ANNEALED ISING MODEL ON CONFIGURATION MODELS

VAN HAO CAN, CRISTIAN GIARDINÀ, CLAUDIO GIBERTI, AND REMCO VAN DER HOFSTAD

ABSTRACT

In this paper, we study the annealed ferromagnetic Ising model on the configuration model. In an annealed system, we take the average on both sides of the ratio defining the Boltzmann-Gibbs measure of the Ising model. In the configuration model, the degrees are specified. Remarkably, when the degrees are deterministic, the critical value of the annealed Ising model is the same as that for the quenched Ising model. For independent and identically distributed (i.i.d.) degrees, instead, the annealed critical value is strictly smaller than that of the quenched Ising model. This identifies the degree structure of the underlying graph as the main driver for the critical value. Furthermore, in both contexts (deterministic or random degrees), we provide the variational expression for the annealed pressure. Interestingly, our rigorous results establish that only part of the heuristic conjectures in the physics literature were correct.

RÉSUMÉ

Dans ce travail, nous étudions le modèle d'Ising ferromagnétique "annealed" sur le modèle de configuration. Dans un système "annealed", nous prenons la moyenne des deux côtés du rapport définissant la mesure de Boltzmann-Gibbs du modèle d'Ising. Dans le modèle de configuration, les degrés sont spécifiés. Remarquablement, lorsque les degrés sont déterministes, la valeur critique du modèle d'Ising "annealed" est le même que celui du modèle d'Ising "quenched". Pour des degrés indépendants et distribués de manière identique (i.i.d.), au contraire, la valeur critique "annealed" est strictement inférieure à celle du modèle d'Ising "quenched". Cela identifie la structure en degrés du graphique sous-jacent comme le principal moteur de la valeur critique. Dans les deux contextes (degrés déterministes ou aléatoires), nous fournissons l'expression variationnelle de la pression "annealed". Fait intéressant, nos résultats rigoureux établissent que seule une partie des conjectures heuristiques de la littérature de physique étaient corrects.

Keywords: Annealed Ising model, configuration model, phase transition, critical temperature.

1. Motivation and results

The Ising model is a paradigmatic model for magnetism, see [28, 29] for an extensive historical account, and the books by Ellis [19], Bovier [3] and Contucci and Giardinà [8] for introductions to statistical mechanics from various perspectives. Recently, the Ising model has gained in popularity since it gives a very simple model for *consensus reaching* in populations. Indeed, when we assume that friends are more likely to have the same opinion rather than opposing ones, then we can use the Ising model on the friendship network to model which opinion will prevail in a population and under which circumstances. As a result, the model has become popular amongst economists and social scientists as well, and later found further applications in neuroscience, etc. See for example Contucci, Gallo and Menconi [7] for an example in social sciences, Kohring [27] for an application in social impact, Fraiman, Balenzuela, Foss and Chialvo [20] for an application to the brain and Bornholdt and Wagner [2] for an example in economics. Whereas the original Ising model was defined on a regular lattice, in recent applications more complex spatial structures (often in the form of a random graph) have been considered to model system inhomogeneities.

Being a paradigmatic model for random variables that are dependent, the Ising model on random graphs has also attracted considerable attention from the mathematics community, starting with Dembo and Montanari [10], see also [9, 12] for an overview of statistical mechanics models on random graphs, as well as [25, Chapter 5] for an extensive discussion of the recent results for the Ising model. Dembo and Montanari [10] prove the existence of the pressure per particle, and identify the critical inverse temperature, in the case where the degree distribution has finite variance. Dommers et al. [15] extend this also to power-law random graphs and identify the critical exponents in [16]. Central limit theorems were obtained in [21]. These results all apply in the general context of the *quenched* Ising model on locally-tree like random graphs, for which there is a quite complete understanding (unfortunately, the non-classical central limit theorem at the critical point is not proved in great generality).

Another important approach to the study of disordered systems is the *annealed* setting, where the pressure per particle is obtained as the logarithm of the averaged partition function (contrary to the quenched setting where one averages the logarithm of the partition function). Essentially the quenched setting concerns thermodynamic properties when a typical (fixed) realization of the disorder is considered, whereas in the annealed setting an average over the disorder distribution is taken so that the spins 'see' an average of the possible random graphs. For an extended discussion about quenched and annealed states, see for instance [21]. Often the annealed setting is easier to handle, and therefore is used as an approximation to the quenched setting. Examples include spin glasses [8], for instance the Sherrington-Kirkpatrick model where the annealed partition function becomes trivial due to the independence of the couplings, and random polymers, where the comparison between quenched and annealed is used to distinguish between relevant and irrelevant disorder [6].

For our setting of the Ising model on a random graph, the fact that the annealed setting is easier than the quenched is not obviously the case in the presence of non-independent edges, and also turns out to be wrong when doing the analysis. Essentially, the study of the quenched Ising model on locally-tree like random graphs is reduced to that of the

quenched Ising model on random trees, that is solved by a fixed point equation. Here we can crucially rely on the fact that the pressure per particle is a continuous functional in the local-weak topology. The situation is quite different if one considers the *annealed* Ising model, since in this case the contribution of graph realizations containing loops can not be ignored. On the contrary, since the partition function is an expectation of an exponentially large random variable, the contribution of atypical graph realizations (having extremely small probabilities) to the average of the partition function might be substantial and prevented, so far, the possibility of a general analysis of the annealed setting.

In the physics literature, the folklore is that the annealed Ising model on a random graph is equivalent to an inhomogeneous mean-field model where the coupling between the spins is proportional to their degrees, see Bianconi [1] and Dorogovtsev et al. [17]. However, this analogy often misses an exact proof, and this paper investigates this relation in detail for general configuration models.

From a mathematical perspective, the annealed Ising model could so far only be solved on two specific random graphs models: the generalized random graph [14], where edges are independent, and the random regular graph [5, 4], where the degrees are all equal (see also [22] where the configuration model with degrees that are either 1 or 2 was studied by mapping this model to a one-dimensional Ising model).

In this paper, we extend the analysis of the annealed Ising model to the *configuration model*, using ideas from, and extending work of Can [5, 4]. The main innovations in our paper are as follows:

  ▷ The model includes degree variability, and dependence between edges that is due to the fact that the edges need to realize a prescribed degree sequence;
  ▷ The model turns out to be exactly solvable and therefore we identify the annealed pressure as well as the critical value;
  ▷ The exact solution is obtained by a careful analysis of the annealed partition function with deterministic degrees using the mapping to a random matching problem derived by the first author in [5]. In the case of i.i.d. degrees, this is complemented by a large deviation analysis for the empirical degree distribution under the annealed Ising model;
  ▷ Surprisingly, the annealed critical value equals the quenched one for *deterministic* degrees, while it is strictly smaller for *independent and identically distributed* (i.i.d.) degrees.

Our solution shows that, as predicted by physicists, the annealed Ising model on the configuration model with prescribed degrees is indeed equivalent to a mean-field model, i.e., the magnetization can be computed in terms of a scalar quantity that solves a mean-field equation, see (2.23). However, the exact mean-field consistency equation turns out to be substantially different from the one predicted in [17]. Also, the prediction of the critical value following from that analysis in general turn out to be wrong.

1.1. Configuration model.

Fix an integer n that will be the number of vertices in the random graph and denote by $[n]$ the set $\{1, 2, \ldots, n\}$. Consider a sequence of degrees $\boldsymbol{d} = (d_i)_{i \in [n]}$. Without loss of generality, we assume that $d_j \geq 1$ for all $j \in [n]$, since when $d_j = 0$, vertex j is isolated and can be removed from the graph. Furthermore, we assume

that the total degree

$$
\ell_n = \sum_{j \in [n]} d_j \tag{1.1}
$$

is even. When ℓ_n is odd we increase the degree d_n by 1. For n large, this will not change the results and we will therefore ignore this effect.

We wish to construct a graph such that $\boldsymbol{d} = (d_i)_{i \in [n]}$ are the degrees of the n vertices. To construct the multigraph where vertex j has degree d_j for all $j \in [n]$, we have n separate vertices and incident to vertex j, we have d_j half-edges. Every half-edge needs to be connected to another half-edge to form an edge, and by forming all edges we build the graph. For this, the half-edges are numbered in an arbitrary order from 1 to ℓ_n. We start by randomly connecting the first half-edge with one of the $\ell_n - 1$ remaining half-edges, chosen uniformly at random. Once paired, two half-edges form a single edge of the multigraph, and the half-edges are removed from the list of half-edges that need to be paired. Hence, a half-edge can be seen as the left or the right half of an edge. We continue the procedure of randomly choosing and pairing the half-edges until all half-edges are connected, and call the resulting graph the *configuration model with degree sequence* $\boldsymbol{d}$, abbreviated as $\mathrm{CM}_n(\boldsymbol{d})$.

1.2. **Ising model.** We start by defining Ising models on finite graphs. Consider a graph sequence $(G_n)_{n \geq 1}$, where $G_n = (V_n, E_n)$, with vertex set $V_n = [n]$ and some random edge set E_n. To each vertex $i \in [n]$ we assign an Ising spin $\sigma_i = \pm 1$. A configuration of spins is denoted by $\sigma = \{\sigma_i : i \in [n]\}$. The ferromagnetic *Ising model on* G_n is then defined by the Boltzmann-Gibbs distribution

$$
\mu_n(\sigma) = \frac{1}{Z_n(\beta, B)} \exp\Big\{ \beta \sum_{(i,j) \in E_n} \sigma_i \sigma_j + B \sum_{i \in [n]} \sigma_i \Big\}. \tag{1.2}
$$

Here, $\beta \geq 0$ is the inverse temperature and B is the external magnetic field. The partition function $Z_n(\beta, B)$ is a normalization factor given by

$$
Z_n(\beta, B) = \sum_{\sigma \in \{-1,+1\}^n} \exp\Big\{ \beta \sum_{(i,j) \in E_n} \sigma_i \sigma_j + B \sum_{i \in [n]} \sigma_i \Big\}. \tag{1.3}
$$

We let $\langle \cdot \rangle_{\mu_n}$ denote the expectation with respect to the Ising measure, i.e., for every function $f : \{-1, +1\}^n \to \mathbb{R}$,

$$
\langle f \rangle_{\mu_n} = \sum_{\sigma \in \{-1,+1\}^n} f(\sigma) \mu_n(\sigma). \tag{1.4}
$$

The main quantity we shall study is the *pressure* per particle, which is defined as

$$
\psi_n(\beta, B) = \frac{1}{n} \log Z_n(\beta, B), \tag{1.5}
$$

in the thermodynamic limit of $n \to \infty$. It turns out that the pressure characterizes the behavior in the Ising model, and the phase transition and related quantities can be retrieved from it by taking appropriate derivatives with respect to the parameters B and β. For example, the *magnetization* $M_n(\beta, B)$ is given by

$$
M_n(\beta, B) = \Big\langle \frac{1}{n} \sum_{i \in [n]} \sigma_i \Big\rangle_{\mu_n} = \frac{\partial}{\partial B} \psi_n(\beta, B). \tag{1.6}
$$

4

When dealing with a random graph, such as $\mathrm{CM}_n(\boldsymbol{d})$, the above variables are all *random*. Then, the above corresponds to the so-called *random quenched* setting. However, we often wish to average out over the randomness of the graph. One way to do this, which is called the *averaged quenched setting*, is to simply take the expectation in (1.2). In this setting, when computing an expectation with respect to the Ising measure we deal with a ratio of random variables (since they are functions of the random graph), and both numerator and denominator grow exponentially in the large graph limit. Typically, such problems are difficult, as one still has to deal with the ratio of large random variables.

A common alternative approach is to take the average over the randomness both in the numerator as well as the denominator, giving rise to the *annealed* measure. In the annealed setting, the measure (1.2) is replaced by

$$(1.7) \qquad \mu_n^{\mathrm{an,d}}(\sigma) = \frac{1}{\mathbb{E}[Z_n(\beta,B)]} \mathbb{E}\Big[\exp\Big\{ \beta \sum_{(i,j)\in E_n} \sigma_i\sigma_j + B \sum_{i\in[n]} \sigma_i \Big\} \Big].$$

Here $\mathbb{E}$ is the expectation over the randomness of the graph. The superscript "d" is added to remember that we are dealing with the configuration model with *deterministic* degrees. Later (see Section 1.4) we shall also consider *random* degrees, and then $\mathbb{E}$ will denote the averages over all the sources of randomness of the graph.

The main quantity of interest is the *annealed pressure*, that is defined as

$$(1.8) \qquad \psi_n^{\mathrm{an,d}}(\beta,B) = \frac{1}{n} \log \mathbb{E}[Z_n(\beta,B)].$$

See [25, Chapter 5] for an extensive discussion of these different settings, as well as an overview of the results on the Ising model on random graphs.

We are interested in the thermodynamic limit of our models, i.e. in the behavior as $n \to \infty$. Define $\varphi^{\mathrm{an,d}}(\beta,B) := \lim_{n\to\infty} \psi_n^{\mathrm{an,d}}(\beta,B)$. For the quenched setting, the thermodynamic limit of the pressure was proved to exist, while for the annealed setting, we prove the existence of $\varphi^{\mathrm{an,d}}(\beta,B)$ in this paper. Being the limit of a sequence of convex functions both in β and in B, the infinite volume pressure $\varphi^{\mathrm{an,d}}(\beta,B)$ is also convex. Criticality of the model manifests itself in the behavior of the *annealed spontaneous magnetization*

$$(1.9) \qquad M^{\mathrm{an,d}}(\beta) := \lim_{B\searrow 0} \frac{1}{B}[\varphi^{\mathrm{an,d}}(\beta,B) - \varphi^{\mathrm{an,d}}(\beta,0)],$$

which is well-defined since convex functions always have right derivates. In terms of these quantities, the *critical inverse temperature* is defined as

$$(1.10) \qquad \beta_c^{\mathrm{an,d}} := \inf\{\beta > 0 \colon M^{\mathrm{an,d}}(\beta) > 0\}.$$

An identical characterization holds in the quenched setting, now in terms of the quenched spontaneous magnetization. Thus, depending on the setting, we can obtain the *quenched* and *annealed critical points* denoted by $\beta_c^{\mathrm{qu,d}}$ and $\beta_c^{\mathrm{an,d}}$, respectively. When $0 < \beta_c < \infty$, we say that the system undergoes a *phase transition* at $\beta = \beta_c$. We will refer to $\beta_c^{\mathrm{qu,d}}$ and $\beta_c^{\mathrm{an,d}}$ as *critical values*, to avoid writing 'critical inverse temperature' many times.

It is possible that the critical value β_c, at which the ferromagnetic phase transition – if any – occurs, is different for the quenched and annealed settings (it is not hard to prove that β_c is the same in the random and averaged quenched setting). However, the general belief in physics is that, for the ferromagnetic Ising model on random graphs, the annealed and quenched cases have the same critical exponents, and are thus in the same universality

class. In this paper, we study the annealed Ising model on the configuration model. We continue by stating our result, both in the case where the degrees are deterministic, as well as the case where they are i.i.d. (see Sections 1.3 and 1.4, respectively).

1.3. **Results for deterministic degrees.** We impose certain *regularity conditions* on the degree sequence $\boldsymbol{d}$. In order to state these assumptions, we introduce some notation. We denote the degree of a uniformly chosen vertex U_n in $[n]$ by $D_n = d_{U_n}$. The random variable D_n has distribution function F_n given by

$$(1.11) \qquad F_n(x) = \frac{1}{n} \sum_{j \in [n]} \mathbb{1}_{\{d_j \leq x\}},$$

which is the *empirical distribution of the degrees.* We assume that the vertex degrees satisfy the following *regularity conditions:*

Condition 1.1 (Regularity conditions for vertex degrees).
(a) Weak convergence of vertex degrees. *There exists a distribution function F such that, as $n \to \infty$,*

$$(1.12) \qquad D_n \xrightarrow{\mathcal{D}} D,$$

where D_n and D have distribution functions F_n and F, respectively, and $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. Further, we assume that $F(0) = 0$, i.e., $\mathbb{P}(D \geq 1) = 1$.
(b) Convergence of average vertex degrees. *As $n \to \infty$,*

$$(1.13) \qquad \mathbb{E}[D_n] \to \mathbb{E}[D] < \infty,$$

where D_n and D have distribution functions F_n and F from part (a), respectively.
(c) Convergence of second moment vertex degrees. *As $n \to \infty$,*

$$(1.14) \qquad \mathbb{E}[D_n^2] \to \mathbb{E}[D^2] \in (0, \infty],$$

where again D_n and D have distribution functions F_n and F from part (a), respectively.
(d) Bound on the maximum degree. *Let $d_{\max}^{(n)} = \max\{d_i : i \in [n]\}$ be the maximum degree of n vertices. As $n \to \infty$,*

$$(1.15) \qquad d_{\max}^{(n)} = o(n/\log n).$$

While Condition 1.1(c) is not necessary for the thermodynamic limits proved in this paper, the second moment of the degrees is needed to identify the critical value of the Ising model on the configuration model, which is why we have added it. Note that we allow for $\mathbb{E}[D^2] = \infty$, in which case the asymptotic degree distribution has an infinite second moment. Condition 1.1(c) rules out cases where $\mathbb{E}[D_n^2]$ oscillates, in which case we cannot define the critical value.

As for Condition 1.1(d), the bound on the maximal degree is quite weak, in that it will be true for most degree distributions, and is used to characterize the pressure per particle. See (2.10) and (2.14) below.

Two canonical examples of such degree sequences that satisfy Condition 1.1 are when we take $d_i = [1 - F]^{-1}(i/n)$ for $i \in [n]$, where F is the distribution function of an integer-valued random variable, and when $(d_i)_{i \in [n]}$ constitutes a *realization* of an i.i.d. sequence of random variables with distribution function F. In the latter case, beware that in the annealed measure in (1.7), we do *not* take the average with respect to the randomness of

the degrees. The convergence results in this paper are then meant to holds *in probability*. See [23, Chapter 7] for an extensive discussion of the configuration model and the Degree Regularity Condition 1.1.

The Ising model on the configuration model with deterministic degrees has been investigated in substantial detail in the quenched setting. One of the main results is that the thermodynamic limits of the pressure and magnetization exist, and that the Ising model has a phase transition with the quenched critical value given by

$$(1.16) \qquad \beta_c^{\mathrm{qu,d}} = \mathrm{atanh}(1/\nu),$$

where

$$(1.17) \qquad \nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}$$

equals the asymptotic expected forward degree in $\mathrm{CM}_n(\boldsymbol{d})$. Our main result for the annealed Ising model on the configuration model with deterministic degrees is the following theorem, which settles the same result in the annealed setting:

Theorem 1.2 (Thermodynamic limit of the annealed pressure: deterministic degrees)**.** *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *where the degree sequence satisfies Conditions 1.1(a)-(b) and (d). Then, for all $\beta \geq 0$ and $B \in \mathbb{R}$, the thermodynamic limit of the annealed pressure exists, i.e., as $n \to \infty$,*

$$(1.18) \qquad \psi_n^{\mathrm{an,d}}(\beta, B) \to \varphi^{\mathrm{an,d}}(\beta, B),$$

where $\varphi^{\mathrm{an,d}}(\beta, B)$ is defined in (2.17) below.

We next identify the critical value of the annealed Ising model on the configuration model:

Theorem 1.3 (Critical value of annealed Ising model: deterministic degrees)**.** *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *where the degree distribution satisfies Conditions 1.1(a)-(d). Then, the critical value $\beta_c^{\mathrm{an,d}}$ equals*

$$(1.19) \qquad \beta_c^{\mathrm{an,d}} = \beta_c^{\mathrm{qu,d}} = \mathrm{atanh}(1/\nu),$$

where ν is the expected forward degree in $\mathrm{CM}_n(\boldsymbol{d})$ *defined in (1.17), and D is the asymptotic degree whose existence is stated in Conditions 1.1(a).*

The above result is quite surprising, as it states that, when the degrees are fixed, the configuration model cannot really favorably arrange itself so as to decrease the critical value of the Ising model on it. Some insight on this result can be obtained by comparing to the corresponding findings for the Ising model on the Generalized Random Graph (GRG), where instead the annealed critical value is smaller that the quenched critical value [22]. The reason for this is to be found in the fact that under the annealed measure the GRG random graph re-arranges itself by increasing the typical total number of edges (see [13]). As a consequence, by the second Griffiths inequality, the critical temperature increases. A re-arrangement of the degree sequence to increase the number of edges is, of course, not possible in the configuration model with prescribed deterministic degrees.

1.4. **Results for i.i.d. degrees.** In this section, we continue with the behavior for i.i.d. degrees, where now we also average over the randomness in the degrees. We denote the finite-volume annealed pressure of the annealed Ising model on the configuration models with degrees that are i.i.d. copies of the random variable D by

$$(1.20) \qquad \psi_n^{\mathrm{an,D}}(\beta, B) = \frac{1}{n} \log \mathbb{E}[Z_n(\beta, B)].$$

The superscript "D" is added to remember that we are dealing with the configuration model with random degrees. We define $\varphi^{\mathrm{an,D}}(\beta, B) := \lim_{n \to \infty} \psi_n^{\mathrm{an,D}}(\beta, B)$. Provided that this limit is not ∞, the annealed spontaneous magnetization and the inverse critical temperature $M^{\mathrm{an,D}}(\beta)$ and $\beta_c^{\mathrm{an,D}}$ are defined in a similar way as in (1.9) and (1.10), respectively. We remark that the *quenched* critical inverse temperature $\beta_c^{\mathrm{qu,D}}$ is equal to $\beta_c^{\mathrm{an,d}}$, where the asymptotic degree distribution D in Condition 1.1(a) is the degree distribution.

We emphasize that the expectation $\mathbb{E}$ in (1.20) includes averaging over two sources of randomness: firstly, for a given realization of the degrees, we average over all possible random graphs that can be formed with that given degrees sequence; secondly, we average over the randomness of the degrees. It turns out that this has a rather dramatic effect, contrary to the setting of deterministic degrees, that is very much alike to the quenched setting. Indeed, when the degrees have sufficiently thick tails, the annealed pressure can even be infinite:

Proposition 1.4 (Infinite pressure for annealed Ising model with i.i.d. degrees). *The pressure of the annealed Ising model on the configuration model with i.i.d. degrees that are copies of the random variable D with distribution function F equals $\psi_n^{\mathrm{an,D}}(\beta, B) = \infty$ precisely when*

$$(1.21) \qquad \mathbb{E}[\mathrm{e}^{\beta D/2}] = \infty.$$

Proposition 1.4 implies that for power-law degrees, the pressure is simply equal to infinity. The boundary occurs when the degree distribution has an exponential tail, in which case, the pressure per particle is infinity when β is large, but not when it is small. In this case, it is natural to believe that the critical value $\beta_c^{\mathrm{an,D}}$ is strictly smaller than the value β where the pressure per particle explodes.

One could say that the anomaly that $\psi_n^{\mathrm{an,D}}(\beta, B) = \infty$ in Proposition 1.4 can only occur due to the lack of a single-edge constraint. Thus, it would be interesting to investigate the setting where multiple edges are merged and self-loops removed, which is sometimes called the *erased configuration model*. In this case, it is not hard to see that even though $\psi_n^{\mathrm{an,D}}(\beta, B) < \infty$ is always true, it also holds that $\psi_n^{\mathrm{an,D}}(\beta, B)$ tends to infinity with n.

From now on, we assume that β and D are fixed so that $\mathbb{E}[\mathrm{e}^{\beta D/2}] < \infty$. Then, the following theorem describes the thermodynamic limit of the pressure per particle for the annealed Ising model on the configuration model with i.i.d. degrees:

Theorem 1.5 (Thermodynamic limit of the annealed pressure: i.i.d. degrees). *Consider the configuration model where the degrees are i.i.d. copies of the random variable D with distribution function F. Then, for all β satisfying $\mathbb{E}[\mathrm{e}^{\beta D/2}] < \infty$ and for all $B \in \mathbb{R}$, the thermodynamic limit of the annealed pressure exists, i.e., as $n \to \infty$,*

$$(1.22) \qquad \psi_n^{\mathrm{an,D}}(\beta, B) \to \varphi^{\mathrm{an,D}}(\beta, B),$$

where $\varphi^{\mathrm{an,D}}(\beta, B)$ is defined in (3.8) below. Further, $\varphi^{\mathrm{an,D}}(\beta, B) > \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{p})$, where $\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{p})$ denotes the pressure of the configuration model with a deterministic degree sequence having asymptotic degree distribution $\boldsymbol{p} = (p_k)_{k \geq 1}$ with $p_k = \mathbb{P}(D = k)$.

Theorem 1.5 implies that it is favorable to increase the edge-density in the configuration model with i.i.d. degrees, and that the above effect is so pronounced that it changes the pressure per particle. We continue by studying the critical value $\beta_c^{\mathrm{an,D}}$, and show that this effect is so large that it actually changes the critical value:

Theorem 1.6 (Critical value of annealed Ising model: i.i.d. degrees). *Consider the configuration model where the degrees are i.i.d. copies of the random variable D with distribution function F. The critical inverse temperature $\beta_c^{\mathrm{an,D}}$ satisfies $\beta_c^{\mathrm{an,D}} \leq \bar{\beta}_c^{\mathrm{D}}$, where $\bar{\beta}_c^{\mathrm{D}}$ is the unique solution to*

$$\bar{\beta}_c^{\mathrm{D}} = \operatorname{atanh}(1/\nu(\boldsymbol{q}(\bar{\beta}_c^{\mathrm{D}}))), \tag{1.23}$$

where

$$\nu(\boldsymbol{q}) = \frac{\sum_{k \geq 1} k(k-1)q_k}{\sum_{k \geq 1} kq_k} \quad and \quad q_k(\beta) = \frac{p_k \cosh(\beta)^{k/2}}{\mathbb{E}[\cosh(\beta)^{D/2}]}, \quad for\ k \in \mathbb{N} \tag{1.24}$$

and $p_k = \mathbb{P}(D = k)$. Assume that $\mathbb{E}[e^{\bar{\beta}_c^{\mathrm{D}} D/2}] < \infty$. Then, the critical value $\beta_c^{\mathrm{an,D}}$ satisfies

$$\beta_c^{\mathrm{an,D}} < \beta_c^{\mathrm{qu,d}} = \operatorname{atanh}(1/\nu), \tag{1.25}$$

unless $D = r$ almost surely for some positive integer r.

The case where $D = r$ corresponds to random regular graphs, and is the only case where Theorems 1.2 and 1.5 actually agree, which explains why this case needs to be excluded in Theorem 1.6. The random regular graph case was also investigated by Dommers et al. in [22] for $r = 2$, by Can in [5, 4], and by Dembo, Montanari, Sly and Sun [11] (see also the remark below [11, Theorem 1], where it is mentioned that the Ising result also holds for odd degree).

The fact that $\beta_c^{\mathrm{an,D}} < \beta_c^{\mathrm{qu,d}}$ can be understood by noting that, by a correlation inequality,

$$\nu(\boldsymbol{q}(\beta)) \geq \nu. \tag{1.26}$$

Indeed, introducing the random variable $D^\star$ with law given by

$$\mathbb{P}(D^\star = k) = \frac{kp_k}{\mathbb{E}[D]}, \tag{1.27}$$

where $p_k = \mathbb{P}(D = k)$, one has

$$\nu(\boldsymbol{q}(\beta)) = \frac{\mathbb{E}[(D^\star - 1)\cosh(\beta)^{D^\star/2}]}{\mathbb{E}[\cosh(\beta)^{D^\star/2}]} \geq \mathbb{E}[D^\star - 1] = \nu, \tag{1.28}$$

with strict inequality unless $D^\star$ is constant or $\beta = 0$. Thus, by the first part of Theorem 1.6,

$$\beta_c^{\mathrm{an,D}} \leq \bar{\beta}_c^{\mathrm{D}} = \operatorname{atanh}(1/\nu(\boldsymbol{q}(\bar{\beta}_c^{\mathrm{D}}))) < \operatorname{atanh}(1/\nu) = \beta_c^{\mathrm{qu,d}}.$$

Remark 1.7 (Equality of $\beta_c^{\mathrm{an,D}}$ and $\bar{\beta}_c^{\mathrm{D}}$). We believe that $\beta_c^{\mathrm{an,D}}$ and $\bar{\beta}_c^{\mathrm{D}}$ are equal, see also Remark 3.4 below, where we explain this intuition more precisely. Unfortunately, due to insufficient analytical control of the variational formulas that define $\varphi^{\mathrm{an,D}}(\beta, B)$, we are unable to prove this.

Remark 1.8 (Specific degree distributions). In Section 3.4, we work out two special cases. For the CM with Poisson degrees with parameter λ, we have $\beta_c^{\mathrm{qu}} = \mathrm{atanh}(1/\lambda)$ for $\lambda > 1$. When $\lambda \leq 1$, there is no giant component [24], so that $\beta_c^{\mathrm{qu}} = \infty$. In the annealed setting, however, we instead obtain that

$$(1.29) \qquad \bar{\beta}_c^{\mathrm{D}} = -\log(2\lambda^2) + \log\left[1 + \sqrt{1 + 4\lambda^4} + \sqrt{2 + 2\sqrt{1 + 4\lambda^4}}\right] < \infty,$$

for all $\lambda > 0$, so that the annealed setting with Poisson i.i.d. degrees behaves markedly different from the quenched setting. In particular, even when $\lambda < 1$, the annealed Ising model still has a *finite* critical value, indicating that the graph, under the annealed Ising model, will have a giant component for *all* $\lambda > 0$.

A similar scenario appears for the CM with Geometric degrees with parameter p. In this case $\beta_c^{\mathrm{qu}} = \mathrm{atanh}(p/[2(1-p)])$, since the expected forward degree is $\nu = 2(1-p)/p$. Thus, when $p \geq \frac{2}{3}$ there is no giant component and $\beta_c^{\mathrm{qu}} = \infty$. On the other hand, in the annealed case with i.i.d. degrees we prove that

$$(1.30) \qquad \bar{\beta}_c^{\mathrm{D}} = \ln\left(x^\star(p)^2 + \sqrt{x^\star(p)^4 - 1}\right) < \infty,$$

where $x^\star(p)$ is the solution of a fourth order algebraic equation, see (3.91), that is larger than 1 for *all* $0 < p < 1$. Also in this case we have a *finite* annealed critical value for all values of the parameter p.

1.5. **Discussion.** In this section, we provide some discussion of our results and state open problems.

Quenched and annealed. For the random r-regular graph, not only do the critical values in the annealed and quenched settings agree, but even their pressures [5]. We believe however this to be exceptional. Generically, i.e. for all configuration models with non constant deterministic degrees, we proved in Theorem 1.3 that the critical values are the same, but we believe that the quenched pressure in [10, 15] and the annealed pressures in Theorem 1.2 are different, as the first involves the solution of an infinite dimensional variational problem for the so-called cavity field distribution, whereas the second involves the solution of a scalar variational problem (see Lemma 2.1). However, it is not clear to us how to prove such a result.

Deterministic degrees: several universality classes. For deterministic degrees, we expect that in the annealed setting the behavior around criticality is described by the degree distribution in a similar way as for the quenched case, as described in [16]. In particular, such a result would imply that in the case where the degrees have a bounded fourth moment, the critical exponents are equal to those for the Curie-Weiss model, while in the strongly inhomogeneous setting of power-law degrees with exponent $\tau \in (3,5)$, they depend on τ. We do not address this question in this paper.

Independent and identically distributed degrees: only one universality class. The situation is quite different in the case of i.i.d. degrees. We believe that $\beta_c^{\mathrm{an,D}} = \bar{\beta}_c^{\mathrm{D}}$ (recall Remark 1.7). When $\mathbb{E}[e^{\beta_c^{\mathrm{an,D}} D/2}] < \infty$, then $\boldsymbol{q}(\beta_c^{\mathrm{an,D}})$ in (1.24) has exponential tails, since $\cosh(\beta_c^{\mathrm{an,D}}) < e^{\beta_c^{\mathrm{an,D}}}$. Therefore, denoting $\boldsymbol{q}(\beta, B)$ to be the empirical degree distribution of the annealed Ising model on the configuration model with i.i.d. degrees,

10

one can expect that for $\beta > \beta_c^{\mathrm{an,D}}$ and $B > 0$ with $\beta - \beta_c^{\mathrm{an,D}}$ and B very small, also $\boldsymbol{q}(\beta, B)$ has exponential tails. As a result, power-law degree distributions cannot occur, which suggests that the critical exponents are all equal to those of the Curie-Weiss model. In this case, there exists only *one* universality class, compared to the several ones for the setting of deterministic degrees. It would be quite surprising should the universality classes for the two settings be so crucially different. See Remark 3.5 for more details.

Generalized random graph with random weights. The analysis developed for the configuration model with i.i.d. degrees can be extended to the generalized random graph model with i.i.d. weights. For generalized random graphs with deterministic weights, it is known that the value of the critical temperature where the annealed phase transition occurs is strictly larger than the value of the quenched critical temperature [22]. The additional randomness of the weights causes a further increase of the annealed critical temperature. As for the critical behavior, by the same argument of the previous item we expect the existence of a unique universality class for the generalized random graph with i.i.d. weights, contrary to the quenched case or the annealed case with deterministic weights that have been shown to have multiple universality classes [16, 14]. Thus adding randomness to the random graph seems to favor a higher degree of universality, in the sense that different universality classes merge together as a consequence of averaging.

Organization of the paper. We prove the results for deterministic degrees in Section 2. More precisely, we identify the partition function for deterministic degrees and prove Theorem 1.2 in Section 2.1, and identify the annealed critical value in Theorem 1.3 in Section 2.2. We prove the results for i.i.d. degrees in Section 3. More precisely, we show when the partition function is finite and prove Proposition 1.4 in Section 3.1, identify the partition function for i.i.d. degrees with finite support and prove Theorem 1.2 in Section 3.2, extend the analysis to infinite support i.i.d. degrees in Section 3.3, and identify the annealed critical value in Theorem 1.6 in Section 3.4.

2. Deterministic degrees: Proof of Theorems 1.2 and 1.3

In this section, we prove our results in Theorems 1.2 and 1.3 on the annealed Ising model on the configuration model with deterministic degrees satisfying Conditions 1.1(a)-(b) and (d). In Section 2.1, we derive the thermodynamic limit of the pressure per particle. In Section 2.2, we identify the critical value.

2.1. **Partition function for deterministic degrees: Proof of Theorem 1.2.** Consider the annealed Ising model on the configuration model $\mathrm{CM}_n(\boldsymbol{d})$. Assume that Conditions 1.1(a)-(b) and (d) hold. For $k \geq 1$, denote

$$(2.1) \qquad n_k = \#\{i \leq n : d_i = k\}, \qquad \text{and} \qquad p_k^{(n)} = \mathbb{P}(D_n = k) = \frac{n_k}{n}.$$

By Condition 1.1(a), $p_k^{(n)} \to p_k = \mathbb{P}(D = k)$. By symmetry, we only consider the case that $B \geq 0$.

We now rewrite the partition function, using ideas from Can [5]. By [5, (3.4)], the expectation of the partition function can be written as

$$(2.2) \qquad \mathbb{E}[Z_n(\beta, B)] = \mathrm{e}^{(\beta \mathbb{E}[D_n]/2 - B)n} \sum_{A \subseteq [n]} \mathrm{e}^{2B|A|} g_\beta(\ell_A, \ell_n),$$

where $|A|$ denotes the number of vertices in A,

$$(2.3) \qquad \ell_A = \sum_{i \in A} d_i$$

is the total degree of the vertices in A and

$$(2.4) \qquad g_\beta(k, m) = \mathbb{E}[e^{-2\beta X(k,m)}].$$

Here, for $k \in [m]$, $X(k, m)$ denotes the number of edges between $[k]$ and $[m] \setminus [k]$ in a configuration model with m vertices of degree 1 only (i.e., a random matching).

The proofs of Can [5, 4] are centered around a careful asymptotic analysis of the function $g_\beta(k, m)$ that we next explain in detail (see in particular [5, Lemma 2.1]). Can applies these ideas to the random regular graph, but also notices that the results hold more generally (see [5, Section 7]). In this paper, we extend the analysis so as to deal with the non-regular case, for which Can proves the existence of the annealed pressure in [5, Proposition 7.3], but does not analyze it further. Indeed, the sequence $g_\beta(k, m)$ satisfies

$$(2.5) \qquad \max_{0 \le k \le m} \left| \frac{\log g_\beta(k, m)}{m} - F_\beta(k/m) \right| = \frac{O(1)}{m},$$

where, denoting $x \wedge y = \min\{x, y\}$ for $x, y \in \mathbb{R}$,

$$(2.6) \qquad F_\beta(t) = \int_0^{t \wedge (1-t)} \log f_\beta(u) du,$$

and

$$(2.7) \qquad f_\beta(u) = \frac{e^{-2\beta}(1 - 2u) + \sqrt{1 + (e^{-4\beta} - 1)(1 - 2u)^2}}{2(1 - u)}.$$

Using these asymptotics, a variational formula for the pressure per particle can be obtained in the thermodynamic limit. We start by rewriting (2.2) as

$$(2.8) \qquad \mathbb{E}[Z_n(\beta, B)] = e^{(\beta \mathbb{E}[D_n]/2 - B)n} \sum_{(j_k) \le (n_k)} \prod_{k \ge 1} \binom{n_k}{j_k} e^{2B \sum_{k \ge 1} j_k} g_\beta\left(\sum_{k \ge 1} k j_k, \ell_n\right).$$

In the above formula, compared to (2.2), j_k is the number of vertices of degree k that are in A (so that indeed $j_k \le n_k$), and $\ell_A = \sum_{k \ge 1} k j_k$. The factor $\binom{n_k}{j_k}$ arises from the number of possible choices of j_k elements out of a total of n_k elements.

Then

$$\psi_n^{\mathrm{an,d}}(\beta, B) = \frac{\beta \mathbb{E}[D_n]}{2} - B$$
$$+ \frac{1}{n} \log \sum_{(j_k) \le (n_k)} \exp\left[\sum_{k \ge 1} \log\binom{n_k}{j_k} + 2B \sum_{k \ge 1} j_k + \log g_\beta\left(\sum_{k \ge 1} k j_k, \ell_n\right)\right].$$

We can bound the above sum from below by taking the maximum over all $(j_k) \le (n_k)$, and from above by the total number of (j_k) with $(j_k) \le (n_k)$ times the maximum. Note that this number is bounded from above by

$$(2.9) \qquad \prod_{k \ge 1} \left(1 + (n_k - 1)\mathbb{1}_{\{n_k \ge 1\}}\right).$$

12

Since $d_{\max}^{(n)} = o(n/\log n)$ by Condition 1.1(d), and since $n_k \leq n$, we can trivially upper bound

$$(2.10) \qquad \prod_{k \geq 1} \left(1 + (n_k - 1)\mathbb{1}_{\{n_k \geq 1\}}\right) \leq n^{d_{\max}^{(n)}} = \mathrm{e}^{d_{\max}^{(n)} \log n} = \mathrm{e}^{o(n)}.$$

Further, let us denote the proportion of vertices of degree k in our set as

$$(2.11) \qquad s_k = j_k/n_k \in [0, 1].$$

Then, the Stirling approximation $n! = \sqrt{2\pi n}\mathrm{e}^{-n}n^n(1 + O(1/n))$ gives the estimate

$$(2.12) \qquad \binom{n}{j} = \frac{n^n}{j^j(n-j)^{n-j}}\mathrm{e}^{O(\log n)} = \mathrm{e}^{nI(j/n)+O(\log n)},$$

where $I(0) = I(1) = 0$ and, for $t \in (0, 1)$,

$$(2.13) \qquad I(t) = -t \log t - (1 - t) \log(1 - t).$$

Using that $\sum_{k \geq 1} \log n_k \mathbb{1}_{\{n_k \geq 1\}} = o(n)$ when $d_{\max}^{(n)} = o(n/\log n)$, Condition 1.1(d) implies

$$\psi_n^{\mathrm{an,d}}(\beta, B) = \frac{\beta \mathbb{E}[D_n]}{2} + \max_{0 \leq s_k \leq 1}\left[\sum_{k \geq 1} p_k^{(n)} I(s_k) + B\left(2\sum_{k \geq 1} s_k p_k^{(n)} - 1\right)\right.$$

$$(2.14) \qquad\qquad \left. + \mathbb{E}[D_n]F_\beta\left(\frac{\sum_{k \geq 1} k p_k^{(n)} s_k}{\mathbb{E}[D_n]}\right)\right] + O\left(\frac{d_{\max}^{(n)} \log n}{n}\right),$$

where the maximum runs over $s_k = j_k/n_k$ and we have used (2.5). We next argue that we can replace the maximum over $s_k = j_k/n_k$ by the supremum over $s_k \in (0, 1)$. The upper bound holds trivially; for the lower bound, we will make use of the explicit solution that will be derived below, which satisfies that $s_k^\star \in (\varepsilon, 1 - \varepsilon)$ for all $k \geq 1$. See Lemma 2.1 below. Since the function $\sum_{k \geq 1} p_k^{(n)} I(s_k) + B\left(2\sum_{k \geq 1} s_k p_k^{(n)} - 1\right) + \mathbb{E}[D_n]F_\beta\left(\frac{\sum_{k \geq 1} k p_k^{(n)} s_k}{\mathbb{E}[D_n]}\right)$ is uniformly continuous for $s_k \in (\varepsilon, 1 - \varepsilon)$, this shows that the maximum will be close to the supremum over $s_k \in (0, 1)$.

Using that $p_k^{(n)} \to p_k$ by Condition 1.1(a) and $\mathbb{E}[D_n] \to \mathbb{E}[D]$ by Condition 1.1(b), we see that the function that is being maximized converges pointwise to

$$(2.15) \qquad G((s_k)_{k \geq 1}) = \sum_{k \geq 1} p_k I(s_k) + B\left(2\sum_{k \geq 1} s_k p_k - 1\right) + \mathbb{E}[D]F_\beta\left(\frac{\sum_{k \geq 1} k p_k s_k}{\mathbb{E}[D]}\right).$$

We thus arrive at

$$(2.16) \qquad \psi_n^{\mathrm{an,d}}(\beta, B) = \frac{\beta \mathbb{E}[D]}{2} + \sup_{s_k \in (0,1)} G((s_k)_{k \geq 1}) + o(1).$$

This proves Theorem 1.2, with

$$(2.17) \qquad \varphi^{\mathrm{an,d}}(\beta, B) = \frac{\beta \mathbb{E}[D]}{2} + \sup_{s_k \in (0,1)} G((s_k)_{k \geq 1}).$$

We first show that the maximizers of G are attainable. To do this, we will show G is a continuous function in an appropriate compact metric space. Let us define

$$\mathcal{A} = [0, 1]^{\mathbb{N}} = \{\boldsymbol{s} = (s_k)_{k \geq 1} \colon 0 \leq s_k \leq 1\}.$$

We define a metric on $\mathcal{A}$ associated with the degree distribution $(p_k)_{k\geq 1}$ as follows. For $\boldsymbol{s} = (s_k)_{k\geq 1}$ and $\boldsymbol{t} = (t_k)_{k\geq 1}$, define the distance

$$d(\boldsymbol{s}, \boldsymbol{t}) = \sum_{k\geq 1} kp_k|s_k - t_k| + \sum_{k\geq 1} 2^{-k}|s_k - t_k|\mathbb{1}_{\{p_k=0\}}.$$

Since $\sum_{k\geq 1} kp_k$ is finite, by standard arguments in functional analysis, we can show that $(\mathcal{A}, d)$ is a compact metric space. Indeed, $d(\boldsymbol{s}, \boldsymbol{t})$ metrizes the product topology [18] according to which $[0,1]^{\mathbb{N}}$ is compact by Tychonoff's theorem. We now show that G is continuous function on $(\mathcal{A}, d)$. Suppose that $\boldsymbol{s}^{(n)}$ converges to $\boldsymbol{s}$ in $(\mathcal{A}, d)$. We shall show $G(\boldsymbol{s}^{(n)}) \to G(\boldsymbol{s})$ as $n \to \infty$. By the definition of G and d,

$$|G(\boldsymbol{s}^{(n)}) - G(\boldsymbol{s})| \leq \sum_{k\geq 1} p_k|I(s_k^{(n)}) - I(s_k)| + 2B\sum_{k\geq 1} p_k|s_k^{(n)} - s_k|$$

$$+ \mathbb{E}[D]\Big|F_\beta\Big(\frac{\sum_{k\geq 1} kp_k s_k^{(n)}}{\mathbb{E}[D]}\Big) - F_\beta\Big(\frac{\sum_{k\geq 1} kp_k s_k}{\mathbb{E}[D]}\Big)\Big|$$

$$(2.18) \qquad \leq \sum_{k\geq 1} p_k|I(s_k^{(n)}) - I(s_k)| + \Big[2B + \sup_{t\in(0,1)} |F_\beta'(t)|\Big]d(\boldsymbol{s}^{(n)}, \boldsymbol{s}).$$

We notice that

$$(2.19) \qquad \sup_{t\in(0,1)} |F_\beta'(t)| = 2\beta,$$

since

$$(2.20) \qquad F_\beta'(t) = \begin{cases} \log f_\beta(t) & \text{if } t \leq 1/2, \\ -\log f_\beta(1-t) & \text{if } t \geq 1/2, \end{cases}$$

and $\mathrm{e}^{-2\beta} \leq f_\beta(t) \leq 1$ for all $t \in (0, \frac{1}{2})$. Therefore, the second term in (2.18) converges to 0 as $n \to \infty$. To show the convergence of the first term, we notice that for all $\ell \geq 1$,

$$(2.21) \qquad \sum_{k\geq 1} p_k|I(s_k^{(n)}) - I(s_k)| \leq \sum_{k\leq \ell} p_k|I(s_k^{(n)}) - I(s_k)| + 2\log 2\sum_{k\geq \ell} p_k,$$

since $0 \leq I(s) \leq \log 2$ for all $s \in [0,1]$. The first term converges to 0 as $n \to \infty$, since $I(\cdot)$ is a continuous function and $s_k^{(n)} \to s_k$ for all $k \leq \ell$; the second term converges to 0 as $\ell \to \infty$. Hence, we conclude that the first term in (2.18) converges to 0 as $n \to \infty$, and thus G is a continuous function on a compact space. Therefore, G attains maximizers at some points $\boldsymbol{s} \in \mathcal{A}$.

In the next lemma, we identify the solution of this optimization problem, and thus give a more explicit formula for $\varphi^{\mathrm{an,d}}(\beta, B)$. This in particular implies that $\varphi^{\mathrm{an,d}}(\beta, B)$ is well defined and thus completes the proof of Theorem 1.2. We remark that the functional to be optimized in (2.15) only depends on those components s_k whose index $k \in \mathbb{N}$ is such that the asymptotic degree distribution $p_k = \mathbb{P}(D = k)$ satisfies $p_k > 0$.

Lemma 2.1 (Solution of the optimization). *Let $r = d_{\min} = \min\{i \geq 1 : p_i > 0\}$ be the minimal asymptotic degree and let $\mathcal{S}_r := \{i \geq r : p_i > 0\}$. Let $B \geq 0$. The maximizer $(s_k^\star)_{k\geq 1}$ to (2.17) satisfies, for all $k \in \mathcal{S}_r$,*

$$(2.22) \qquad s_k^\star = \frac{1}{w^k \mathrm{e}^{-2B} + 1},$$

14

where, for $B > 0$, $w = w(\beta, B)$ is a solution in $(\mathrm{e}^{-2\beta}, 1)$ to

$$(2.23) \qquad \frac{1 - \mathrm{e}^{-2\beta} w}{1 + w^2 - 2\mathrm{e}^{-2\beta} w} = \mathbb{E}\left[\left(1 + w^{D^\star} \mathrm{e}^{-2B}\right)^{-1}\right],$$

where $D^\star$ is the random variable defined in (1.27).

Proof. We first show that if $(s_k^\star)_{k \geq 1}$ is a maximizer of G, then for all $k \in \mathcal{S}_r$,

    (a) $1/2 \leq s_k^\star < 1$,
    (b) $\partial G((s_k^\star)_{k \geq 1})/\partial s_k = 0$.

We prove (a) by contradiction. Assume that there is an index $\ell \in \mathcal{S}_r$ such that $s_\ell^\star < \frac{1}{2}$. To show that $(s_k^\star)_{k \geq 1}$ is not a maximizer, it suffices to prove that $G((u_k^\star)_{k \geq 1}) > G((s_k^\star)_{k \geq 1})$, with $u_k^\star = s_k^\star \vee (1 - s_k^\star)$ for $k \geq 1$, where $x \vee y = \max\{x, y\}$. We observe that $u_k^\star \geq \frac{1}{2}$ for all k, so $\sum_{k \geq 1} k p_k u_k^\star / \mathbb{E}[D] \geq \frac{1}{2}$. Recall that $F_\beta(t) = F_\beta(1 - t)$ for $t \in [0, 1]$, and F_β is decreasing in $(0, \frac{1}{2}]$ and increasing in $[\frac{1}{2}, 1)$, while

$$(2.24) \qquad \left|\frac{1}{\mathbb{E}[D]} \sum_{k \geq 1} k p_k u_k^\star - \frac{1}{2}\right| = \left|\frac{1}{\mathbb{E}[D]} \sum_{k \geq 1} k p_k (s_k^\star \vee (1 - s_k^\star) - \frac{1}{2})\right|$$

$$= \frac{1}{\mathbb{E}[D]} \sum_{k \geq 1} k p_k |s_k^\star - \frac{1}{2}| \geq \left|\frac{1}{\mathbb{E}[D]} \sum_{k \geq 1} k p_k (s_k^\star - \frac{1}{2})\right|,$$

so that $\sum_{k \geq 1} k p_k u_k^\star / \mathbb{E}[D]$ is further from $\frac{1}{2}$ than $\sum_{k \geq 1} k p_k s_k^\star / \mathbb{E}[D]$. Hence,

$$(2.25) \qquad F_\beta\left(\sum_{k \geq 1} k p_k u_k^\star / \mathbb{E}[D]\right) \geq F_\beta\left(\sum_{k \geq 1} k p_k s_k^\star / \mathbb{E}[D]\right).$$

Furthermore,

$$(2.26) \qquad B\left(2 \sum_{k \geq 1} p_k u_k^\star - 1\right) > B\left(2 \sum_{k \geq 1} p_k s_k^\star - 1\right).$$

Finally, $I(u_k^\star) = I(s_k^\star)$ for all $k \geq 1$. We conclude that $G((u_k^\star)_{k \geq 1}) > G((s_k^\star)_{k \geq 1})$, which is a contradiction with $(s_k^\star)_{k \geq 1}$ being a maximizer.

Now assume that there is an index $\ell \in \mathcal{S}_r$, such that $s_\ell^\star = 1$. We compute

$$(2.27) \qquad \frac{\partial G}{\partial s_k} = p_k\left(I'(s_k) + 2B + k F_\beta'\left(\frac{\sum_{j \geq 1} j p_j s_j}{\mathbb{E}[D]}\right)\right).$$

Therefore, $\partial G((s_k^\star)_{k \geq 1})/\partial s_\ell = -\infty$ for $s_\ell = 1$, since $I'(1) = -\infty$ and F_β' is a bounded function (see (2.19)). Hence, there exists a small positive constant ε, such that $G((u_k^\star)_{k \geq 1}) > G((s_k^\star)_{k \geq 1})$, with $u_k^\star = s_k^\star$ for $k \neq \ell$ and $u_\ell^\star = 1 - \varepsilon$. This completes the proof of (a).

The proof of (b) is similar. Assume that there exists a maximizer $(s_k^\star)_{k \geq 1}$ satisfying $\frac{\partial G((s_k^\star)_{k \geq 1})}{\partial s_\ell} \neq 0$ for some $\ell \in \mathcal{S}_r$. Thanks to part (a), $1/2 \leq s_\ell^\star < 1$. Hence, there exists $(u_k^\star)_{k \geq 1}$, with $u_k^\star = s_k^\star$ for $k \neq \ell$ and $u_\ell^\star \in (s_\ell^\star - \varepsilon, s_\ell^\star + \varepsilon)$ with ε small enough, such that $G((u_k^\star)_{k \geq 1}) > G((s_k^\star)_{k \geq 1})$, which again gives a contradiction. Thus, also part (b) follows.

By part (b), all maximizers are solutions to the systems of equations $\partial G/\partial s_k = 0$ for all $k \in \mathcal{S}_r$. By part (a), all maximizers $(s_k)_{k \geq 1}$ satisfy $s_k \geq 1/2$ for all $k \in \mathcal{S}_r$, so that also $\sum_{k \geq 1} k p_k s_k / \mathbb{E}[D] \geq 1/2$. Hence,

$$(2.28) \qquad F_\beta'\left(\frac{\sum_{k \geq 1} k p_k s_k}{\mathbb{E}[D]}\right) = -\log f_\beta\left(1 - \frac{\sum_{k \geq 1} k p_k s_k}{\mathbb{E}[D]}\right),$$

15

by (2.20). Combining this equation with (2.27) and (b), we conclude that the maximizers in (2.16) are solutions of

$$(2.29) \qquad I'(s_k) + 2B - k \log f_\beta \left(1 - \frac{\sum_{j \geq 1} j p_j s_j}{\mathbb{E}[D]}\right) = 0, \qquad k \in \mathcal{S}_r.$$

In (2.29), we multiply the equation for s_r with $r = d_{\min}$ by k and substract the equation for s_k with $k \in \mathcal{S}_r \setminus \{r\}$ multiplied by r, which gives

$$(2.30) \qquad k I'(s_r) - r I'(s_k) + 2(k - r)B = 0.$$

Note that

$$(2.31) \qquad I'(s) = \log\left(\frac{1-s}{s}\right).$$

Thus, we arrive at, now writing $s = s_r$,

$$(2.32) \qquad \left(\frac{1-s}{s}\right)^k = \left(\frac{1-s_k}{s_k}\right)^r e^{-2B(k-r)},$$

which is equivalent to

$$(2.33) \qquad \frac{1-s_k}{s_k} = \left(\frac{1-s}{s}\right)^{k/r} e^{2(k-r)B/r}.$$

We conclude that

$$(2.34) \qquad s_k = s_k(s) = \left[\left(e^{2B}\frac{1-s}{s}\right)^{k/r} e^{-2B} + 1\right]^{-1}.$$

Defining

$$(2.35) \qquad w = \left(e^{2B}\frac{1-s}{s}\right)^{1/r},$$

we thus arrive at

$$(2.36) \qquad s_k = \left[w^k e^{-2B} + 1\right]^{-1}.$$

This proves (2.22) for all $k \in \mathcal{S}_r$.

In terms of w and using (2.36), equation (2.29) with $k = r$ becomes

$$(2.37) \qquad \log(w^r e^{-2B}) + 2B - r \log f_\beta \left(1 - \frac{\sum_{j \geq 1} j p_j \left[w^j e^{-2B} + 1\right]^{-1}}{\mathbb{E}[D]}\right) = 0,$$

which is equivalent to

$$(2.38) \qquad w = f_\beta \left(1 - \mathbb{E}\left(\left[1 + w^{D^\star} e^{-2B}\right]^{-1}\right)\right),$$

where $D^\star$ is the size-biased distribution of D given by

$$(2.39) \qquad \mathbb{P}(D^\star = k) = \frac{k p_k}{\mathbb{E}[D]}.$$

Note that f_β takes values in $(e^{-2\beta}, 1]$, so $w \in (e^{-2\beta}, 1]$. We now compute the inverse function of f_β as

$$(2.40) \qquad f_\beta^{-1}(y) = \frac{y^2 - e^{-2\beta}y}{1 + y^2 - 2e^{-2\beta}y}.$$

Then (2.38) becomes

$$(2.41) \qquad \frac{1 - e^{-2\beta}w}{1 + w^2 - 2e^{-2\beta}w} = \mathbb{E}\left[\left(1 + w^{D^\star}e^{-2B}\right)^{-1}\right],$$

which is (2.23).

We can see that if $B > 0$, then $w = 1$ is not a solution to the above equation. Thus, we only need to consider the solution in $(e^{-2\beta}, 1)$. $\qquad\square$

Proof of Theorem 1.2. Due to Lemma 2.1, combined with (2.17), we see that $\varphi^{\text{an,d}}(\beta, B)$ is well defined, and thus (2.17) completes the proof. $\qquad\square$

Below, we will rely on the fact that if we know the solution of (2.23), then we can find the annealed pressure by (2.17) and (2.22). This will be used to identify the critical value.

2.2. Critical value for deterministic degrees: Proof of Theorem 1.3.
We aim to show that the critical inverse temperature is

$$(2.42) \qquad \beta_c^{\text{an,d}} = \beta_c^{\text{qu,d}} = \operatorname{atanh}(1/\nu) = \frac{1}{2}\log\left(\frac{\mathbb{E}[D^\star]}{\mathbb{E}[D^\star] - 2}\right),$$

where ν is defined in (1.17) and the fact that $\nu = \mathbb{E}[D^\star] - 1$ is used.

We have already shown that

$$\varphi^{\text{an,d}}(\beta, B) = \frac{\beta\mathbb{E}[D]}{2} + \sup_{s_k \in (0,1)} G((s_k)_{k \geq 1}).$$

By Lemma 2.1,

$$\mathcal{M}_G = \{(s_k)_{k \in \mathcal{S}_r} : 0 \leq s_k \leq 1, (s_k)_{k \in \mathcal{S}_r} \text{ is a maximizer of } G(\cdot)\}$$

$$(2.43) \qquad \subseteq \{(s_k(w, B))_{k \in \mathcal{S}_r} : e^{-2\beta} \leq w \leq 1, w \text{ is a solution to } (2.23)\},$$

where

$$(2.44) \qquad s_k(w, B) = \frac{e^{2B}}{e^{2B} + w^k}.$$

and we recall that $\mathcal{S}_r := \{i \geq r : p_i > 0\}$ with $r = d_{\min} = \min\{i \geq 1 : p_i > 0\}$ the minimal asymptotic degree.

Lemma 2.2 (Spontaneous magnetization). *The spontaneous magnetization defined in (1.9) satisfies that*

(a) $M^{\text{an,d}}(\beta) = 0$ *when* $\beta < \beta_c^{\text{qu,d}}$;
(b) $M^{\text{an,d}}(\beta) > 0$ *when* $\beta > \beta_c^{\text{qu,d}}$.

Proof of Lemma 2.2 We start with part (a). Assume that $\beta < \beta_c^{\text{qu,d}}$. Then

$$(2.45) \qquad c = e^{-2\beta} > e^{-2\beta_c^{\text{qu,d}}} = \frac{m - 2}{m}, \qquad \text{with} \quad m = \mathbb{E}[D^\star].$$

Define
$$a = \mathrm{e}^{2B}.$$

The equation (2.23) becomes

(2.46)
$$g(w) := \mathbb{E}\left[\frac{a}{a + w^{D^\star}}\right] - \frac{1 - cw}{1 + w^2 - 2cw} = 0.$$

We observe that for any $w \in (0, 1)$, the function $x \mapsto a/(a + w^x)$ is concave. Therefore, by Jensen's inequality,

(2.47)
$$g(w) \le h_a(w) := \frac{a}{a + w^m} - \frac{1 - cw}{1 + w^2 - 2cw}.$$

For $w \in (0, 1)$,

(2.48)
$$h_a(w) = 0 \quad \Leftrightarrow \quad \tilde{h}_a(w) := cw^m - w^{m-1} + aw - ac = 0.$$

We notice that $h_a(0) = 0, h_a'(0^+) < 0$ and $g(w) \le h_a(w)$. Hence,

$$
\begin{aligned}
\underline{w} \ &:= \ \inf\{w \in (\mathrm{e}^{-2\beta}, 1): g(w) = 0\} \\
&\ge \ \inf\{w \in (\mathrm{e}^{-2\beta}, 1): h_a(w) = 0\} = \inf\{w \in (\mathrm{e}^{-2\beta}, 1): \tilde{h}_a(w) = 0\} =: \underline{w}_a.
\end{aligned}
$$

We have

$$
\begin{aligned}
\tilde{h}_a'(w) \ &= \ cmw^{m-1} - (m-1)w^{m-2} + a \\
\tilde{h}_a''(w) \ &= \ (m-1)w^{m-3}(cmw - (m-2)).
\end{aligned}
$$

Since $c > (m-2)/m$, one has $(m-2)/(cm) \in (0, 1)$. Thus, $\tilde{h}_a''(v)$ changes its sign from negative to positive when w crosses $(m-2)/(cm)$. Hence, $\tilde{h}_a'(v)$ attains the minimum value at $(m-2)/(cm)$. Using $c > (m-2)/m$, we obtain that

(2.49)
$$\tilde{h}_a'\left(\tfrac{m-2}{cm}\right) = a - \left(\tfrac{m-2}{cm}\right)^{m-2}$$

is positive for any $a \ge 1$, therefore $\tilde{h}_a$ is an increasing function in $[0, 1]$, and $\tilde{h}_a(w) = 0$ has a unique solution, say $\underline{w}_a$. Moreover, the unique solution w of $\tilde{h}_1(w) = 0$ is 1. We can check that $\frac{\partial \tilde{h}_1}{\partial w}(1) = (mc - (m-2)) \ne 0$. Therefore, the implicit function theorem gives that $\underline{w}_a$ is a differentiable function in a, satisfying that $\underline{w}_a \to 1$ as $a \to 1$ and

(2.50)
$$\frac{\partial \underline{w}_a}{\partial a}\bigg|_{a=1} = -\frac{\frac{\partial \tilde{h}_a(w)}{\partial a}\big|_{(w,a)=(1,1)}}{\frac{\partial \tilde{h}_a(w)}{\partial w}\big|_{(w,a)=(1,1)}} = -\frac{1}{mc - (m-2)}.$$

Thus, $a \mapsto \underline{w}_a$ is differentiable in $a = 1$ with bounded derivative. Further, under the assumption that $\beta < \beta_c^{\mathrm{qu,d}}$, if $B = 0$ then the equation (2.23) has the unique solution $w = 1$ (this follows combining the fact that the unique solution $\underline{w}_1$ of $\tilde{h}_1(w) = 0$ is 1 and the fact that the solution w of (2.23) is bounded by $\underline{w}_1 \le w \le 1$). Therefore,

(2.51)
$$\varphi^{\mathrm{an,d}}(\beta, 0) = \beta \frac{\mathbb{E}(D)}{2} + G_0((s_k(1, 0))_{k \in \mathcal{S}_r}),$$

where $G_0(\cdot)$ denotes the function $G(\cdot)$ when $B = 0$. We remark that

$$
\begin{aligned}
M^{\mathrm{an,d}}(\beta) &= \lim_{B \searrow 0} \frac{\varphi^{\mathrm{an,d}}(\beta, B) - \varphi^{\mathrm{an,d}}(\beta, 0)}{B} \\
&= \lim_{B \searrow 0} \frac{G((s_k(w^\star, B))_{k \in \mathcal{S}_r}) - G_0((s_k(1,0))_{k \in \mathcal{S}_r})}{B} \\
&= \lim_{B \searrow 0} \frac{G((s_k(w^\star, B))_{k \in \mathcal{S}_r}) - G((s_k(1,B))_{k \in \mathcal{S}_r})}{B} \\
&\quad + \lim_{B \searrow 0} \frac{G((s_k(1,B))_{k \in \mathcal{S}_r})) - G_0((s_k(1,0))_{k \in \mathcal{S}_r})}{B},
\end{aligned}
$$
(2.52)

with $w^\star$ a solution of (2.23). Using the fact that $(s_k(1,0))_{k \in \mathcal{S}_r} = (1/2)_{k \in \mathcal{S}_r}$ is a stationary point of G_0, we can show that

$$
\begin{aligned}
&\lim_{B \searrow 0} \frac{G((s_k(1,B))_{k \in \mathcal{S}_r})) - G_0((s_k(1,0))_{k \in \mathcal{S}_r})}{B} \\
(2.53) \quad &= \frac{\partial G\big((1/2)_{k \in \mathcal{S}_r}\big)}{\partial B} \Big|_{B=0} + \sum_{k \in \mathcal{S}_r} \frac{\partial G\big((s_k)_{k \in \mathcal{S}_r}\big)}{\partial s_k} \Big|_{(s_k)_{k \in \mathcal{S}_r} = (1/2)_{k \in \mathcal{S}_r}, B=0} \times \frac{\partial s_k}{\partial B} \Big|_{B=0} = 0,
\end{aligned}
$$

where we recall that the function $G = G_{\beta,B}$ both explicitly depends on the external field B, as well as implicitly depends on B through the s_k variables. By the mean-value theorem, there is some $\tilde{w} \in (w^\star, 1)$, such that $G((s_k(w^\star, B))_{k \in \mathcal{S}_r}) - G((s_k(1,B))_{k \in \mathcal{S}_r}) = G'((s_k(\tilde{w}, B))_{k \in \mathcal{S}_r})(w^\star - 1)$. Therefore, using the facts that $\underline{w}_a \leq \underline{w} \leq w^\star \leq \tilde{w} \leq 1$ and that $\underline{w}_a \to 1$ as $a \searrow 1$ (or equivalently $B \searrow 0$),

$$
\begin{aligned}
\lim_{B \searrow 0} \frac{|G((s_k(w^\star, B))_{k \in \mathcal{S}_r}) - G((s_k(1,B))_{k \in \mathcal{S}_r})|}{B} &= \lim_{B \searrow 0} \frac{|G'((s_k(\tilde{w}, B))_{k \in \mathcal{S}_r})||w^\star - 1|}{B} \\
\leq \lim_{B \searrow 0} \frac{|G'((s_k(\tilde{w}, B))_{k \in \mathcal{S}_r})|(1 - \underline{w})}{B} &\leq \lim_{B \searrow 0} \frac{|G'((s_k(\tilde{w}, B))_{k \in \mathcal{S}_r})|(1 - \underline{w}_a)}{B}
\end{aligned}
$$
$$
(2.54) = \lim_{B \searrow 0} |G'((s_k(\tilde{w}, B))_{k \in \mathcal{S}_r})| \lim_{a \searrow 1} \frac{(1 - \underline{w}_a)}{a - 1} \lim_{B \searrow 0} \frac{\mathrm{e}^{2B} - 1}{B},
$$

by the chain rule. The right-hand side vanishes by (2.50), since the first term converges to 0 since $G'((s_k(1,0))_{k \in \mathcal{S}_r}) = 0$ and since the second and third terms are bounded. This, together with (2.52), shows that $M^{\mathrm{an,d}}(\beta) = 0$, which completes the proof of (a).

We now prove (b). Assume that $\beta > \beta_c^{\mathrm{qu,d}}$. We will prove that there exist positive constants $\theta = \theta(\beta)$ and $\delta = \delta(\beta)$, such that, for all $0 < B \leq \delta$,

$$
(2.55) \qquad \mathcal{M}_G \subseteq \{(s_k(w, B))_{k \in \mathcal{S}_r} : \mathrm{e}^{-2\beta} \leq w \leq 1 - \theta, w \text{ is a solution to } (2.23)\}.
$$

Let $w^\star \in (\mathrm{e}^{-2\beta}, 1)$ be any solution to (2.23). Then $(s_k(w^\star, B))_{k \in \mathcal{S}_r}$ is a stationary point of G, and thus

$$
(2.56) \qquad G'((s_k(w^\star, B)) = \sum_{k \in \mathcal{S}_r} \frac{\partial G\big((s_k(w, B))_{k \in \mathcal{S}_r}\big)}{\partial s_k} \Big|_{w = w^\star} \times s_k'(w^\star, B)) = 0.
$$

19

Moreover, we can compute that

$$G''((s_k(w^\star, B)) = \sum_{i,k \in \mathcal{S}_r} \frac{\partial^2 G\big((s_k(w, B))_{k \in \mathcal{S}_r}\big)}{\partial s_k \partial s_i} \bigg|_{w=w^\star} \times s_k'(w^\star, B) s_i'(w^\star, B)$$

$$+ \sum_{k \in \mathcal{S}_r} \frac{\partial G\big((s_k(w, B))_{k \in \mathcal{S}_r}\big)}{\partial s_k} \bigg|_{w=w^\star} \times s_k''(w^\star, B)$$

$$(2.57) \qquad = \sum_{i,k \in \mathcal{S}_r} \frac{\partial^2 G\big((s_k(w, B))_{k \in \mathcal{S}_r}\big)}{\partial s_k \partial s_i} \bigg|_{w=w^\star} \times s_k'(w^\star, B) s_i'(w^\star, B) =: H(w^\star, B),$$

with

$$H(w, B) = \sum_{k \in \mathcal{S}_r} p_k I''(s_k)(s_k'(w, B))^2 - \frac{f_\beta'\left(1 - \frac{\sum_{k \geq 1} k s_k p_k}{\mathbb{E}[D]}\right)}{f_\beta\left(1 - \frac{\sum_{k \geq 1} k s_k p_k}{\mathbb{E}[D]}\right)} \frac{1}{\mathbb{E}[D]} \left(\sum_{k \in \mathcal{S}_r} k p_k s_k'(w, B)\right)^2.$$

When $w \to 1$ and $B \to 0$ (or equivalently $a \to 1$), for all $k \in \mathcal{S}_r$,

$$(2.58) \qquad s_k(w, B) \to \tfrac{1}{2}, \quad s_k'(w, B) \to -\frac{k}{4}, \quad I''(s_k(w, B)) \to -4.$$

Moreover,

$$\frac{f_\beta'(1/2)}{f_\beta(1/2)} = 2(1 - \mathrm{e}^{-2\beta}).$$

Therefore, as $w \to 1$ and $B \to 0$,

$$H(w, B) \to H(1, 0) \;=\; -\frac{1}{4} \sum k^2 p_k + \frac{1 - \mathrm{e}^{-2\beta}}{8\mathbb{E}[D]} \left(\sum p_k k^2\right)^2$$

$$(2.59) \qquad\qquad > \; -\frac{1}{4}\mathbb{E}[D^2] + \frac{2\mathbb{E}[D]}{8\mathbb{E}[D^2]\mathbb{E}[D]}\mathbb{E}[D^2]^2 = 0,$$

since, for $\beta > \beta_c^{\mathrm{qu,d}}$,

$$\mathrm{e}^{-2\beta} < \mathrm{e}^{-2\beta_c^{\mathrm{qu,d}}} = 1 - \frac{2}{\mathbb{E}[D^\star]} = 1 - \frac{2\mathbb{E}[D]}{\mathbb{E}[D^2]}.$$

This implies that there exist positive constants θ and δ, such that $H(w, B) > 0$ for all $1 - \theta \leq w \leq 1$ and $0 \leq B \leq \delta$. Therefore, if $w^\star \in (1 - \theta, 1)$ is a solution to (2.23), then $G'((s_k(w^\star, B)) = 0$ and $G''((s_k(w^\star, B)) > 0$. Hence, $(s_k(w^\star, B))_{k \in \mathcal{S}_r}$ is not a maximizer. Combining this with (2.43), we get (2.55).

The above implies that if $0 < B \leq \delta$, for any maximizer $(s_k(w^\star, B))_{k \in \mathcal{S}_r}$, one has

$$\sum_{k \in \mathcal{S}_r} p_k s_k(w^\star) \;=\; \sum_{k \in \mathcal{S}_r} p_k \frac{a}{a + (w^\star)^k} = \mathbb{E}\left(\frac{a}{a + (w^\star)^{\mathrm{D}}}\right)$$

$$(2.60) \qquad\qquad \geq \; \mathbb{E}\left(\frac{1}{1 + (1 - \theta)^{\mathrm{D}}}\right) = \frac{1}{2} + \eta,$$

for some $\eta > 0$. This implies in particular that

$$(2.61) \qquad \mathcal{M}_G \subseteq \left\{(s_k)_{k \in \mathcal{S}_r} : \sum_{k \in \mathcal{S}_r} p_k s_k \geq \tfrac{1}{2} + \eta\right\}.$$

20

On the other hand, we observe that with $\sigma_+ = \{v\colon \sigma_v = +1\}$,

$$M_n^{\mathrm{an,d}}(\beta, B) := \sum_\sigma \left(\frac{2|\sigma_+|}{n} - 1\right)\mu_n^{\mathrm{an,d}}(\sigma)$$

$$\geq \quad \eta\mu_n^{\mathrm{an,d}}\left(\sigma\colon |\sigma_+| > n(1+\eta)/2\right) - \mu_n^{\mathrm{an,d}}\left(\sigma\colon |\sigma_+| \leq n(1+\eta)/2\right)$$

$$(2.62) \qquad = \quad \eta - (\eta + 1)\mu_n^{\mathrm{an,d}}\left(\sigma\colon |\sigma_+| \leq n(1+\eta)/2\right).$$

Further,

$$(2.63) \qquad \mu_n^{\mathrm{an,d}}\left(\sigma\colon |\sigma_+| \leq n(1+\eta)/2\right) = \frac{\mathbb{E}[Y_n(\beta, B)]}{\mathbb{E}[Z_n(\beta, B)]},$$

with, as in (1.3),

$$(2.64) \qquad Y_n(\beta, B) = \sum_{\frac{|\sigma_+|}{n} \leq \frac{1+\eta}{2}} \exp\left\{\beta \sum_{(i,j)\in E_n} \sigma_i\sigma_j + B\sum_{i\in[n]} \sigma_i\right\}.$$

Using the same argument as for the partition function, it follows that

$$\frac{1}{n}\log \mathbb{E}[Y_n(\beta, B)] = \max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}_\eta} G((s_k)_{k\in\mathcal{S}_r}) + o(1),$$

where

$$\mathcal{A}_\eta = \left\{(s_k)_{k\in\mathcal{S}_r} \in \mathcal{A}\colon \sum_{k\in\mathcal{S}_r} p_k s_k \leq \tfrac{1+\eta}{2}\right\}.$$

We conclude that, uniformly in $B > 0$
$$(2.65)$$
$$\limsup_{n\to\infty} M_n^{\mathrm{an,d}}(\beta, B) \geq \eta - (\eta+1)\lim_{n\to\infty}\exp\left[-n\left(\max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}} G((s_k)_{k\in\mathcal{S}_r}) - \max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}_\eta} G((s_k)_{k\in\mathcal{S}_r})\right)\right].$$

Since $\mathcal{A}_\eta$ is closed in $\mathcal{A}$, and $\mathcal{A}$ is a compact space, also $\mathcal{A}_\eta$ is a compact set. Hence, on $\mathcal{A}_\eta$ the continuous function G attains the maximal value at some point $\mathbf{s}_\eta \in \mathcal{A}_\eta$. By (2.61), all maximizers of G are not in $\mathcal{A}_\eta$. Therefore,
$$(2.66)$$
$$\max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}} G((s_k)_{k\in\mathcal{S}_r}) - \max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}_\eta} G((s_k)_{k\in\mathcal{S}_r}) = \max_{(s_k)_{k\in\mathcal{S}_r}\in\mathcal{A}} G((s_k)_{k\in\mathcal{S}_r}) - G(\mathbf{s}_\eta) = \varepsilon > 0,$$

for some ε. As a consequence, (2.65) implies that uniformly in $B > 0$

$$(2.67) \qquad \limsup_{n\to\infty} M_n^{\mathrm{an,d}}(\beta, B) \geq \eta > 0.$$

Since $(\psi_n^{\mathrm{an,d}}(\beta, B))_{n\geq 1}$ is a sequence of convex functions in $B \in \mathbb{R}$ converging to $\varphi^{\mathrm{an,d}}(\beta, B)$, we have (see [30, Proposition I.3.2])

$$(2.68) \qquad (D_B^+\varphi^{\mathrm{an,d}})(\beta, B) \geq \limsup_{n\to\infty}(D_B^+\psi_n^{\mathrm{an,d}})(\beta, B)$$

where $(D_B^+\varphi^{\mathrm{an,d}})(\beta, B)$ denotes the right derivative w.r.t. B, i.e.

$$(2.69) \qquad (D_B^+\varphi^{\mathrm{an,d}})(\beta, B) := \lim_{h\searrow 0} \frac{\varphi^{\mathrm{an,d}}(\beta, B+h) - \varphi^{\mathrm{an,d}}(\beta, B)}{h}.$$

On the other hand, we have differentiability of the pressure at finite volume $\psi_n^{\mathrm{an,d}}(\beta, B)$ and thus on the right hand side of (2.68) we can replace the right derivative D_B^+ with the derivative D_B, obtaining

$$(2.70) \qquad (D_B^+ \varphi^{\mathrm{an,d}})(\beta, B) \geq \limsup_{n \to \infty} (D_B \psi_n^{\mathrm{an,d}})(\beta, B) = \limsup_{n \to \infty} M_n^{\mathrm{an,d}}(\beta, B).$$

By (2.67) we thus have

$$(2.71) \qquad (D_B^+ \varphi^{\mathrm{an,d}})(\beta, B) \geq \eta > 0.$$

Now for a real convex function f we have continuity from above of the right derivative (see, e.g., [26, Theorem 1.1.7])

$$(2.72) \qquad \lim_{h \to 0+} D^+ f(x + h) = D^+ f(x).$$

As a consequence, taking the limit $B \searrow 0$ in (2.71), we conclude that if $\beta > \beta_c^{\mathrm{qu,d}}$ then

$$(2.73) \qquad M^{\mathrm{an,d}}(\beta) > 0.$$

This completes the proof of part (b).

$\square$

Proof of Theorem 1.3. The proof follows directly from Lemma 2.2. $\square$

3. Proofs for i.i.d. degrees

In this section, we investigate the annealed Ising model with i.i.d. degrees. In Section 3.1, we prove Proposition 1.4. In Section 3.2, we identify the partition function and prove Theorem 1.5 for degree distributions with finite supports. In Section 3.3, we extend the analysis to general degree distributions. In Section 3.4, we identify an upper bound for the annealed critical value for configuration model with i.i.d. degrees and prove Theorem 1.6.

3.1. Finiteness of pressure per particle for i.i.d. degrees: Proof of Proposition 1.4. We first note that

$$(3.1) \qquad \mathbb{E}[Z_n(\beta, B)] \geq \mathbb{E}\left[\exp\left\{ \beta \sum_{(i,j) \in E_n} 1 + B \sum_{i \in [n]} 1 \right\} \right] = \mathbb{E}\left[\exp\left\{ \beta |E_n| + nB \right\} \right]$$
$$= \mathrm{e}^{nB} \mathbb{E}\left[\mathrm{e}^{\beta \sum_{i \in [n]} D_i / 2} \right].$$

The latter is finite precisely when $\mathbb{E}[\mathrm{e}^{\beta D / 2}] < \infty$, as required.

Further, we can similarly bound from above

$$(3.2) \qquad \mathbb{E}[Z_n(\beta, B)] \leq 2^n \mathbb{E}\left[\exp\left\{ \beta \sum_{(i,j) \in E_n} 1 + |B| \sum_{i \in [n]} 1 \right\} \right]$$
$$= 2^n \mathrm{e}^{n|B|} \mathbb{E}\left[\exp\left\{ \beta |E_n| \right\} \right]$$
$$= 2^n \mathrm{e}^{n|B|} \mathbb{E}\left[\mathrm{e}^{\beta \sum_{i \in [n]} D_i / 2} \right].$$

Again, this is finite precisely when $\mathbb{E}[\mathrm{e}^{\beta D / 2}] < \infty$. The latter bound also implies that $\limsup_{n \to \infty} \psi_n^{\mathrm{an,D}}(\beta, B) < \infty$ when $\mathbb{E}[\mathrm{e}^{\beta D / 2}] < \infty$, which completes the proof. $\square$

3.2. **Identification of the partition function for i.i.d. degrees: Proof of Theorem 1.5 for finite supports.** In this section, we prove Theorem 1.5 for degree distributions having a finite support. We will extend the proof to general degree distributions satisfying $\mathbb{E}[\mathrm{e}^{\beta D/2}] < \infty$ below. Let us consider the annealed Ising model on the configuration model with i.i.d. sequence of degrees given by the distribution $\boldsymbol{p}$, i.e., the sequence $(D_i)_{i \in [n]}$ are i.i.d. random variables with distribution $\mathbb{P}(D = i) = p_i$, with $p_i = 0$ if $i > i_0$, for a given i_0 (finite support condition). For any probability distribution $\boldsymbol{q} = (q_k)_{k \geq 1}$, let $\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q})$ be the annealed pressure of the Ising model on the configuration model with degree distribution $\boldsymbol{q}$.

Let $\boldsymbol{P}^{(n)}$ be the empirical degree distribution, i.e.,

$$(3.3) \qquad P_k^{(n)} = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{D_i = k\}}.$$

Then

$$(3.4) \qquad \mathbb{E}\left[Z_n(\beta, B)\right] = \mathbb{E}\left(\mathrm{e}^{n \psi_n^{\mathrm{an,d}}(\beta, B; \boldsymbol{P}^{(n)})}\right).$$

We note that $\psi_n^{\mathrm{an,d}}(\beta, B; \boldsymbol{p}) = \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{p}) + o(1)$ uniformly in $\boldsymbol{p}$, since $\boldsymbol{p}$ only takes finitely many values. See in particular (2.14), which gives a uniform bound for $\boldsymbol{p}$ only taking finitely many values, and (2.16), which shows that the error in the continuum approximation of the values of $(s_k)_{k \in \mathcal{S}_r}$ over which we maximize is uniform as well. By Varadhan's Lemma,

$$(3.5) \qquad -\lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}\left[Z_n(\beta, B)\right] = \inf_{\boldsymbol{q}}[H(\boldsymbol{q} \mid \boldsymbol{p}) - \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q})],$$

where we define the relative entropy

$$(3.6) \qquad H(\boldsymbol{q} \mid \boldsymbol{p}) = \sum_{i \geq 1} q_i \log\left(q_i/p_i\right)$$

and we use that the functional $\boldsymbol{q} \mapsto \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q})$ is bounded and continuous, since, e.g.

$$(3.7) \qquad \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}) - \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}') \leq \sup_{(s_k)_{k \in \mathcal{S}_r}} [G_{\beta,B}((s_k)_{k \in \mathcal{S}_r}; \boldsymbol{q}) - G_{\beta,B}((s_k)_{k \in \mathcal{S}_r}; \boldsymbol{q}')],$$

where now we also make the dependence on the degree distribution explicit in the notation. Since $\boldsymbol{q} \mapsto G_{\beta,B}((s_k)_{k \in \mathcal{S}_r}; \boldsymbol{q})$ is uniformly continuous, the upper continuity of $\boldsymbol{q} \mapsto \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q})$ follows. The lower continuity can be proved in an identical manner. We conclude that

$$(3.8) \qquad \varphi^{\mathrm{an,D}}(\beta, B) = \sup_{\boldsymbol{q}}[\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}) - H(\boldsymbol{q} \mid \boldsymbol{p})].$$

This identifies the pressure of the annealed Ising model on the configuration model with i.i.d. degrees, and thus proves Theorem 1.5 for degree distributions having finite support. $\qquad \square$

3.3. **Identification of the partition function for i.i.d. degrees: Proof of Theorem 1.5 for infinite supports.** When $\mathbb{E}[\mathrm{e}^{\beta D/2}] = \infty$, we know that the pressure per particle is infinite. In this section, we investigate what happens otherwise, by extending the analysis in the previous section to such degree distributions with infinite support satisfying $\mathbb{E}[\mathrm{e}^{\beta D/2}] < \infty$. We proceed by upper and lower bounds.

Proof of lower bound in Theorem 1.5 for infinite supports. We start with the lower bound, which is the easiest. Fix k large, to be determined later on. We use that

$$\mathbb{E}\left[Z_n(\beta, B)\right] \geq \mathbb{E}\left(e^{n\psi_n^{\mathrm{an,d}}(\beta,B;\boldsymbol{P}^{(n)})}\mathbb{1}_{\{P_l^{(n)}=0\ \forall l\geq k\}}\right)$$

$$(3.9) \qquad = \left(\sum_{i=1}^{k-1} p_i\right)^n \mathbb{E}\left[e^{n\psi_n^{\mathrm{an,d}}(\beta,B;\boldsymbol{P}^{(n)})} \mid D_i < k\ \forall i \in [n]\right].$$

Fix $\varepsilon > 0$, and use that $\left(\sum_{i=1}^{k-1} p_i\right)^n \geq e^{-\varepsilon n}$ for $k \geq k(\varepsilon)$ sufficiently large. It thus suffices to investigate the conditional expectation. Here, we use that, conditionally on $D_i < k\ \forall i \in [n]$, the sequence $(D_i)_{i\in[n]}$ is i.i.d. with probability mass function $\tilde{p}_i = p_i/\sum_{l=1}^{k-1} p_l$, $i = 1, \ldots, k-1$. Thus, we can use the analysis in Section 3.2 to arrive at

$$(3.10) \qquad \liminf_{n\to\infty} \frac{1}{n}\log\mathbb{E}\left[Z_n(\beta,B)\right] \geq \sup_{\boldsymbol{q}}[\varphi^{\mathrm{an,d}}(\beta,B;\boldsymbol{q}) - H(\boldsymbol{q}\mid\widetilde{\boldsymbol{p}})]-\varepsilon.$$

Since

$$H(\boldsymbol{q}\mid\widetilde{\boldsymbol{p}}) = \sum_{i=1}^{k-1} q_i \log\left(q_i/\tilde{p}_i\right) = \sum_{i=1}^{k-1} q_i \log\left(q_i/p_i\right) - \log\left(\sum_{i=1}^{k-1} p_i\right)\sum_{i=1}^{k-1} q_i$$

$$(3.11) \qquad = \sum_{i=1}^{k-1} q_i \log\left(q_i/p_i\right) + o(1),$$

we arrive at

$$(3.12) \qquad \liminf_{k\to\infty}\liminf_{n\to\infty} \frac{1}{n}\log\mathbb{E}\left[Z_n(\beta,B)\right] \geq \sup_{\boldsymbol{q}}[\varphi^{\mathrm{an,d}}(\beta,B;\boldsymbol{q}) - H(\boldsymbol{q}\mid\boldsymbol{p})] - \varepsilon,$$

with $H(\boldsymbol{q}\mid\boldsymbol{p}) = \sum_{i=1}^{\infty} q_i \log\left(q_i/p_i\right)$ and the supremum is over all probability distributions on $\mathbb{N}$. This establishes the lower bound. $\square$

Proof of upper bound in Theorem 1.5 for infinite supports. The upper bound is substantially more involved, and will follow from three lemmas. We start by giving an alternative representation of the pressure per particle.

It has been shown in [5, (7.2) and (7.3)] that

$$(3.13) \qquad \mathbb{E}[Z_n(\beta, B)] = e^{O(1/n)} \sum_{j=0}^{n} \binom{n}{j} e^{B(2j-n)}\mathbb{E}\left(e^{\ell_n[\beta/2+F_\beta(\ell_j/\ell_n)]}\right),$$

where for all $j \leq n$,

$$(3.14) \qquad \ell_j = D_1 + \cdots + D_j$$

consists of the partial sums of the degrees. Our first lemma investigates some properties of the function F_β:

Lemma 3.1 (Properties of F_β). *Let a, b, c, d be non-negative real numbers satisfying $a \leq b, c \leq d$ and $b - a \leq d - c$. Then*

$$(3.15) \qquad dF_\beta(c/d) - bF_\beta(a/b) \leq 0.$$

24

Proof. Recall that F_β satisfies

$$F_\beta(t) = F_\beta(1-t) \leq 0 \ \forall t \in [0,1], \tag{3.16}$$

$$F'_\beta(t) \leq 0 \ \forall t \in [0, \tfrac{1}{2}], \qquad F''_\beta(t) \geq 0 \ \forall t \in [0,1]. \tag{3.17}$$

We notice that it is sufficient to prove the lemma for the case $\frac{a}{b}, \frac{c}{d} \leq \frac{1}{2}$. Indeed, let $\bar{a} = b - a$ and $\bar{c} = d - c$. Then $\bar{a} \leq \bar{c}$, and thus $(a \wedge \bar{a}) \leq b, (c \wedge \bar{c}) \leq d$. By (3.16), $F(a/b) = F((a \wedge \bar{a})/b)$ and $F(c/d) = F((c \wedge \bar{c})/d)$. Hence, we only need to prove (3.15) for the case that $\frac{a}{b}, \frac{c}{d} \leq \frac{1}{2}$.

Assume that $\frac{a}{b}, \frac{c}{d} \leq \frac{1}{2}$. We show that for all $t \in [0,1]$,

$$x(t) = F_\beta(t) - tF'_\beta(t) \leq 0. \tag{3.18}$$

Indeed, using (3.16) and (3.17), $x'(t) = -tF''_\beta(t) \leq 0$ for all $t \in [0,1]$. Thus $x(t) \leq x(0) = 0$ for all $t \in [0,1]$.

Since $F''_\beta(t) \geq 0$, by a Taylor expansion, we obtain that $F_\beta(x) - F_\beta(y) \leq F'_\beta(x)(x-y)$ for all $0 \leq x, y \leq 1$. Hence,

$$dF_\beta(c/d) - bF_\beta(a/b) = (d-b)F_\beta(c/d) + b(F_\beta(c/d) - F_\beta(a/b))$$

$$\leq (d-b)F_\beta(c/d) + bF'_\beta(c/d)\left(\frac{c}{d} - \frac{a}{b}\right). \tag{3.19}$$

Observe that, since $a \leq c$,

$$b\left(\frac{c}{d} - \frac{a}{b}\right) = \frac{cb - ad}{d} \geq \frac{cb - cd}{d} = -(d-b)\frac{c}{d}. \tag{3.20}$$

Further $F'_\beta(c/d) \leq 0$, since $c/d \leq 1/2$. Hence, using the last two inequalities, we get

$$dF_\beta(c/d) - bF_\beta(a/b) \leq (d-b)\Big(F_\beta(c/d) - (c/d)F'_\beta(c/d)\Big) \leq 0, \tag{3.21}$$

by (3.18), as required. $\qquad\square$

We use Lemma 3.1 to truncate the degree distribution, as we explain now. For any $k \geq 1$, we consider a truncated sequence $(D_i^{(k)})_{i \geq 1}$ defined by $D_i^{(k)} = D_i \wedge k$. We define correspondingly the partition function $Z_n^{(k)}(\beta, B)$ and $\ell_j^{(k)} = D_1^{(k)} + \cdots + D_j^{(k)}$ for $j \leq n$. Then we also have

$$\mathbb{E}[Z_n^{(k)}(\beta, B)] = e^{O(1/n)} \sum_{j=0}^{n} \binom{n}{j} e^{B(2j-n)} \mathbb{E}\left(e^{\ell_n^{(k)}[\beta/2 + F_\beta(\ell_j^{(k)}/\ell_n^{(k)})]}\right). \tag{3.22}$$

Moreover, it is clear that for all $j \leq n$,

$$\ell_j^{(k)} \leq \ell_n^{(k)}, \ell_j \leq \ell_n, \quad \text{and} \quad \ell_n^{(k)} - \ell_j^{(k)} = \sum_{i=j+1}^{n}(D_i \wedge k) \leq \sum_{i=j+1}^{n} D_i = \ell_n - \ell_j.$$

Therefore, applying Lemma 3.1, we obtain that, for all $j \leq n$,

$$\ell_n F_\beta(\ell_j/\ell_n) \leq \ell_n^{(k)} F_\beta(\ell_j^{(k)}/\ell_n^{(k)}),$$

which implies, using (3.13), that

$$\mathbb{E}[Z_n(\beta, B)] \leq e^{O(1/n)} \sum_{j=0}^{n} \binom{n}{j} e^{B(2j-n)} \mathbb{E}\left(e^{\ell_n^{(k)}[\beta/2 + F_\beta(\ell_j^{(k)}/\ell_n^{(k)})]} e^{\beta(\ell_n - \ell_n^{(k)})/2}\right). \tag{3.23}$$

Define

$$(3.24) \qquad N_{\geq k} = \sum_{i \in [n]} \mathbb{1}_{\{D_i \geq k\}}.$$

Then, by conditioning on $N_{\geq k}$, we obtain

$$(3.25)$$

$$\mathbb{E}[Z_n(\beta, B)] \leq \mathrm{e}^{O(1/n)} \sum_{m,j=0}^{n} \binom{n}{j} \mathrm{e}^{B(2j-n)} \mathbb{E}\left[\mathrm{e}^{\ell_n^{(k)}[\beta/2 + F_\beta(\ell_j^{(k)}/\ell_n^{(k)})]} \mathrm{e}^{\beta(\ell_n - \ell_n^{(k)})/2} \mid N_{\geq k} = m \right] \mathbb{P}(N_{\geq k} = m).$$

Conditionally on $N_{\geq k} = m$, the vectors $(D_i \wedge k)_{i \in [n]}$ and $\ell_n - \ell_n^{(k)} = \sum_{i \in [n]}(D_i - k)\mathbb{1}_{\{D_i \geq k\}}$ are independent, so that

$$\mathbb{E}[Z_n(\beta, B)] \leq \mathrm{e}^{O(1/n)} \sum_{m,j=0}^{n} \binom{n}{j} \mathrm{e}^{B(2j-n)} \mathbb{E}\left[\mathrm{e}^{\ell_n^{(k)}[\beta/2 + F_\beta(\ell_j^{(k)}/\ell_n^{(k)})]} \mid N_{\geq k} = m \right]$$

$$(3.26) \qquad\qquad \times \mathbb{E}\left[\mathrm{e}^{\beta(\ell_n - \ell_n^{(k)})/2} \mid N_{\geq k} = m \right] \mathbb{P}(N_{\geq k} = m).$$

We compute

$$(3.27) \qquad \mathbb{E}\left[\mathrm{e}^{\beta(\ell_n - \ell_n^{(k)})/2} \mid N_{\geq k} = m \right] = \mathbb{E}[\mathrm{e}^{\beta(D-k)/2} \mid D \geq k]^m.$$

In the following lemma, we investigate the other conditional expectation:

Lemma 3.2 (Conditional pressure per particle). *For any $q_{\geq k} \in [0,1]$ and $k \geq 1$,*

$$\lim_{n \to \infty} \frac{1}{n} \log \sum_{j=0}^{n} \binom{n}{j} \mathrm{e}^{B(2j-n)} \mathbb{E}\left[\mathrm{e}^{\ell_n^{(k)}[\beta/2 + F_\beta(\ell_j^{(k)}/\ell_n^{(k)})]} \mid N_{\geq k} = \lceil n q_{\geq k} \rceil \right]$$

$$(3.28) \qquad\qquad = \sup_{\widetilde{q}} \left[\varphi^{\mathrm{an,d}}(\beta, B; \widetilde{q}) - \widetilde{H}(\widetilde{q} \mid \widetilde{p}) \right],$$

where $\widetilde{p} = (p_1, \ldots, p_{k-1}, p_{\geq k})$ and $\widetilde{q} = (q_1, \ldots, q_{k-1}, q_{\geq k})$, and

$$(3.29) \qquad \widetilde{H}(\widetilde{q} \mid \widetilde{p}) = \sum_{i=1}^{k-1} q_i \log(q_i/p_i) - (1 - q_{\geq k}) \log\big((1 - q_{\geq k})/(1 - p_{\geq k})\big).$$

Proof. This follows immediately from Varadhan's lemma, as in (3.8), noting that the vector $(D_i \wedge k)_{i \in [n]}$ satisfies a conditional large deviation principle with rate function $\widetilde{H}(\widetilde{q} \mid \widetilde{p})$. Indeed, conditionally on $N_{\geq k} = \lceil n q_{\geq k} \rceil$, this sequence consists of $q_{\geq k} n$ values k, and the remaining $(1 - q_{\geq k})n$ values are i.i.d. with distribution $\widetilde{p}$. The rate function of this sequence can be computed using the ratio of probabilities as

$$\widetilde{H}(\widetilde{q} \mid \widetilde{p}) = -\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(N_1 = q_1 n, \ldots, N_{k-1} = q_{k-1} n, N_{\geq k} = q_{\geq k} n) + \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(N_{\geq k} = q_{\geq k} n)$$

$$= \left[\sum_{i=1}^{k-1} q_i \log(q_i/p_i) + q_{\geq k} \log(q_{\geq k}/p_{\geq k}) \right]$$

$$(3.30)$$

$$\qquad - \left[q_{\geq k} \log(q_{\geq k}/p_{\geq k}) + (1 - q_{\geq k}) \log\big((1 - q_{\geq k})/(1 - p_{\geq k})\big) \right],$$

and the $q_{\geq k} \log(q_{\geq k}/p_{\geq k})$ terms cancel, as required. $\qquad\square$

We conclude that

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}[Z_n(\beta,B)]\leq\sup_{\widetilde{\boldsymbol{q}}}\Big[\varphi^{\mathrm{an,d}}(\beta,B;\widetilde{\boldsymbol{q}})-\widetilde{H}(\widetilde{\boldsymbol{q}}\mid\widetilde{\boldsymbol{p}})+q_{\geq k}\log\mathbb{E}[\mathrm{e}^{\beta(D-k)/2}\mid D\geq k]$$

$$-q_{\geq k}\log(q_{\geq k}/p_{\geq k})-(1-q_{\geq k})\log((1-q_{\geq k})/(1-p_{\geq k}))\Big]$$

$$=\sup_{\widetilde{\boldsymbol{q}}}\Big[\varphi^{\mathrm{an,d}}(\beta,B;\widetilde{\boldsymbol{q}})-\sum_{i=1}^{k-1}q_i\log(p_i/q_i)-q_{\geq k}\log(q_{\geq k})$$

$$(3.31)\qquad\qquad +q_{\geq k}\log\mathbb{E}[\mathrm{e}^{\beta(D-k)/2}\mathbb{1}_{\{D\geq k\}}]\Big].$$

When k is sufficiently large $\mathbb{E}[\mathrm{e}^{\beta(D-k)/2}\mathbb{1}_{\{D\geq k\}}]\leq 1$ by dominated convergence, so that we arrive at

$$(3.32)\quad\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}[Z_n(\beta,B)]\leq\sup_{\widetilde{\boldsymbol{q}}}\Big[\varphi^{\mathrm{an,d}}(\beta,B;\widetilde{\boldsymbol{q}})-\sum_{i=1}^{k-1}q_i\log(p_i/q_i)-q_{\geq k}\log(q_{\geq k})\Big].$$

We aim to let $k\to\infty$ in the above expression. We see that, as $k\to\infty$,

$$(3.33)\qquad\qquad\sum_{i=1}^{k-1}q_i\log(p_i/q_i)-q_{\geq k}\log(q_{\geq k})\to\sum_{i=1}^{\infty}q_i\log(p_i/q_i)$$

for every proper probability mass function $\boldsymbol{q}$. What remains to prove is that the optimal $\boldsymbol{q}$ does not put mass in infinity:

Lemma 3.3 (Optimizer does not put mass at infinity). *For any $\varepsilon>0$,*

$$(3.34)\qquad\qquad\liminf_{k\to\infty}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}[Z_n(\beta,B)\mathbb{1}_{\{\sum_{i\in[n]}\mathbb{1}_{\{D_i\geq k\}}\geq\varepsilon n\}}]=-\infty.$$

Proof. We bound as in (3.2),

$$\mathbb{E}\Big[Z_n(\beta,B)\mathbb{1}_{\{\sum_{i\in[n]}\mathbb{1}_{\{D_i\geq k\}}\geq\varepsilon n\}}\Big]\leq 2^n\mathrm{e}^{n|B|}\mathbb{E}\Big[\mathrm{e}^{\beta\sum_{i\in[n]}D_i/2}\mathbb{1}_{\{\sum_{i\in[n]}\mathbb{1}_{\{D_i\geq k\}}\geq\varepsilon n\}}\Big]$$

$$(3.35)\qquad\qquad =2^n\mathrm{e}^{n|B|}\sum_{m\geq\varepsilon n}\binom{n}{m}p_{\geq k}^m(1-p_{\geq k})^{n-m}\mathbb{E}\Big[\mathrm{e}^{\beta\sum_{i\in[n]}D_i/2}\mid N_{\geq k}=m\Big].$$

We compute

$$(3.36)\qquad\mathbb{E}\Big[\mathrm{e}^{\beta\sum_{i\in[n]}D_i/2}\mid N_{\geq k}=m\Big]=\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\mid D\geq k\Big]^m\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\mid D<k\Big]^{n-m}.$$

This yields

$$\mathbb{E}\Big[Z_n(\beta,B)\mathbb{1}_{\{\sum_{i\in[n]}\mathbb{1}_{\{D_i\geq k\}}\geq\varepsilon n\}}\Big]\leq 2^n\mathrm{e}^{n|B|}\sum_{m\geq\varepsilon n}\binom{n}{m}\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\mathbb{1}_{\{D\geq k\}}\Big]^m\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\mathbb{1}_{\{D<k\}}\Big]^{n-m}$$

$$(3.37)\qquad\qquad =2^n\mathrm{e}^{n|B|}\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\Big]^n\sum_{m\geq\varepsilon n}\binom{n}{m}r_k^m(1-r_k)^{n-m},$$

where

$$(3.38)\qquad\qquad r_k=\frac{\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\mathbb{1}_{\{D\geq k\}}\Big]}{\mathbb{E}\Big[\mathrm{e}^{\beta D/2}\Big]}.$$

27

Since $\mathbb{E}\big[e^{\beta D/2}\big] < \infty$, we have that $\lim_{k\to\infty} r_k = 0$, so that

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{m \geq \varepsilon n} \binom{n}{m} r_k^m (1 - r_k)^{n-m} = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathsf{Bin}(n, r_k) \geq \varepsilon n)$$

$$= -\varepsilon \log(\varepsilon/r_k) - (1 - \varepsilon) \log((1 - \varepsilon)/(1 - r_k))$$

(3.39)
$$\to -\infty,$$

for any $\varepsilon > 0$ as $k \to \infty$. $\qquad\square$

We arrive at

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[Z_n(\beta, B)] \leq \sup_{\boldsymbol{q}} \Big[\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}) - \sum_{i=1}^{\infty} q_i \log(p_i/q_i)\Big]$$

(3.40)
$$= \sup_{\boldsymbol{q}} \Big[\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}) - H(\boldsymbol{q} \mid \boldsymbol{p})\Big].$$

This proves the required upper bound. $\qquad\square$

Together with the corresponding lower bound proved in (3.12), we arrive at

(3.41)
$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[Z_n(\beta, B)\right] = \sup_{\boldsymbol{q}}[\varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}) - H(\boldsymbol{q} \mid \boldsymbol{p})],$$

which establishes the formula for the annealed pressure for i.i.d. degrees having infinite support. $\qquad\square$

3.4. Annealed critical value for configuration model with i.i.d. degrees: Proof of Theorem 1.6. We first notice that

(3.42)
$$\varphi^{\mathrm{an,D}}(\beta, B) = \sup_{w, \boldsymbol{q}} R_{\beta, B}(w, \boldsymbol{q}),$$

where

(3.43)
$$R_{\beta, B}(w, \boldsymbol{q}) = -H(\boldsymbol{q} \mid \boldsymbol{p}) + r_{\beta, B}(w, \boldsymbol{q})$$

and

(3.44)
$$r_{\beta, B}(w, \boldsymbol{q}) = \frac{\beta \mathbb{E}[D(\boldsymbol{q})]}{2} + G_{\beta, B}((s_k(w, B))_{k \geq 1}; \boldsymbol{q}).$$

Here we explicitly write the parameters $(\beta, B, \boldsymbol{q})$ of all functions involved. We claim that $R_{\beta, B}(\cdot)$ is a continuous function in a compact space $[0, 1] \times \mathcal{P}$, with $\mathcal{P}$ the space of discrete probability measures. Thus, there exists a non-empty set of maximizers.

In the proof below, we fix $\beta < \beta_c^{\mathrm{an,D}}$. For $B > 0$, define

(3.45) $\quad \bar{w}(B) = \sup\{w^{\star}: \text{ there exists } \boldsymbol{q}^{\star} \text{ such that } (w^{\star}, \boldsymbol{q}^{\star}) \text{ is a maximizer of } R_{\beta, B}(\cdot)\}.$

The proof consists of four steps.

Step 1: An optimizer $w^\star(B)$ that converges to 1. We first show that $\bar{w}(0^+) = \limsup_{B\to 0^+} \bar{w}(B) = 1$ by contradiction. Assume that $\bar{w}(0^+) < 1$. Then there exists $\delta, \varepsilon > 0$, such that when $0 < B < \delta$, we have $w^\star = w^\star(B) < 1 - \varepsilon$ for all maximizers $(w^\star, \boldsymbol{q}^\star)$ of $R_{\beta,B}$. Since $[0,1] \times \mathcal{P}$ is compact, as $B \searrow 0$, we can take a subsequence of $(w^\star, \boldsymbol{q}^\star)$ converging to a limit $(\tilde{w}, \tilde{\boldsymbol{q}})$ with $\tilde{w} \le 1 - \varepsilon$. Therefore,

$$(3.46) \qquad \varphi^{\mathrm{an,D}}(\beta, 0^+) = R_{\beta,0}(\tilde{w}, \tilde{\boldsymbol{q}}) = -H(\tilde{\boldsymbol{q}} \mid \boldsymbol{p}) + \frac{\beta \mathbb{E}[D(\tilde{\boldsymbol{q}})]}{2} + G_{\beta,0}((s_k(\tilde{w}, 0))_{k\ge 1}; \tilde{\boldsymbol{q}}).$$

On the other hand,

$$(3.47) \qquad \varphi^{\mathrm{an,D}}(\beta, B) \ge -H(\tilde{\boldsymbol{q}} \mid \boldsymbol{p}) + \frac{\beta \mathbb{E}[D(\tilde{\boldsymbol{q}})]}{2} + G_{\beta,B}((s_k(\tilde{w}, B))_{k\ge 1}; \tilde{\boldsymbol{q}}).$$

Thus,

$$
\begin{aligned}
M^{\mathrm{an,D}}(\beta) &= \lim_{B\searrow 0} \frac{\varphi^{\mathrm{an,D}}(\beta, B) - \varphi^{\mathrm{an,D}}(\beta, 0)}{B} \\
&\ge \lim_{B\searrow 0} \frac{G_{\beta,B}((s_k(\tilde{w}, B))_{k\ge 1}; \tilde{\boldsymbol{q}}) - G_{\beta,0}((s_k(\tilde{w}, 0))_{k\ge 1}; \tilde{\boldsymbol{q}})}{B} \\
&= \lim_{B\searrow 0} \frac{G_{\beta,B}((s_k(\tilde{w}, B))_{k\ge 1}; \tilde{\boldsymbol{q}}) - G_{\beta,B}((s_k(\tilde{w}, 0))_{k\ge 1}; \tilde{\boldsymbol{q}})}{B} \\
&\quad + \lim_{B\searrow 0} \frac{G_{\beta,B}((s_k(\tilde{w}, 0))_{k\ge 1}; \tilde{\boldsymbol{q}}) - G_{\beta,0}((s_k(\tilde{w}, 0))_{k\ge 1}; \tilde{\boldsymbol{q}})}{B}.
\end{aligned}
$$

The first term converges to

$$(3.48) \qquad \sum_{k\ge 1} \frac{\partial}{\partial s_k} G_{\beta,0}((s_k(\tilde{w}, B))_{k\ge 1})\big|_{B=0} \frac{\partial}{\partial B} s_k(\tilde{w}, B) = 0,$$

since, by assumption, $(s_k(\tilde{w}, 0))_{k\ge 1}$ is a stationary point of $G_{\beta,0}((s_k)_{k\ge 1})$. The second term converges to $\sum_{k\ge 1} \tilde{q}_k (2 s_k(\tilde{w}, 0) - 1)$, so that

$$
\begin{aligned}
M^{\mathrm{an,D}}(\beta) &\ge \sum_{k\ge 1} \tilde{q}_k (2 s_k(\tilde{w}, 0) - 1) = 2 \sum_{k\ge 1} \frac{\tilde{q}_k}{1 + \tilde{w}^k} - 1 \\
&= \mathbb{E}\left[\frac{2}{1 + w^{\tilde{D}}}\right] - 1 = \mathbb{E}\left[\frac{1 - w^{\tilde{D}}}{1 + w^{\tilde{D}}}\right] \\
&\ge \mathbb{E}\left[\frac{1 - (1-\varepsilon)^{\tilde{D}}}{1 + (1-\varepsilon)^{\tilde{D}}}\right] > 0.
\end{aligned}
$$

Here we have used that $\tilde{w} \le 1 - \varepsilon$ and $\tilde{D}$- the random variable with law $\tilde{\boldsymbol{q}}$ is not 0 almost surely. By (1.10) applied to $M^{\mathrm{an,D}}(\beta)$ and $\beta_c^{\mathrm{an,D}}$, this contradicts the assumption that $\beta < \beta_c^{\mathrm{an,D}}$.

Step 2: A Lagrange multiplier equation. We have now proved that $\bar{w}(0^+) = 1$, and thus we can assume that $w^\star(B) \to 1$ as $B \searrow 0$ for one sequence of maximizers $(w^\star(B))_{B \geq 0}$. Now we analyze the limit of $\boldsymbol{q}^\star = \boldsymbol{q}^\star(B)$ as $B \searrow 0$. Since $(w^\star(B), \boldsymbol{q}^\star(B))$ is a maximizer of $R_{\beta,B}$, $(w^\star(B), \boldsymbol{q}^\star(B))$ satisfies

$$(3.49) \qquad \frac{\partial R_{\beta,B}(w^\star(B), \boldsymbol{q}^\star(B))}{\partial q_i} = \lambda,$$

for some Lagrange multiplier λ that is due to the restriction that $\boldsymbol{q}^\star(B)$ is a probability measure, and all i for which $p_i > 0$. If $\boldsymbol{p}$ only takes finitely many values, the above system indeed provides the critical points of $R_{\beta,B}$.

The above Lagrange equation is standard when the degree distribution has finite support, but requires some extra arguments when $\boldsymbol{p}$ has infinite support. In this paragraph, we fix (β, B) are omit them from the notation, as we aim to show that (3.49) holds for all (β, B). We now provide these arguments. First, we prove that the optimizer $\boldsymbol{q}^\star$ satisfies that $q_i^\star > 0$ whenever $p_i > 0$. For this, we fix $i \geq 1$ for which $p_i > 0$ and a $j \geq 1$ for which $q_j^\star > 0$. Assume that $q_i^\star = 0$, and consider, instead, the probability distribution $\boldsymbol{q}(\varepsilon) = \boldsymbol{q}^\star + \varepsilon \delta_i - \varepsilon \delta_j$, where δ_i is the Kronecker-delta on $i \in \mathbb{N}$. Since $(w^\star, \boldsymbol{q}^\star)$ is the maximizer, we must have that

$$(3.50) \qquad H(\boldsymbol{q}(\varepsilon) \mid \boldsymbol{p}) - r(w^\star, \boldsymbol{q}(\varepsilon)) \leq H(\boldsymbol{q}^\star \mid \boldsymbol{p}) - r(w^\star, \boldsymbol{q}^\star).$$

By the identification of r, the function the function $q_i \mapsto r(w^\star, \boldsymbol{q})$ is differentiable. However,

$$(3.51) \quad H(\boldsymbol{q}(\varepsilon) \mid \boldsymbol{p}) - H(\boldsymbol{q}^\star \mid \boldsymbol{p}) = \varepsilon \log(\varepsilon/p_i) + (q_j^\star - \varepsilon) \log((q_j^\star - \varepsilon)/p_j) - q_j^\star \log((q_j^\star - \varepsilon)/p_j),$$

which is $\varepsilon \log(\varepsilon)(1 + o(1))$, and which has a derivative $-\infty$ at $\varepsilon = 0$. This contradicts with the fact that $\boldsymbol{q}^\star$ is the maximizer. We conclude that the optimizer $\boldsymbol{q}^\star$ satisfies that $q_i^\star > 0$ whenever $p_i > 0$.

We continue by proving that (3.49) holds. Fix i, j for which $p_i, p_j > 0$. By the previous argument, also $q_i^\star, q_j^\star > 0$. Pick ε so small that $\boldsymbol{q}(\varepsilon) = \boldsymbol{q}^\star + \varepsilon \delta_i - \varepsilon \delta_j$ is a probability measure. Then, by the fact that $(w^\star, \boldsymbol{q}^\star)$ is the maximizer of $H(\boldsymbol{q} \mid \boldsymbol{p}) - r(w, \boldsymbol{q})$, we obtain that

$$(3.52) \qquad \frac{\partial}{\partial \varepsilon}[H(\boldsymbol{q}(\varepsilon) \mid \boldsymbol{p}) - r(w^\star, \boldsymbol{q}(\varepsilon))]_{\varepsilon=0} = 0,$$

which is equivalent to

$$(3.53) \qquad \frac{\partial}{\partial q_i}\Big[H(\boldsymbol{q} \mid \boldsymbol{p}) - r(w^\star, \boldsymbol{q})\Big]_{\boldsymbol{q}=\boldsymbol{q}^\star} = \frac{\partial}{\partial q_j}\Big[H(\boldsymbol{q} \mid \boldsymbol{p}) - r(w^\star, \boldsymbol{q})\Big]_{\boldsymbol{q}=\boldsymbol{q}^\star}.$$

Since this is true for all i, j for which $p_i, p_j > 0$, we get (3.49).

Step 3: Identification of a limit of the optimizer $\boldsymbol{q}^\star(B)$. In this step, we show that, along an appropriate subsequence, $\boldsymbol{q}^\star(B)$ converges to an explicit limit. Note that $\frac{\partial}{\partial q_i}H(\boldsymbol{q} \mid \boldsymbol{p}) = \log(q_i/p_i) + 1$. We can thus alternatively write (3.49) as

$$(3.54) \qquad \log(q_i^\star(B)/p_i) - \frac{\partial}{\partial q_i}r_{\beta,B}(w^\star(B), \boldsymbol{q})\big|_{\boldsymbol{q}=\boldsymbol{q}^\star(B)} = \lambda.$$

By (3.44),

$$(3.55) \qquad \frac{\partial}{\partial q_i} r_{\beta,B}(w^\star(B), \boldsymbol{q}) = \frac{\partial}{\partial q_i} [\beta \mathbb{E}[D(\boldsymbol{q})]/2 + G_{\beta,B}((s_k(w^\star))_{k\geq 1}; \boldsymbol{q})].$$

Further, by (2.15),

$$(3.56) \quad G_{\beta,B}((s_k)_{k\geq 1}; \boldsymbol{q}) = \sum_{k\geq 1} q_k I(s_k) + B\Big(2 \sum_{k\geq 1} s_k q_k - 1\Big) + \mathbb{E}[D(\boldsymbol{q})] F_\beta\Big(\frac{\sum_{k\geq 1} kq_k s_k}{\mathbb{E}[D(\boldsymbol{q})]}\Big),$$

Thus,

$$\frac{\partial}{\partial q_i} G_{\beta,B}((s_k)_{k\geq 1}; \boldsymbol{q}) = I(s_i) + 2s_i B + i F_\beta\Big(\frac{\sum_{k\geq 1} kq_k s_k}{\mathbb{E}[D(\boldsymbol{q})]}\Big)$$

$$(3.57) \qquad\qquad + \mathbb{E}[D(\boldsymbol{q})] F'_\beta\Big(\frac{\sum_{k\geq 1} kq_k s_k}{\mathbb{E}[D(\boldsymbol{q})]}\Big)\Big(\frac{is_i}{\mathbb{E}[D(\boldsymbol{q})]} - \frac{i \sum_{k\geq 1} kq_k s_k}{\mathbb{E}[D(\boldsymbol{q})]^2}\Big).$$

Take $B \searrow 0$. Extract a subsequence along which $\boldsymbol{q}^\star(B)$ converges to a limit, say $\boldsymbol{q}(\beta)$ and $w^\star(B) \to 1$. Along this subsequence, $(s_k(w^\star(B)))_{k\geq 1} \to (1/2)_{k\geq 1}$. Then, using that $F_\beta(\frac{1}{2}) = \log\big(\frac{1}{2}[1 + e^{-2\beta}]\big)$, (3.57) becomes

$$(3.58) \qquad \frac{\partial}{\partial q_i} G_{\beta,0}((s_k(1))_{k\geq 1}; \boldsymbol{q}(\beta)) = \log 2 + i F_\beta(\tfrac{1}{2}) = \log 2 + \frac{i}{2} \log\big(\tfrac{1}{2}[1 + e^{-2\beta}]\big).$$

We conclude that (3.49) becomes

$$(3.59) \quad \log(q_i(\beta)/p_i) + \beta i/2 + \frac{i}{2} \log\big(\tfrac{1}{2}[1 + e^{-2\beta}]\big) = \log(q_i(\beta)/p_i) + \frac{i}{2} \log\cosh(\beta) = \lambda,$$

so that

$$(3.60) \qquad\qquad q_i(\beta) = p_i \cosh(\beta)^{i/2}/c(\beta),$$

where $c(\beta) = \mathbb{E}[\cosh(\beta)^{D/2}] < \infty$, since $\mathbb{E}[e^{\beta D/2}] < \infty$.

Step 4: Completion of the proof of Theorem 1.6. For all $B \geq 0$, we have

$$(3.61) \qquad\qquad \varphi^{\mathrm{an,D}}(\beta, B) \geq -H(\boldsymbol{q}(\beta) \mid \boldsymbol{p}) + \varphi^{\mathrm{an,d}}(\beta, B; \boldsymbol{q}(\beta)).$$

Moreover, since $w^\star(B) \to 1$ and $\boldsymbol{q}^\star(B) \to \boldsymbol{q}(\beta)$,

$$\varphi^{\mathrm{an,D}}(\beta, 0^+) = \lim_{B\searrow 0} R(w^\star, \boldsymbol{q}^\star) = R(1, \boldsymbol{q}(\beta)) \quad = \quad -H(\boldsymbol{q}(\beta) \mid \boldsymbol{p}) + r(1, \boldsymbol{q}(\beta))$$

$$(3.62) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \leq \quad -H(\boldsymbol{q}(\beta) \mid \boldsymbol{p}) + \varphi^{\mathrm{an,d}}(\beta, 0; \boldsymbol{q}(\beta)).$$

Combining the last two inequalities, we obtain, using the definition of the spontaneous magnetization

$$M^{\mathrm{an,D}}(\beta) = \lim_{B\searrow 0} \frac{\varphi^{\mathrm{an,D}}(\beta, B) - \varphi^{\mathrm{an,D}}(\beta, 0)}{B} \quad \geq \quad \lim_{B\searrow 0} \frac{\varphi^{\mathrm{an,d}}(\beta, B) - \varphi^{\mathrm{an,d}}(\beta, 0)}{B}$$

$$(3.63) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \quad M^{\mathrm{an,d}}(\beta; \boldsymbol{q}(\beta)).$$

Since $\beta < \beta_c^{\mathrm{an,D}}$, $M^{\mathrm{an,D}}(\beta) = 0$. Since $M^{\mathrm{an,d}}(\beta; \boldsymbol{q}(\beta)) \geq 0$ we conclude that $M^{\mathrm{an,d}}(\beta; \boldsymbol{q}(\beta)) = 0$. This is equivalent to

$$(3.64) \qquad\qquad\qquad \beta \leq \mathrm{atanh}(1/\nu(\boldsymbol{q}(\beta))),$$

where we write

$$(3.65) \qquad \nu(\boldsymbol{q}) = \frac{\sum_{k \geq 1} k(k-1) q_k}{\sum_{k \geq 1} k q_k}.$$

We also notice that $\beta \mapsto \operatorname{atanh}(1/\nu(\boldsymbol{q}(\beta)))$ is monotonically decreasing. Hence the equation $\beta = \operatorname{atanh}(1/\nu(\boldsymbol{q}(\beta)))$ has a unique solution, denoted by $\bar{\beta}_c^D$. Moreover, if $\beta \leq \operatorname{atanh}(1/\nu(\boldsymbol{q}(\beta)))$ then $\beta \leq \bar{\beta}_c^D$. Thus we have proved that if $\beta < \beta_c^{\mathrm{an},D}$, then $\beta \leq \bar{\beta}_c^D$. Therefore, $\beta_c^{\mathrm{an},D} \leq \bar{\beta}_c^D$. Furthermore, for any $\beta \geq 0$,

$$(3.66) \qquad \nu(\boldsymbol{q}(\beta)) \geq \nu,$$

since

$$(3.67) \qquad \nu(\boldsymbol{q}(\beta)) = \frac{\mathbb{E}[(D^\star - 1)\cosh(\beta)^{D^\star/2}]}{\mathbb{E}[\cosh(\beta)^{D^\star/2}]} \geq \mathbb{E}[D^\star - 1] = \nu,$$

with strict inequality unless $D^\star$ is constant or $\beta = 0$. Thus, we obtain that when $\bar{\beta}_c^D > 0$ and the degrees are not a.s. constant,

$$(3.68) \qquad \beta_c^{\mathrm{an},D} \leq \bar{\beta}_c^D = \operatorname{atanh}(1/\nu(\boldsymbol{q}(\bar{\beta}_c^D))) < \operatorname{atanh}(1/\nu) = \beta_c^{\mathrm{qu}}.$$

This completes the proof of Theorem 1.6. $\qquad\qquad\square$

Remark 3.4 (Equality $\beta_c^{\mathrm{an},D} = \bar{\beta}_c^D$). We believe that $\beta_c^{\mathrm{an},D} = \bar{\beta}_c^D$, but lack a proof of this fact. Indeed, we expect that the magnetization $M^{\mathrm{an},D}(\beta, B) \approx 0$ when $(\beta, B) \approx (\beta_c^{\mathrm{an},D}, 0)$ with $\beta \geq \beta_c^{\mathrm{an},D}, B \geq 0$. This, in turn, would mean that there exists an optimizer $w^\star(\beta, B)$ that is close to 1, so that there also exists an optimizer $(s_k^\star(\beta, B))_{k \geq 1}$ that is close to $\frac{1}{2}$. As a result, there should also be an optimizer $\boldsymbol{q}_k(\beta, B)$ that is close to $\boldsymbol{q}(\beta)$, which should thus be close to critical. This should mean that $\beta_c^{\mathrm{an},D} = \operatorname{atanh}(1/\nu(\boldsymbol{q}(\beta_c^{\mathrm{an},D})))$. Our proof, however, does not yield this, since we have insufficient control over the optimizers.

Remark 3.5 (Critical behavior). Assume that $\beta_c^{\mathrm{an},D} = \bar{\beta}_c^D$, as we argue above, as well as $\mathbb{E}[\mathrm{e}^{\beta_c^{\mathrm{an},D} D/2}] < \infty$. Then, the solution $q_i(\beta)$ in (3.60) has exponential tails, since $\cosh(\beta) < \mathrm{e}^\beta$. Thus, by (3.57), one can expect the solution $\boldsymbol{q}(\beta, B)$ to have exponential tails as well for $B > 0$ small enough and $\beta > \beta_c^{\mathrm{an},D}$ with $\beta - \beta_c^{\mathrm{an},D}$ sufficiently small. As a result, one can expect that the critical behavior is related to that of an annealed system with a degree distribution with exponential tails, which should be equal to those for the Curie-Weiss model. Since this is true for *all* degree distributions with $\mathbb{E}[\mathrm{e}^{\beta_c^{\mathrm{an},D} D/2}] < \infty$, this suggests that the power-law universality class does not arise for i.i.d. degrees, which is quite surprising.

Example 3.6 (Poisson degrees). Let us compute the critical inverse temperature when D is the Poisson distribution with intensity $\lambda > 0$, that is $\mathbb{P}(D = i) = p_i$, $i \geq 0$, with

$$(3.69) \qquad p_i = \mathrm{e}^{-\lambda} \frac{\lambda^i}{i!}.$$

From (3.60) we get the optimizer for all $\beta > 0$:

$$(3.70) \qquad q_i(\beta) = \frac{p_i \cosh(\beta)^{i/2}}{c(\beta)} = \frac{1}{\mathrm{e}^\lambda c(\beta)} \frac{(\lambda \sqrt{\cosh(\beta)})^i}{i!}.$$

Hence, the expected forward degree corresponding to distribution $\boldsymbol{q}(\beta) = (q_i(\beta))_i$ is

$$(3.71) \qquad \nu(\boldsymbol{q}(\beta)) = \frac{\sum_{i \geq 0} i(i-1) q_i(\beta)}{\sum_{i \geq 0} i q_i(\beta)} = \frac{\sum_{i \geq 2} (\lambda \sqrt{\cosh(\beta)})^i / (i-2)!}{\sum_{i \geq 1} (\lambda \sqrt{\cosh(\beta)})^i / (i-1)!} = \lambda \sqrt{\cosh(\beta)}.$$

The upper bound $\bar{\beta}_c^{\mathrm{D}}$ satisfies (1.23), or equivalently

$$\tanh(\bar{\beta}_c^{\mathrm{D}}) = \left(\lambda \sqrt{\cosh(\bar{\beta}_c^{\mathrm{D}})}\right)^{-1},$$

that can be rewritten as

$$(3.72) \qquad \cosh(\bar{\beta}_c^{\mathrm{D}}) = \lambda^2 (\cosh(\bar{\beta}_c^{\mathrm{D}})^2 - 1).$$

Solving this equation for $\cosh(\bar{\beta}_c^{\mathrm{D}})$ and inverting the hyperbolic cosine, we obtain

$$(3.73) \qquad \beta_c^{\mathrm{an,D}} \leq \bar{\beta}_c^{\mathrm{D}} = -\log(2\lambda^2) + \log\left[1 + \sqrt{1 + 4\lambda^4} + \sqrt{2 + 2\sqrt{1 + 4\lambda^4}}\right].$$

On the other hand, the quenched critical value is $\beta_c^{\mathrm{qu,d}} = \mathrm{atanh}(1/\nu(\boldsymbol{p}))$, where $\nu(\boldsymbol{p}) = \lambda$ (as can be readily seen by setting $\beta = 0$ in (3.71)), that is

$$(3.74) \qquad \beta_c^{\mathrm{qu,d}} = \tfrac{1}{2} \log\left(\frac{\lambda+1}{\lambda-1}\right).$$

Let us observe that, while $\beta_c^{\mathrm{an,D}}$ exists for any $\lambda > 0$, the quenched critical value is finite only if $\lambda > 1$, and as $\lambda \searrow 1$, $\beta_c^{\mathrm{qu,d}} \to \infty$, while $\beta_c^{\mathrm{an,D}} \leq \bar{\beta}_c^{\mathrm{D}} \to \log[(1 + \sqrt{5} + \sqrt{2 + 2\sqrt{5}})/2]$.

Example 3.7 (Geometric degrees). We next investigate the case of geometric degrees, i.e., we assume that $\mathbb{P}(D = i) = p_i$ with

$$(3.75) \qquad p_i = (1-p)^{i-1} p, \quad i \geq 1,$$

for some $0 < p < 1$.

As a first remark, to make sure that the annealed pressure is finite, we use Prop. 1.4, and thus require

$$(3.76) \qquad \mathbb{E}[e^{\beta D/2}] < \infty,$$

which amounts to

$$(3.77) \qquad \beta < -2 \log(1 - p).$$

Then, by (3.60),

$$(3.78) \qquad q_i(\beta) = \frac{(1-p)^{i-1} p (\sqrt{\cosh(\beta)})^i}{c(\beta)}, \quad i \geq 1.$$

Hence,

$$\sum_{i \geq 1} i q_i(\beta) = \frac{p\sqrt{\cosh(\beta)}}{c(\beta)} \sum_{i \geq 1} i((1-p)\sqrt{\cosh(\beta)})^{i-1}$$

$$(3.79) \qquad = \frac{p\sqrt{\cosh(\beta)}}{c(\beta)} \left(1 - (1-p)\sqrt{\cosh(\beta)}\right)^{-2},$$

33

where the summability condition $(1-p)\sqrt{\cosh(\beta)} < 1$ is guaranteed by the assumption (3.77). Furthermore,

$$(3.80) \qquad \sum_{i\geq 1} i(i-1)q_i(\beta) = \frac{p\sqrt{\cosh(\beta)}}{c(\beta)} \frac{2(1-p)\sqrt{\cosh(\beta)}}{\left(1-(1-p)\sqrt{\cosh(\beta)}\right)^3}.$$

Then, from (3.65), we obtain the expected forward degree

$$(3.81) \qquad \nu(\mathbf{q}(\beta)) = \frac{2(1-p)\sqrt{\cosh(\beta)}}{1-(1-p)\sqrt{\cosh(\beta)}}.$$

Thus, in this case, the equation for $\bar{\beta}_c^{\mathrm{D}}$ in (1.23) reads as

$$(3.82) \qquad \tanh(\bar{\beta}_c^{\mathrm{D}}) = \left(2(1-p)\sqrt{\cosh(\bar{\beta}_c^{\mathrm{D}})}\right)^{-1} - \frac{1}{2}.$$

This equation looks similar to the corresponding equation for the Poisson case, but is in fact quite different. Setting, for the sake of notation, $x = \sqrt{\cosh(\bar{\beta}_c^{\mathrm{D}})}$ and $r = 1-p$, we can rewrite (3.82) as

$$(3.83) \qquad \frac{\sqrt{x^4-1}}{x^2} = \frac{1-rx}{2rx}$$

that, after some manipulation, can be transformed into

$$(3.84) \qquad x^2(3r^2x^4 + 2rx^3 - x^2 - 4r^2) = 0.$$

Recalling the definition of x, we conclude that the equation for the critical $\bar{\beta}_c^{\mathrm{D}}$ is

$$(3.85) \qquad 3r^2x^4 + 2rx^3 - x^2 - 4r^2 = 0, \quad \text{with} \quad x \geq 1.$$

The polynomial $f_r(x) = 3r^2x^4 + 2rx^3 - x^2 - 4r^2$ (with $r > 0$) has a local negative maximum at $x = 0$ and for $x > 0$ a unique negative local minimum at $x_+(r) = \frac{1}{4r}(\frac{\sqrt{33}}{3} - 1)$. Since $f_r'(x) > 0$ for $x > x_+(r)$, we conclude that there exists a unique positive solution $x^\star(r) \in (x_+(r), \infty)$ to the equation $f_r(x) = 0$. Now we have to check that $x^\star(r)$ corresponds to an admissible solution to (3.85), that is, $1 \leq x^\star(r) < r^{-1}$ for all $0 < r \leq 1$, i.e., for all $0 \leq p < 1$. Observe that we can bound $f_r(x)$ for $x < 1$ as

$$(3.86) \qquad f_r(x) = r^2(3x^4 - 4) + 2rx^3 - x^2 < r^2(3x^4 - 4) + x^2(2r - 1),$$

and, since $3x^4 - 4 < -1$ and $x^2 < 1$,

$$(3.87) \qquad r^2(3x^4 - 4) + x^2(2r - 1) < -r^2 + 2r - 1 = -(1-r)^2 \equiv -p^2.$$

Thus, $f_r(x) < 0$ if $x < 1$ (with $f_r(1) = 0$ if and only if $r = 1$) and this shows that $x^\star(r) \geq 1$. Now we show that $x^*(r) < r^{-1}$ by contradiction. Dropping the dependence of $x^\star(r)$ on r, we write

$$(3.88) \qquad f_r(x^\star) = 3(rx^\star)^2 x^{\star 2} + 2(rx^\star)x^{\star 2} - x^{\star 2} - 4r^2,$$

and assuming $x^\star > 1$ we obtain,

$$(3.89) \qquad f_r(x^\star) > 4x^{\star 2} - 4r^2 > 0,$$

(where the last inequality follows from the fact that $x^\star > 1$ and $r < 1$) in contradiction with the fact that $x^\star$ is the root of $f_r(x^\star) = 0$. The previous inequality proves the claim.

Therefore we conclude that, for any $0 < p < 1$, the critical inverse temperature with i.i.d. geometric degrees is bounded from above by

$$(3.90) \qquad \bar{\beta}_c^{\mathrm{D}} = \ln \left(x^\star(p)^2 + \sqrt{x^\star(p)^4 - 1} \right),$$

where $x^\star(p)$ is the unique solution to the fourth order equation

$$(3.91) \qquad 3(1-p)^2 x^4 + 2(1-p)x^3 - x^2 - 4(1-p)^2 = 0.$$

On the other hand, the quenched critical value with geometric degrees is by (1.19) equal to

$$(3.92) \qquad \beta_c^{\mathrm{qu,d}} = \operatorname{atanh} \left(\frac{p}{2(1-p)} \right).$$

Then, we have $\beta_c^{\mathrm{qu,d}} = \infty$ for $p \geq \frac{2}{3}$, while in the annealed case with i.i.d. degrees there is a finite critical value for any $0 < p < 1$.

References

[1] G. Bianconi. Mean field solution of the Ising model on a Barabási-Albert network. *Phys. Lett.* A303, 166 (2002).

[2] S. Bornholdt and F. Wagner. Stability of money: phase transitions in an Ising economy. *Physica A: Statistical Mechanics and its Applications*, **316**(1):453–468, (2002).

[3] A. Bovier. *Statistical mechanics of disordered systems.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, (2006).

[4] V.H. Can. Critical behavior of the annealed Ising model on random regular graphs. *Journal of Statistical Physics*, **169**:480–503, (2017).

[5] V.H. Can. Annealed limit theorems for the Ising model on random regular graphs. *The Annals of Applied Probability*, **29**(3):1398–1445, (2019).

[6] F. Caravenna, F. den Hollander and N. Pétrélis. Lectures on random polymers. *Clay Mathematics Proceedings* **15**, 319–393, (2012).

[7] P. Contucci, I. Gallo, and G. Menconi. Phase transitions in social sciences: two-population mean field theory. *International Journal of Modern Physics B*, **22**(14):2199–2212, (2008).

[8] P. Contucci and C. Giardinà. *Perspectives on spin glasses.* Cambridge University Press, Cambridge, (2013).

[9] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Brazilian Journal of Probability and Statistics*, **24**(2):137–211, (2010).

[10] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *The Annals of Applied Probability*, **20**(2):565–592, (2010).

[11] A. Dembo, A. Montanari, A. Sly, and N. Sun. The replica symmetric solution for Potts models on d-regular graphs. *Communication in Mathematical Physics*, **327**(2):551–575, (2014)

[12] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree-like graphs, *The Annals of Probability*, **41**(6): 4162–4213, (2013)

[13] S. Dommers, C. Giardinà, C. Giberti, and R. van der Hofstad. Large deviations for the annealed Ising model on inhomogeneous random graphs: spins and degrees *Journal of Statistical Physics*, **173**: 1045–1081 (2018).

[14] S. Dommers, C. Giardinà, C. Giberti, R. van der Hofstad, and M.L. Prioriello. Ising critical behavior of inhomogeneous Curie-Weiss models and annealed random graphs. *Communication in Mathematical Physics*, **348**(1):221–263, (2016).

[15] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising models on power-law random graphs. *Journal of Statistical Physics*, **141**(4):638–660, (2010).

[16] S. Dommers, C. Giardinà, and R. van der Hofstad. Ising critical exponents on random trees and graphs. *Communication in Mathematical Physics*, **328**(1):355–395, (2014).

[17] S. Dorogovtsev, A. Goltsev and J. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, **80**, (4)1275, (2008).

[18] J. Dugundji. *Topology.* Allyn and Bacon, Boston, (1966).

[19] R.S. Ellis. *Entropy, large deviations, and statistical mechanics*, volume **271** of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, (1985).

[20] D. Fraiman, P. Balenzuela, J. Foss, and D. Chialvo. Ising-like dynamics in large-scale functional brain networks. *Physical Review E*, **79**(6):061922, (2009).

[21] C. Giardinà, C. Giberti, R. van der Hofstad, and M.L. Prioriello. Quenched central limit theorems for the Ising model on random graphs. *Journal of Statistical Physics*, **160**(6):1623–1657, (2015).

[22] C. Giardinà, C. Giberti, R. van der Hofstad, and M. L. Prioriello. Annealed central limit theorems for the Ising model on random graphs. *ALEA Latin American Journal of Probability and Mathematical Statistics*, **13**(1):121–161, (2016).

[23] R. van der Hofstad. *Random graphs and complex networks. Volume 1.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, (2017).

[24] R. van der Hofstad. *Random graphs and complex networks. Volume 2.* (2020+). In preparation, see `http://www.win.tue.nl/∼rhofstad/NotesRGCNII.pdf`.

[25] R. van der Hofstad. *Stochastic processes on random graphs.* (2020+). In preparation, see `http://www.win.tue.nl/∼rhofstad/SaintFlour_SPoRG.pdf`.

[26] L. Hörmander *The notion of convexity*, volume **127** of Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, (1994).

[27] GA Kohring. Ising models of social impact: the role of cumulative advantage. *Journal de Physique I*, **6**(2):301–308, (1996).

[28] M. Niss. History of the Lenz-Ising model 1920–1950: from ferromagnetic to cooperative phenomena. *Arch. Hist. Exact Sci.*, **59**(3):267–318, (2005).

[29] M. Niss. History of the Lenz-Ising model 1950–1965: from irrelevance to relevance. *Arch. Hist. Exact Sci.*, **63**(3):243–287, (2009).

[30] B. Simon. *Statistical mechanics of lattice gases.* Princeton University Press, (1993).

Van Hao Can, Research Institute for Mathematical Sciences, Kyoto University, 606–8502 Kyoto, Japan & Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, 10307 Hanoi, Vietnam, cvhao89@gmail.com

Cristian Giardinà, University of Modena and Reggio Emilia, via G. Campi 213/b, 41125 Modena, Italy, cristian.giardina@unimore.it

Claudio Giberti, University of Modena and Reggio Emilia, via G. Amendola 2, Pad. Morselli, 42122 Reggio E., Italy, claudio.giberti@unimore.it

Remco van der Hofstad, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, r.w.v.d.hofstad@tue.nl