \Theta)^2 - \sigma^2 n^2\right| \geq 2\sigma^2 n^2\right) \\ &\quad \leq 2\exp\left(-\frac{1}{C}\min\left\{\frac{4\sigma^4 n^4}{\|H\|_{HS}^2}, \frac{2\sigma^2 n^2}{\|H\|_{op}}\right\}\right) \leq \tilde{C} \mathrm{e}^{-\tilde{c}\sqrt{n}}. \end{aligned} \tag{5.19}$$

5.2 Concentration for the largest eigenvalue under the canonical ensemble

After applying A once to $\vec{\mathbb{1}}$, we must find a suitable normalization in order to isolate λ_1. This is given by

$$\frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} = \frac{\langle \vec{K}, \vec{K}\rangle}{\langle \vec{\mathbb{1}}, \vec{K}\rangle} = \frac{\|A\vec{\mathbb{1}}\|}{\langle \vec{\mathbb{1}}, A\vec{\mathbb{1}}\rangle} = \lambda_1 + \frac{\|A\vec{r}\|^2 - \lambda_1 \langle \vec{r}, A\vec{r}\rangle}{\sum_{i=1}^n K_i}. \tag{5.20}$$

In [11], it was shown that $\sum_{i=1}^n K_i^2 / \sum_{i=1}^n K_i$ approximates λ_1 with high probability, in the sense that for any $x > 0$,

$$\mathbb{P}_{\mathrm{can}}\left(\left|\frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \frac{\sum_{i=1}^n K_i}{n} - \frac{\sigma^2}{\mu}\right| \geq \frac{3\sigma^2 x}{\sqrt{n}}\right) \leq \frac{1}{x^2}, \tag{5.21}$$

which with the choice $x = \sqrt{n}$ leads to an upper bound of order $1/n$. As it turns out, however, in order to transfer the estimates to the microcanonical ensemble via (3.4), we need the upper bound to hold with probability $o(1/n)$. This result is covered by the following lemma.

Lemma 5.3. *Let $\vec{K}$ be as before, and $\mu = p$, $\sigma^2 = p(1-p)$. For any $\gamma > 0$ there exist $\gamma', \gamma_1, \gamma_2$ satisfying $c_1\gamma_1, \gamma_2 > 1$, with c_1 the constant in (5.4), such that*

$$\mathbb{P}_{\mathrm{can}}\left(\left|\frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \frac{\sum_{i=1}^n K_i}{n} - \frac{\sigma^2}{\mu}\right| \geq \frac{\gamma}{\sqrt{n}}\right) \leq \frac{\gamma'}{n^{\min\{c_1\gamma_1, \gamma_2\}}}. \tag{5.22}$$

Proof. First note that

$$\mathbb{E}_{\text{can}}\left[\frac{\sum_{i=1}^n K_i}{n}\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}_{\text{can}}[K_i] = (n-1)p = \Theta \tag{5.23}$$

and write

$$\frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \frac{\sum_{i=1}^n K_i}{n} = \frac{\sum_{i=1}^n (K_i - \Theta)^2}{\sum_{i=1}^n K_i} - \frac{\left(n^{-1}\sum_{i=1}^n K_i - \Theta\right)^2}{n^{-1}\sum_{i=1} K_i}. \tag{5.24}$$

To analyse the first ratio in (5.24), note that

$$\sum_{i=1}^n K_i = \sum_{i,j\in[n]} a_{ij} = 2\sum_{i,j\in[n],j>i} a_{ij}. \tag{5.25}$$

Applying Hoeffding's inequality (see e.g. [5,6]), we have

$$\mathbb{P}_{\text{can}}\left(\left|\sum_{i,j\in[n],j>i} a_{ij} - \frac{n(n-1)}{2}\mu\right| \geq t\right) \leq 2\exp\left(-\frac{4t^2}{n(n-1)}\right). \tag{5.26}$$

Take $t = n\sqrt{\gamma_2 \log n}$ in (5.26) with $\gamma_2 > 1$ and apply Lemma 5.1 with $t = n^{3/2}\gamma_1 \log n$, with $\gamma_1 c_1 > 1$ and c_1 the constant in the exponential bound of (5.4). Then, for some $\gamma > 0$,

$$\frac{\sum_{i=1}^n (K_i - \Theta)^2}{\sum_{i=1}^n K_i} \leq \frac{n(n-1)\sigma^2 + n^{3/2}\gamma_1 \log n}{n(n-1)\mu + n\sqrt{\gamma_2 \log n}} \leq \frac{\sigma^2}{\mu} + \frac{\gamma}{\sqrt{n}} \tag{5.27}$$

with probability at least $1 - 1/n^{\gamma_1 c_1} - 1/n^{\gamma_2}$. Similarly, the probability of

$$\frac{\sum_{i=1}^n (K_i - \Theta)^2}{\sum_{i=1}^n K_i} \geq \frac{\sigma^2}{\mu} - \frac{\gamma}{\sqrt{n}} \tag{5.28}$$

is bounded from below by $1 - 1/n^{\gamma_1 c_1} - 1/n^{\gamma_2}$. Hence

$$\mathbb{P}_{\text{can}}\left(\left|\frac{\sum_{i=1}^n (K_i - \Theta)^2}{\sum_{i=1}^n K_i} - \frac{\sigma^2}{\mu}\right| \geq \frac{\gamma}{\sqrt{n}}\right) \leq \frac{\gamma'}{n^{\min\{\gamma_1 c_1, \gamma_2\}}}. \tag{5.29}$$

To analyse the second ratio in (5.24), we write

$$\begin{aligned}\left(n^{-1}\sum_{i=1}^n K_i - \Theta\right)^2 &= \left(\frac{1}{n}\sum_{i,j\in[n]} (a_{ij} - \mathbb{E}_{\text{can}}[a_{ij}])\right)^2 \\ &= \frac{1}{n^2}\left(2\sum_{i,j\in[n],j>i} (a_{ij} - \mathbb{E}_{\text{can}}[a_{ij}])\right)^2,\end{aligned} \tag{5.30}$$

and apply Hoeffding's inequality with $t = O(n^2)$ twice. This gives

$$\mathbb{P}_{\text{can}}\left(\frac{\left(n^{-1}\sum_{i=1}^n K_i - \Theta\right)^2}{n^{-1}\sum_{i=1}^n K_i} > \frac{\tilde{\gamma}}{n}\right) \leq \tilde{\gamma}_2 e^{-\tilde{\gamma}_1 n^2}, \qquad \tilde{\gamma} > 0, \tag{5.31}$$

where $\tilde{\gamma}_2$ and $\tilde{\gamma}_1$ are suitable constants. Applying the union bound to the complementary events, we obtain (5.22). $\square$

5.3 Transfer to the microcanonical ensemble

Next we use the transfer method to pass the property characterised by the event in (5.22) to the microcanonical ensemble. Indeed, using the notation of Section 3, we identify

$$\left| \frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \frac{\sum_{i=1}^n K_i}{n} - \frac{\sigma^2}{\mu} \right| \geq \frac{\gamma}{\sqrt{n}} \tag{5.32}$$

as the event $\mathcal{E}_{\mathcal{P}}^c$, i.e., the set of graphs that do not possess the property that we would like to pass on. The fact that $\mathbb{P}_{\text{can}}\left(\mathcal{E}_{\mathcal{P}}^c\right)$ tends to zero faster than $\mathbb{P}_{\text{can}}(\Gamma_{\vec{C}^\star})$ (as $n \to \infty$, that is) tells us that also $\mathbb{P}_{\text{mic}}\left(\mathcal{E}_{\mathcal{P}}^c\right)$ tends to zero, and implies that

$$\lim_{n\to\infty} \mathbb{P}_{\text{mic}} \left(\left| \frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \frac{\sum_{i=1}^n K_i}{n} - \frac{\sigma^2}{\mu} \right| \leq \frac{\gamma}{\sqrt{n}} \right) = 1. \tag{5.33}$$

Thus, in the microcanonical ensemble $\sum_{i=1}^n K_i^2 / \sum_{i=1}^n K_i$ concentrates around $n^{-1}\sum_{i=1}^n K_i + \sigma^2/\mu$ with an error of order $1/\sqrt{n}$. However, we need to also see what $n^{-1}\sum_{i=1}^n K_i + \sigma^2/\mu$ is in the microcanonical ensemble. The term σ^2/μ, a constant equal to $1-p$, is in accordance with the constraint in the microcanonical ensemble. For the other term we have

$$n^{-1} \sum_{i=1}^n K_i = (n-1)p. \tag{5.34}$$

The two together give precisely the expected value in the canonical ensemble, as follows from Proposition 1.1. Hence we only need to show that $\sum_{i=1}^n K_i^2 / \sum_{i=1}^n K_i$ concentrates around λ_1 also in the microcanonical ensemble, for which we can once more use the transfer method.

Lemma 5.4. *For any $\eta > 0$, there exist ζ and Λ such that*

$$\mathbb{P}_{\text{can}} \left(\left| \frac{\sum_{i=1}^n K_i^2}{\sum_{i=1}^n K_i} - \lambda_1 \right| \geq \frac{\eta}{\sqrt{n}} \right) \leq \Lambda e^{-\zeta\sqrt{n}}. \tag{5.35}$$

Proof. We need to show that the last term in (5.20),

$$\frac{\|A\vec{r}\|^2 - \lambda_1 \langle \vec{r}, A\vec{r} \rangle}{\sum_{i=1}^n K_i}, \tag{5.36}$$

is small. First we show that $\|\vec{r}\|$ is bounded in probability. Indeed,

$$\sum_{i=1}^n (K_i - \Theta)^2 = (\lambda_1 - \Theta)^2 \|\vec{v_1}\|^2 + \|A\vec{r} - \Theta\vec{r}\|^2. \tag{5.37}$$

Since $(\lambda_1 - \Theta)^2 \|\vec{v_1}\|^2 \geq 0$, we have $\|A\vec{r} - \Theta\vec{r}\|^2 < \sum_{i=1}^n (K_i - \Theta)^2$. By the Courant-Fisher theorem [20, Theorem 1.3.2], we get that $\|A\vec{r} - \Theta\vec{r}\| \geq |\Theta - \lambda_2| \|\vec{r}\|$ (indeed, $|\Theta - \lambda_i| \geq |\Theta - \lambda_2|$ for $i > 2$). Next, we need a concentration inequality for λ_2. Use [2, Thm. 1] plus the fact that the largest eigenvalue of a centred matrix is of order $O(\sigma\sqrt{n})$ almost surely [17], [20, Thm. 2.3.24], [22, Thm. 1.3]. Again use the Courant-Fisher theorem to pass to the non-centred case [11, Lem. 1]. We find that, for any $\beta > 2$ and for $\tilde{n}$ large enough,

$$\mathbb{P}_{\text{can}} \left(\max_{i>1} |\lambda_i| \geq \beta\sigma\sqrt{n} \right) \leq 4e^{-\zeta_1 n}, \quad n > \tilde{n} \tag{5.38}$$

where ζ_1 is a suitable constant. Since $\max_{i>1}|\lambda_i| \geq \lambda_2 \geq 0$, we can bound $\lambda_2 \leq \beta\sigma\sqrt{n}$ with high probability. Using (5.19), we have

$$\mathbb{P}_{\text{can}}\left(\|\vec{r}\|^2 < \frac{(K_i-\Theta)^2}{(\Theta-\lambda_2)^2} < \frac{n^2\sigma^2}{(\mu n-\beta\sigma\sqrt{n})^2} < \frac{4\sigma^2}{\mu^2}\right) \geq 1-4e^{-\zeta_1 n}-\tilde{C}e^{-\tilde{c}\sqrt{n}}, \tag{5.39}$$

as a consequence of the union bound applied to the last term of $\mathbb{P}\left(\cap_n \mathcal{E}_n\right) = 1-\mathbb{P}(\cup_n[\mathcal{E}_n]^c)$, with $[\mathcal{E}_n]^c$ denoting the events described by Lemma 5.1 and (5.38). Thus, we have

$$\mathbb{P}_{\text{can}}\left(\|\vec{r}\|^2 \geq \frac{4\sigma^2}{\mu^2}\right) \leq \tilde{C}_1 e^{-\tilde{c}_1\sqrt{n}}, \tag{5.40}$$

where $\tilde{C}_1$ and $\tilde{c}_1$ are suitable constants.

All the other terms in (5.36) can be obtained by repeatedly using (5.40), (5.38) and (5.19). Note that in order to get (5.40) we have used both (5.38) and (5.19), and the events that these inequalities identify. Thus, using (5.38) twice, we obtain

$$\mathbb{P}_{\text{can}}\left(\|A\vec{r}\|^2 \leq \lambda_2^2\|\vec{r}\|^2 \leq \frac{50\,\sigma^4}{\mu^2}n\right) \geq 1-4e^{-\zeta_1 n}-\tilde{C}e^{-\tilde{c}\sqrt{n}}. \tag{5.41}$$

Therefore

$$\mathbb{P}_{\text{can}}\left(\|A\vec{r}\|^2 \geq \frac{50\,\sigma^4}{\mu^2}n\right) \leq \tilde{C}_2 e^{-\tilde{c}_2\sqrt{n}}, \tag{5.42}$$

where $\tilde{C}_2$ and $\tilde{c}_2$ are suitable constants. In the same way we can bound

$$|\langle\vec{r}, A\vec{r}\rangle| \leq \|\vec{r}\|\|A\vec{r}\|, \tag{5.43}$$

which yields

$$\mathbb{P}_{\text{can}}\left(|\langle\vec{r}, A\vec{r}\rangle| \geq \frac{2\sqrt{50}\sigma^3}{\mu^2}\sqrt{n}\right) \leq \tilde{C}_3 e^{-\tilde{c}_3\sqrt{n}}. \tag{5.44}$$

Now, using the trivial deterministic bound $\lambda_1 \leq \max_i \sum_j |a_{ij}| < n$ and Hoeffding's inequality on $\sum_{i=1}^n K_i = 2\sum_{j>i} a_{ij}$, we can conclude that, for any $\eta > 0$,

$$\mathbb{P}_{\text{can}}\left(\left|\frac{\|A\vec{r}\|^2-\lambda_1\langle\vec{r}, A\vec{r}\rangle}{\sum_{i=1}^n K_i}\right| \geq \frac{\eta}{\sqrt{n}}\right) \leq \Lambda e^{-\zeta\sqrt{n}}, \tag{5.45}$$

where ζ and Λ are suitable constants. Thus, recalling (5.20), we have settled (5.35). □

We thus find that the probability in the canonical ensemble of the event in (5.35) is $o\,(1/n)$. We can therefore conclude via (3.4) that the ratio $\sum_{i=1}^n K_i^2 / \sum_{i=1}^n K_i$ approximates λ_1 with a vanishing error also in the microcanonical ensemble. Together with the result of Lemma 5.3 and (5.34), we conclude that

$$\lim_{n\to\infty}\left(\mathbb{E}_{\text{can}}\left[\lambda_1\right]-\mathbb{E}_{\text{mic}}\left[\lambda_1\right]\right) = 0, \tag{5.46}$$

because of Proposition 1.1. This settles Theorem 1.2(2).

Remark 5.5. The constants in the right-hand side of (5.22) can be chosen freely. By Lemma 5.4, this means that for any choice of constraint for which $S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}}) = O(\log n)$ and the canonical ensemble is the Erdős-Rényi random graph, we have that λ_1 is close to $n^{-1}\sum_{i=1}^n K_i + \frac{\sigma^2}{\mu}$ in both ensembles. If the constraint does not prevent $\mathbb{E}_{\mathrm{mic}}[n^{-1}\sum_{i=1}^n K_i + \frac{\sigma^2}{\mu}]$ to take the value $(n-1)p + (1-p)$, then we have the same result as in Theorem 1.2(2), which supports the working hypothesis put forward in Section 1. Indeed, as shown in Section 3, $S_n(\mathbb{P}_{\mathrm{mic}} \parallel \mathbb{P}_{\mathrm{can}}) = o(n)$ is the condition for EE.

References

[1] R. Adamczak and P. Wolff. Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. *Probability Theory and Related Fields*, 162(3-4):531–586, 2015.

[2] N. Alon, M. Krivelevich, and Van H. Vu. On the concentration of eigenvalues of random symmetric matrices. *Israel Journal of Mathematics*, 131(1):259–267, December 2002.

[3] J. Alt, R. Ducatez, and A. Knowles. Extremal eigenvalues of critical Erdős-Rényi graphs. *arXiv:1905.03243 [math-ph]*, 2020.

[4] F. Benaych-Georges, C. Bordenave, and A. Knowles. Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs. *arXiv:1704.02953 [math]*, 2017.

[5] B. Bercu, B. Delyon, and E. Rio. *Concentration Inequalities for Sums and Martingales*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2015.

[6] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[7] A. Chakrabarty, R.S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs. *Random Matrices: Theory and Applications (to appear)*, 2019.

[8] F. den Hollander, M. Mandjes, A. Roccaverde, and N.J. Starreveld. Ensemble equivalence for dense graphs. *Electronic Journal of Probability*, 23:Paper no. 12, 1–26, 2018.

[9] S. Dhara and S. Sen. Large deviation for uniform graphs with given degrees. *arXiv:1904.07666 [math]*, 2020.

[10] T. Ding, X.and Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *The Annals of Applied Probability*, 20(6):2086–2117, 2010.

[11] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[12] D. Garlaschelli, F. den Hollander, and A. Roccaverde. Covariance structure behind breaking of ensemble equivalence in random graphs. *Journal of Statistical Physics*, 173(3-4):644–662, 2018.

[13] D.L. Hanson and F.T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 42(3):1079–1083, 1971.

[14] M. Krivelevich and B. Sudakov. The largest eigenvalue of sparse random graphs. *Combinatorics, Probability and Computing*, 12(01):61–72, 2003.

[15] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[16] A. Roccaverde. *Breaking of Ensemble Equivalence for Complex Networks*. PhD thesis, Leiden University, 2018.

[17] Alexander Soshnikov. Universality at the Edge of the Spectrum in Wigner Random Matrices. *Communications in Mathematical Physics*, 207(3):697–733, November 1999.

[18] T. Squartini, J. de Mol, F. den Hollander, and D. Garlaschelli. Breaking of ensemble equivalence in networks. *Physical Review Letters*, 115(26):268701, 2015.

[19] T. Squartini and D. Garlaschelli. Reconnecting statistical physics and combinatorics beyond ensemble equivalence. *arXiv:1710.11422 [cond-mat.stat-mech]*, 2020.

[20] T. Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 2012.

[21] H. Touchette. Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels. *Journal of Statistical Physics*, 159(5):987–1016, 2015.

[22] Van H. Vu. Spectral norm of random matrices. *Combinatorica*, 27(6):721–736, November 2007.

[23] Y. Zhu. A graphon approach to limiting spectral distributions of Wigner-type matrices. *Random Structures & Algorithms*, 56(1):251–279, 2020.