Spin models on random graphs

DOI:
10.6100/IR750979

Document status and date:
Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Spin models on random graphs

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN: 978-90-386-3348-0

Printed by Wöhrmann Print Service, Zutphen, the Netherlands.
Cover design by Oranje Vormgevers.

Spin models on random graphs

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op dinsdag 26 maart 2013 om 16.00 uur

door

Sander Dommers

geboren te Dordrecht

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. R.W. van der Hofstad

Copromotor:
prof.dr. C. Giardinà

Summary

Spin models on random graphs

In the past decades complex networks and their behavior have attracted much attention. In the real world many of such networks can be found, for instance as social, information, technological and biological networks. An interesting property that many of them share is that they are *scale free*. Such networks have many nodes with a moderate amount of links, but also a significant amount of nodes with a very high number of links. The latter type of nodes are called *hubs* and play an important role in the behavior of the network. To model scale free networks, we use power-law random graphs. This means that their degree sequences obey a *power law*, i.e., the fraction of vertices that have k neighbors is proportional to $k^{-\tau}$ for some $\tau > 1$.

Not only the structure of these networks is interesting, also the behavior of processes living on these networks is a fascinating subject. Processes one can think of are opinion formation, the spread of information and the spread of viruses. It is especially interesting if these processes undergo a so-called *phase transition*, i.e., a minor change in the circumstances suddenly results in completely different behavior. Hubs in scale free networks again have a large influence on processes living on them. The relation between the structure of the network and processes living on the network is the main topic of this thesis.

We focus on *spin models*, i.e., *Ising* and *Potts* models. In physics, these are traditionally used as simple models to study magnetism. When studied on a random graph, the spins can, for example, be considered as opinions. In that case the *ferromagnetic* or *antiferromagnetic* interactions can be seen as the tendency of two connected persons in a social network to agree or disagree, respectively.

In this thesis we study two models: the ferromagnetic Ising model on power-law random graphs and the antiferromagnetic Potts model on the Erdős-Rényi random graph. For the first model we derive an explicit formula for the thermodynamic limit of the *pressure*, generalizing a result of Dembo and Montanari to random graphs with power-law exponent $\tau > 2$, for which the variance of degrees is potentially infinite. We furthermore identify the thermodynamic limit of the magnetization, internal energy and susceptibility.

For this same model, we also study the phase transition. We identify the *critical temperature* and compute the *critical exponents* of the magnetization and susceptibility. These exponents are *universal* in the sense that they only depend on the power-law exponent τ and not on any other detail of the degree distribution.

The proofs rely on the locally tree-like structure of the random graph. This means that the local neighborhood of a randomly chosen vertex behaves like a branching process.

Correlation inequalities are used to show that it suffices to study the behavior of the Ising model on these branching processes to obtain the results for the random graph. To compute the critical temperature and critical exponents we derive upper and lower bounds on the magnetization and susceptibility. These bounds are essentially Taylor approximations, but for power-law exponents $\tau \leq 5$ a more detailed analysis is necessary.

We also study the case where the power-law exponent $\tau \in (1, 2)$ for which the mean degree is infinite and the graph is no longer locally tree-like. We can, however, still say something about the magnetization of this model.

For the antiferromagnetic Potts model we use an interpolation scheme to show that the thermodynamic limit exists. For this model the correlation inequalities do not hold, thus making it more difficult to study. We derive an extended variational principle and use to it give upper bounds on the pressure. Furthermore, we use a constrained second-moment method to show that the high-temperature solution is correct for high enough temperature. We also show that this solution cannot be correct for low temperatures by showing that the entropy becomes negative if it were to be correct, thus identifying a *phase transition*.

Contents

1
INTRODUCTION

Over one billion people use Facebook each month and over 140 billion friendship connections have been made between them [49]. Almost 10 billion electronic devices are able to connect to the Internet [66] exceeding the number of living people. Airports around the world are connected with thousands of flights between them every day [62]. These are just three examples of huge *complex networks*, a set of nodes with links between them, that can be found in the real world, but many more exist. Many things are going on in such networks. For example, people influence each other in social networks, information is spread via the Internet [75], and viruses spread globally via the airline network [95]. This thesis studies mathematical models for such complex networks and processes living on them.

1.1 Examples of complex networks

In [87, 88], Newman gives extensive overviews of many complex networks. He divides them into four categories: social, information, technological and biological networks. Facebook is just one example of a *social network*, but certainly not the only example in this category. A friendship network in the offline world is also a social network, although this is a bit harder to define: it is not always clear when two people should be considered to be friends and even two persons might disagree themselves whether they should be considered friends. A more well-defined model is the handshake network in which two persons are linked when they ever shook hands. Social networks can also be seen in a broader sense, for example an innovation network with links between producers, users, and governmental bodies [55].

An example of an *information network* is the World Wide Web (WWW), not to be confused with the Internet. It consists of web pages with so-called hyperlinks between them. Scientists cite work of others, giving rise to a citation network of articles.

The Internet [50] is a *technological network*. It is a network of networks of computers and other devices connected mainly with physical cables, although also wireless connections have become common recently. Other examples of technological networks include the power grid [7] and road and rail networks.

Not only man-made networks can be thought of, also in nature networks can be identified. Examples of such *biological networks* include protein-protein interactions [70], metabolic networks [51] and neural networks of the brain [101].

1.2 Properties of complex networks

Although networks have different origins, many of them have similar properties. For instance, many of them show *scale free* behavior [87]. This means that there is not such a thing as a *typical* node in the network, but the network is highly heterogeneous. For example, many people only have a moderate amount of friends but a significant amount of people have a huge number of friends. In fact, the number of links nodes have, called the *degree*, can be found on all scales in such networks. The high degree vertices in a network are called *hubs* and play an important role in the behavior of these networks. This heterogeneity is different from, e.g., the heights of men in a certain country. These heights are often quite close to the average and large deviations from this average are uncommon.

The second property many networks share is that they are *small worlds*: if you select two nodes at random from the network the number of links you have to follow to go from one node to the other is very small, certainly compared to the size of the network. For example, the average distance on Facebook in May 2011 was estimated to be just 4.74 [12].

The scale free and small world properties have a large influence on the behavior of processes living on these networks. They make it possible for example for information, but also viruses, to spread very fast through the network. Also the hubs in the network are much more influential than other vertices in opinion formation.

This makes the behavior of processes living on these networks a fascinating subject. Especially if they undergo a so-called *phase transition*, i.e., a minor change in the circumstances suddenly results in completely different behavior. Examples of such phase transitions include the sudden East European revolution in 1989 [73] and the unusual swine flu outbreak in 2009 [26].

1.3 Random graphs

To better understand the behavior of networks we try to model them mathematically with *random graphs*. These consist of a set of *vertices* representing the nodes of a network and *edges* placed randomly between them according to some specified rules representing the links in the network. Many models have been proposed how to exactly construct such random graphs.

The simplest way of constructing a random graph was proposed by Erdős and Rényi in [46]: start with a number of vertices and then connect each pair of vertices independently with a certain predetermined probability. Although this model seems very simple, it shows very rich behavior [46, 93, 9]. Unfortunately, however, this model does not show the scale-free behavior that is observed in real-world networks.

To overcome this, many alternative random graph models have been proposed. Examples include the inhomogeneous random graph, where vertices are assigned a weight and edges are connected with a probability that depends on these weights, see [24] for an extensive overview of such models. By choosing the weights properly, scale-free behavior can be obtained. Another option is to study the *configuration model*, first introduced in [22]. In this model, the degree distribution is an input to the model: every vertex gets assigned a random number of half-edges according to this distribution and these half-edges are then paired up uniformly at random. In order to get a scale-free graph we choose a *power-law degree distribution*, i.e., the probability that a randomly selected vertex has degree k is proportional to $k^{-\tau}$ for some $\tau > 1$.

The Erdős-Rényi random graph, the inhomogeneous random graph and the configuration model are introduced more formally in Chapter 2. Many more random graph models have been proposed, see, e.g., [23, 63, 69] for extensive reviews on random graphs and their properties. In this thesis, however, we focus on the three above-mentioned models.

1.4 Spin models

To study processes on networks many models have been proposed, see [42] for an overview of models studied in the physics literature. A canonical model to study cooperative behavior is the Ising model, see [89, 90, 91] for its history. In this model every vertex gets assigned a *spin* which can be either in the *up* position or in the *down* position. Originally, the Ising model on a lattice was proposed as a simple model to study the behavior of magnetic materials such as iron. When this model is studied on random graphs, a spin can be thought of, for example, as the opinion of an individual. It is more likely that a certain person has the same opinion as his friends then that his opinion differs. This is what we call a *ferromagnetic* interaction: the spins of neighbors in the graph tend to align.

Ising models have for example been proposed to study integration of immigrants into society [29, 30, 31] and the spread of innovations [82]. Also the brain is suggested to show behavior similar to that of the Ising model [53].

When more than two opinions are possible, the model to study is the Potts model. Here, spins can take $q \geq 2$ values, usually called *colors*. This name comes from the graph coloring problem where the objective is to assign colors to the vertices of a graph in such a way that no two vertices that are connected by an edge have the same color. This is an extreme case of spins having the tendency not to align, or an *antiferromagnetic* interaction. In a social context an antiferromagnetic interaction could mean for example that people prefer to have an opinion different from the persons they dislike.

In the physics literature, the *critical behavior* of many processes on complex networks has been studied, see [42] for an overview. Many of these results have not been mathematically rigorously proved. One of the few models for which rigorous results have been obtained is the contact process [25, 83], where the predictions of physicists, in fact, turned out not to be correct. A mathematical treatment of other models is therefore necessary.

1.5 Contributions and overview

In Part I of this thesis we focus on the ferromagnetic Ising model on power-law random graphs and in Part II on the antiferromagnetic Potts model on the Erdős-Rényi random graph. First we define the random graph models in Chapter 2 and the spin models and the associated thermodynamics in Chapter 3.

The first part starts with Chapters 4 and 5, which are based on [38]. In these chapters we derive an explicit formula for the thermodynamic limit of the *pressure*, generalizing the result of Dembo and Montanari in [34] to random graphs with power-law exponent $\tau > 2$. For such graphs the variance of the degrees is potentially infinite, a case which is not covered in [34]. Also the thermodynamic limits of the magnetization, internal energy and susceptibility are identified. The proofs rely on the locally tree-like structure of the random graph. This means that the local neighborhood of a randomly chosen vertex behaves like a branching process. Correlation inequalities are used to show that it suffices to study the behavior of the Ising model on these branching processes to obtain the results for the random graph.

As an intermezzo, we present new results in Chapter 6 on what happens when the mean degree is infinite, i.e., when the power-law exponent $\tau \in (1,2)$. In this case, the graph is not locally-tree like anymore. Still we can say something about the magnetization in this model.

Chapters 7, 8 and 9 are based on [39]. In these chapters we study the phase transition for the case where $\tau > 2$. We identify the *critical temperature* in Chapter 7 and compute the *critical exponents* $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ describing the critical behavior of the magnetization in Chapter 8 and the critical exponent $\boldsymbol{\gamma}$ describing the critical behavior of the susceptibility in Chapter 9.

To compute the critical temperature and critical exponents we derive upper and lower bounds on the magnetization and susceptibility. These bounds are essentially Taylor approximations, but for power-law exponents $\tau \leq 5$ a more detailed analysis is necessary.

In Part II we turn to the antiferromagnetic Potts model on the Erdős-Rényi random graph. For the antiferromagnet, the correlation inequalities are no longer valid, making this model much harder to analyze. Instead, we use an interpolation scheme to show that the thermodynamic limit of the pressure exists in Chapter 10, which is based on [27]. In Chapter 11, also based on [27], we derive an *extended variational principle* and use this to give two upper bounds on the pressure: the high-temperature solution and the replica-symmetric solution. Finally, in Chapter 12, which is based on the newer version [28], we employ a constrained second-moment method to show that the high-temperature solution is indeed correct for high enough temperatures. We also prove that this solution cannot be correct for low temperatures by showing that the entropy becomes negative if it were to be correct, thus identifying a *phase transition*.

2

RANDOM GRAPHS

In this chapter we introduce three examples of random graph models: the *Erdős-Rényi random graph*, the *inhomogeneous random graph* and the *configuration model* (CM). For an extensive overview of properties of these, and other, random graph models, see [63].

2.1 Erdős-Rényi random graphs

Fix $p \in [0,1]$. Then, the Erdős-Rényi random graph sequence $(G_N^{\mathrm{ER}}(p))_{N\geq 1}$ is constructed as follows. The vertex set of $G_N^{\mathrm{ER}}(p)$ is given by $V_N = [N] \equiv \{1,\dots,N\}$. For every pair $i,j \in [N]$, $i \neq j$, the edge $\{i,j\} \in E_N$ with probability p and is not in the edge set with probability $1-p$.

Often, we choose $p = c/N$. In this case, the degree of a random vertex in the graph has distribution Bin$(N-1, c/N)$, which converges to a Poisson(c) random variable in the limit $N \to \infty$.

The above construction is due to Gilbert [56]. This model differs slightly from the model introduced by Erdős and Rényi in [46] where the graph is constructed by randomly selecting pN edges from the possible $\binom{N}{2}$ edges. Hence, the number of edges in their model is fixed instead of binomial as in our setting.

A third version of the Erdős-Rényi random graph is constructed as follows. Let $(J_{i,j})_{i,j\in[N]}$ be i.i.d. Poisson random variables with parameter $c/2N$. Then, the number of edges between vertices i and j is $J_{i,j} + J_{j,i}$ for $i \neq j$ and $J_{i,i}$ for $i = j$. We denote the resulting graph by $G_N^{\mathrm{ERP}}(c)$). This construction might not produce a simple graph, because self-loops and multiple edges might occur. We can explicitly compute the probability that the graph is simple:

$$\begin{aligned}
&\mathbb{P}[G_N^{\mathrm{ERP}}(c) \text{ is simple}] = \mathbb{P}[\text{no self-loops}]\mathbb{P}[\text{no multiple edges}] \\
&\quad = \mathbb{P}[\text{no self-loops at vertex 1}]^N \mathbb{P}[\text{at most 1 edge between vertices 1 and 2}]^{N(N-1)/2} \\
&\quad = e^{-\frac{c}{2N}N}\left(e^{-\frac{c}{N}} + \frac{c}{N}e^{-\frac{c}{N}}\right)^{N(N-1)/2} \longrightarrow e^{-c/2-c^2/4}, \qquad (2.1)
\end{aligned}$$

for $N \to \infty$. It can in fact be shown that the number of self-loops and multiple edges converge to independent Poisson random variables with parameters $c/2$ and $c^2/4$, respectively. Hence, the number of self-loops and multiple edges will be $O_{\mathbb{P}}(1)$.

It is this last model that we use in Part II. It will become clear there, why this is useful.

2.2 Inhomogeneous random graphs

A natural way to make the graph more inhomogeneous is to give different weights to the vertices in the graph and then conditionally on these weights connect a pair of vertices with a probability depending on these weights, see [24] for results on a very general setting of such models. We focus on the *generalized random graph* which is constructed as follows. Let the weights $(W_i)_{i\in[N]}$ be a sequence of *independent and identically distributed* (i.i.d.) random variables with some distribution W. Then, conditionally on these weights, the probability that the edge $\{i,j\}$ is in the edge set is equal to

$$p_{ij} = \frac{W_i W_j}{W_i W_j + \sum_{k\in[N]} W_k}. \tag{2.2}$$

When these weights are chosen such that they obey a power law with exponent τ, then also the degrees obey a power law with this exponent [63].

2.3 Configuration model

In the configuration model an undirected random graph with N vertices is constructed as follows. Let $(D_i)_{i\in[N]}$ be a sequence of i.i.d. random variables with distribution D for some distribution D on the nonnegative integers. Let vertex $i \in [N]$ be a vertex with D_i half-edges, also called *stubs*, attached to it, i.e., vertex i has degree D_i. Let $L_N = \sum_{i=1}^N D_i$ be the total degree, which we assume to be even in order to be able to construct a graph. When L_N is odd we increase the degree of D_N by 1. For N large, this will hardly change the results and we therefore ignore this effect.

Now connect one of the half-edges uniformly at random to one of the remaining $L_N - 1$ half-edges. Repeat this procedure until all half-edges have been connected. We denote the resulting graph by $G_N^{\mathrm{CM}}\left((D_i)_{i\in[N]}\right)$.

Note that the above construction will not necessarily result in a simple graph. Both self-loops and multiple edges may occur. When $\mathbb{E}[D^2] < \infty$, the probability that the graph is simple can be explicitly calculated in the limit $N \to \infty$ [63]:

$$\mathbb{P}\left[G_N^{\mathrm{CM}}\left((D_i)_{i\in[N]}\right) \text{ is simple}\right] \longrightarrow e^{-\nu/2-\nu^2/4}, \tag{2.3}$$

where

$$\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}. \tag{2.4}$$

Note that this is the same as for the Poissonian Erdős-Rényi random graph for which $\nu = c$. In fact the number of self-loops and multiple edges, in the limit $N \to \infty$, converge in distribution to independent Poisson random variables with means $\nu/2$ and $\nu^2/4$, respectively.

Two ways to get a simple graph are to delete all loops and multiple edges, this model is also called the *erased* configuration model, and to perform the configuration model until it produces a simple graph, which is also called the *repeated* configuration model. For the latter, only a finite number of tries is necessary with probability 1 when $\mathbb{E}[D^2] < \infty$.

2.4 Degree distributions

Let the random variable D have distribution $P = (p_k)_{k\geq 1}$, i.e., $\mathbb{P}[D = k] = p_k$, for $k = 1, 2, \ldots$. We define its *forward degree distribution* by

$$\rho_k = \frac{(k+1)p_{k+1}}{\mathbb{E}[D]}, \tag{2.5}$$

where we assume that $\mathbb{E}[D] < \infty$. Let K be a random variable with $\mathbb{P}[K = k] = \rho_k$ and write $\nu = \mathbb{E}[K] = \mathbb{E}[D(D-1)]/\mathbb{E}[D]$. Moreover, for a probability distribution $(q_k)_{k\geq 0}$ on the non-negative integers, we write $q_{\geq k} = \sum_{\ell\geq k} q_\ell$.

In Part I of this thesis we need to make assumptions on the degree distribution. The first assumption needed in Chapters 4 and 5 is that the degree distribution has *strongly finite mean*, which is defined as follows:

Definition 2.1 (Strongly finite mean degree distribution). *We say that the degree distribution P has* strongly finite mean *when there exist constants $\tau > 2$ and $C_p > 0$ such that*

$$p_{\geq k} \leq C_p k^{-(\tau-1)}. \tag{2.6}$$

For technical reasons, we assume in Chapters 4 and 5, without loss of generality, that $\tau \in (2, 3)$. Note that all distributions P where

$$p_{\geq k} = C_p k^{-(\tau-1)} L(k), \tag{2.7}$$

for $C_p > 0, \tau > 2$ and some slowly varying function $L(k)$, have *strongly finite mean*, because by Potter's theorem [52, Lemma 2, p.277] any slowly varying function $L(k)$ can be bounded above and below by an arbitrary small power of k. Also distributions which have a lighter tail than a power law, e.g., the Poisson distribution, have *strongly finite mean*.

In Chapters 8 and 9 we need to be more precise about the degree distribution. We pay special attention to the case where the degree distribution precisely obeys a *power law* as defined in the following definition. Our results turn out to depend sensitively on the exact value of the power-law exponent.

Definition 2.2 (Power law). *We say that the distribution $P = (p_k)_{k\geq 1}$ obeys a* power law with exponent τ *when there exist constants $C_p \geq c_p > 0$ such that, for all $k = 1, 2, \ldots$,*

$$c_p k^{-(\tau-1)} \leq p_{\geq k} \leq C_p k^{-(\tau-1)}. \tag{2.8}$$

2.5 Local tree-likeness

In Part I we often assume that the graph sequence is *locally tree-like* and *uniformly sparse*. We now define these notions formally.

The random rooted tree $\mathscr{T}(D, K, \ell)$ is a branching process with ℓ generations, where the root offspring is distributed as D and the vertices in each next generation have offsprings that are independent of the root offspring and are i.i.d. copies of the random variable K. We write $\mathscr{T}(K, \ell)$ when the offspring at the root has the same distribution as K.

We write that an event $\mathscr{A}$ holds *almost surely* (a.s.) if $\mathbb{P}[\mathscr{A}] = 1$. The ball of radius r around vertex i, $B_i(r)$, is defined as the graph induced by the vertices at graph distance at most r from vertex i. For two rooted trees $\mathscr{T}_1$ and $\mathscr{T}_2$, we write that $\mathscr{T}_1 \simeq \mathscr{T}_2$, when there exists a bijective map from the vertices of $\mathscr{T}_1$ to those of $\mathscr{T}_2$ that preserves the adjacency relations.

Definition 2.3 (Local convergence to homogeneous random trees). *Let $\mathbb{P}_N$ denote the law induced on the ball $B_i(t)$ in G_N centered at a uniformly chosen vertex $i \in [N]$. We say that the graph sequence $(G_N)_{N\geq 1}$ is* locally tree-like *with asymptotic degree distributed as D when, for any rooted tree $\mathscr{T}$ with t generations*

$$\lim_{N\to\infty} \mathbb{P}_N[B_i(t) \simeq \mathscr{T}] = \mathbb{P}[\mathscr{T}(D,K,t) \simeq \mathscr{T}]. \tag{2.9}$$

Note that this implies that the degree of a uniformly chosen vertex of the graph has an asymptotic degree distributed as D.

Definition 2.4 (Uniform sparsity). *We say that the graph sequence $(G_N)_{N\geq 1}$ is* uniformly sparse *when, a.s.,*

$$\lim_{\ell\to\infty} \limsup_{N\to\infty} \frac{1}{N} \sum_{i\in[N]} D_i \mathbb{1}_{\{D_i \geq \ell\}} = 0, \tag{2.10}$$

where D_i is the degree of vertex i and $\mathbb{1}_{\mathscr{A}}$ denotes the indicator of the event $\mathscr{A}$.

Note that uniform sparsity follows if $\frac{1}{N}\sum_{i\in[N]} D_i \to \mathbb{E}[D]$ a.s., by the weak convergence of the degree of a uniform vertex. An immediate consequence of the local convergence and the uniform sparsity condition is, that, a.s.,

$$\begin{aligned}
\lim_{N\to\infty} \frac{|E_N|}{N} &= \lim_{N\to\infty} \frac{1}{2N} \sum_{i\in[N]} \sum_{k=1}^{\infty} k \mathbb{1}_{\{D_i=k\}} \\
&= \frac{1}{2} \lim_{\ell\to\infty} \lim_{N\to\infty} \left(\sum_{k=1}^{\ell-1} k \frac{\sum_{i\in[N]} \mathbb{1}_{\{D_i=k\}}}{N} + \frac{1}{N} \sum_{i\in[N]} D_i \mathbb{1}_{\{D_i\geq\ell\}} \right) \\
&= \frac{1}{2} \lim_{\ell\to\infty} \sum_{k=1}^{\ell-1} k P_k = \mathbb{E}[D]/2 < \infty.
\end{aligned} \tag{2.11}$$

The Erdős-Rényi random graph, the configuration model and the inhomogeneous random graph all produce locally tree-like and uniformly sparse graph sequences. For the first two models this is proved in [35]. This proof can be adapted for the last model.

It is important to note, however, that random trees, e.g., the random tree $\mathscr{T}(D,K,\ell)$, is *not* locally tree-like according to the definition above. This is because when selecting a random vertex from a (finite) tree, it is very likely that this is a vertex close to the boundary. This, for example, means that in the local neighborhood of a randomly selected vertex many branches of the tree will die out very soon.

2.6 Giant component

One immediate result that can be obtained from the local tree-likeness and the uniform sparsity is the size of the largest connected component, denoted by $\mathscr{C}_{\max}$:

Theorem 2.5 (Size of the giant component). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P and is uniformly sparse. Let D have distribution P and let θ be the survival probability of the tree $\mathscr{T}(D,K,\infty)$, i.e.,*

$$\theta = \mathbb{P}[|\mathscr{T}(D,K,\infty)| = \infty]. \tag{2.12}$$

Then,

$$\frac{|\mathscr{C}_{\max}|}{N} \xrightarrow{p} \theta. \tag{2.13}$$

This theorem was proved in [68]. It is well known that $\theta > 0$ if and only if (*iff*) $\nu > 1$. Hence, there is a giant component, i.e., a connected component of size $\Theta(N)$ in the graph iff $\nu > 1$.

2.7 Degree properties

We now prove some properties of degree sequences that obey a power-law. We first show that when the degree distribution obeys a power law, also the forward degree distribution obeys a power law, where the power-law exponent is one less:

Lemma 2.6 (Tail probabilities of $(\rho_k)_{k\geq 0}$). *Assume that* (2.8) *holds for some* $\tau > 2$. *Then, for the forward degree distribution defined in (2.5), there exist* $0 < c_\rho \leq C_\rho$ *such that, for all* $k \geq 1$,

$$c_\rho k^{-(\tau-2)} \leq \rho_{\geq k} \leq C_\rho k^{-(\tau-2)}. \tag{2.14}$$

Proof. The lower bound follows directly from the fact that $\rho_{\geq k} \geq (k+1)p_{\geq k+1}/\mathbb{E}[D]$, and (2.8). For the upper bound, we note that for any probability distribution $(q_k)_{k\geq 0}$ on the non-negative integers, we have the partial summation identity

$$\sum_{k\geq 0} q_k f(k) = f(0) + \sum_{\ell\geq 1} q_{\geq \ell}[f(\ell) - f(\ell-1)], \tag{2.15}$$

provided that either $f(k)q_{\geq k} \to 0$ as $k \to \infty$ or $k \mapsto f(k)$ is either non-decreasing or non-increasing. Indeed,

$$\sum_{k\geq 0} q_k f(k) = f(0) + \sum_{k\geq 0} q_k[f(k) - f(0)] = f(0) + \sum_{k\geq 0} q_k \sum_{\ell=1}^{k}[f(\ell) - f(\ell-1)]. \tag{2.16}$$

Interchanging the summation order (which is allowed by our assumptions) provides the proof.

We start by proving bounds on $\rho_{\geq k}$. We rewrite

$$\rho_{\geq k} = \sum_{\ell\geq k} \frac{(\ell+1)p_{\ell+1}}{\mathbb{E}[D]} = \sum_{\ell\geq 0} f(\ell)p_{\ell+1}, \tag{2.17}$$

where $f(\ell) = (\ell+1)\mathbb{1}_{\{\ell\geq k\}}/\mathbb{E}[D]$. By (2.15) with $q_\ell = p_{\ell+1}$, for $k \geq 1$ so that $f(0) = 0$,

$$\rho_{\geq k} = \sum_{\ell\geq 1}[f(\ell) - f(\ell-1)]p_{\geq \ell+1} = \frac{(k+1)}{p_{\geq k+1}}\mathbb{E}[D] + \frac{1}{\mathbb{E}[D]} \sum_{\ell\geq k+1} p_{\geq \ell+1}. \tag{2.18}$$

From (2.8), it follows that

$$\rho_{\geq k} \leq \frac{C_p}{\mathbb{E}[D]}(k+1)^{-(\tau-2)} + \sum_{\ell \geq k+1} \frac{C_p}{\mathbb{E}[D]}(\ell+1)^{-(\tau-1)}, \tag{2.19}$$

so that there exists a constant C_ρ such that

$$\rho_{\geq k} \leq C_\rho k^{-(\tau-2)}. \tag{2.20}$$

□

We often split the analysis into two cases: one where the degrees are small and one where the degrees are large. For this we need bounds on truncated moments which are the content of the next lemma.

Lemma 2.7 (Truncated moments of D). *Assume that* (2.8) *holds for some* $\tau > 1$. *Then there exist constants* $C_{a,\tau} > 0$ *such that, as* $\ell \to \infty$,

$$\mathbb{E}\left[D^a \mathbb{1}_{\{D\leq \ell\}}\right] \leq \begin{cases} C_{a,\tau}\ell^{a-(\tau-1)} & \text{when } a > \tau - 1, \\ C_{\tau-1,\tau}\log \ell & \text{when } a = \tau - 1. \end{cases} \tag{2.21}$$

and, when $a < \tau - 1$,

$$\mathbb{E}\left[D^a \mathbb{1}_{\{D>\ell\}}\right] \leq C_{a,\tau}\ell^{a-(\tau-1)}. \tag{2.22}$$

Proof. We start by bounding the truncated moments of D. We rewrite, using (2.15) and with $f(k) = k^a \mathbb{1}_{\{k\leq \ell\}}$,

$$\mathbb{E}\left[D^a \mathbb{1}_{\{D\leq \ell\}}\right] = \sum_{k=0}^{\infty} f(k)p_k = \sum_{k=1}^{\infty}[f(k)-f(k-1)]p_{\geq k} \leq \sum_{k=1}^{\lfloor \ell \rfloor}[k^a-(k-1)^a]p_{\geq k}. \tag{2.23}$$

Using $k^a - (k-1)^a = a\int_{k-1}^{k} x^{a-1}dx \leq ak^{a-1}$, we arrive at

$$\mathbb{E}\left[D^a \mathbb{1}_{\{D\leq \ell\}}\right] \leq aC_p \sum_{k=1}^{\lfloor \ell \rfloor} k^{a-1}k^{-(\tau-1)} \leq aC_p \sum_{k=1}^{\lfloor \ell \rfloor+1} k^{a-\tau}. \tag{2.24}$$

Note that $k \mapsto k^{a-\tau}$ is either increasing or decreasing. Hence,

$$\sum_{k=1}^{\lfloor \ell \rfloor+1} k^{a-\tau} \leq \int_1^{\ell+2} k^{a-\tau}\mathrm{d}k. \tag{2.25}$$

For $a > \tau - 1$,

$$\int_1^{\ell+2} k^{a-\tau}\mathrm{d}k \leq \frac{2}{a+1-\tau}\ell^{a-(\tau-1)}, \tag{2.26}$$

whereas for $a = \tau - 1$,

$$\int_1^{\ell+2} k^{a-\tau}\mathrm{d}k \leq 2\log \ell. \tag{2.27}$$

Similarly, for $a < \tau - 2$,

$$\mathbb{E}\left[D^a \mathbb{1}_{\{D>\ell\}}\right] = \lceil \ell \rceil^a p_{\geq \ell} + \sum_{k>\ell} [k^a - (k-1)^a] p_{\geq k} \tag{2.28}$$

$$\leq C_p \lceil \ell \rceil^{a-(\tau-1)} + a C_p \sum_{\lfloor \ell \rfloor + 1}^{\infty} k^{a-1} (k+1)^{-(\tau-1)} \leq C_{a,\tau} \ell^{a-(\tau-1)}.$$

□

Because of Lemma 2.6, the same statements hold for the truncated moments of K when τ is replaced by $\tau - 1$.

3

SPIN MODELS

In this chapter we formally introduce the spin models that are studied in this thesis. We start by defining the Ising model in Section 3.1. Then, in Section 3.2, we introduce various thermodynamic quantities, the critical temperature and the critical exponents quantifying the behavior of the thermodynamic quantities near this phase transition. In Section 3.4, we state two important correlation inequalities for the ferromagnetic Ising model. Finally, we introduce the Potts model in Section 3.5 and discuss phase transitions in this setting.

3.1 Ising model

We start by defining Ising models on finite graphs. Consider a graph $G_N = (V_N, E_N)$, with vertex set $V_N = [N]$ and with edge set E_N. To each vertex $i \in [N]$ an Ising spin $\sigma_i = \pm 1$ is assigned. A configuration of spins is denoted by $\sigma = (\sigma_i)_{i \in [N]}$. The *Ising model on G_N* is then defined by the Boltzmann-Gibbs measure

$$\mu_N(\sigma) = \frac{1}{Z_N(\beta, \underline{B})} \exp\left\{\beta \sum_{(i,j)\in E_N} J_{i,j}\sigma_i\sigma_j + \sum_{i\in[N]} B_i\sigma_i\right\}. \tag{3.1}$$

Here, $\beta \geq 0$ is the inverse temperature and $\underline{B}$ the vector of external magnetic fields $\underline{B} = (B_i)_{i\in[N]} \in \mathbb{R}^N$. For a uniform external field we write B instead of $\underline{B}$, i.e., $B_i = B$ for all $i \in [N]$.

Note that the inverse temperature β does not multiply the external field. This turns out to be technically convenient and does not change the results, because we are only looking at systems at equilibrium, and hence this would just be a reparametrization.

If $J_{i,j} \geq 0$ for all $(i,j) \in E_N$ we call the model *ferromagnetic* and if $J_{i,j} \leq 0$ for all $(i,j) \in E_N$ we call the model *antiferromagnetic*. Models where the $J_{i,j}$ have mixed signs are also possible. If the interactions are random with a distribution that is symmetric around zero, the model is called a *spin glass*. Models with mixed signs are not considered in this thesis. In Chapters 4 until 9 we focus on the ferromagnetic Ising model with $J_{i,j} = +1$ if $(i,j) \in E_N$.

The *partition function* $Z_N(\beta, \underline{B})$ is the normalization constant in (3.1), i.e.,

$$Z_N(\beta, \underline{B}) = \sum_{\sigma \in \{-1,+1\}^N} \exp\left\{ \beta \sum_{(i,j) \in E_N} J_{i,j} \sigma_i \sigma_j + \sum_{i \in [N]} B_i \sigma_i \right\}. \tag{3.2}$$

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

$$\langle f(\sigma) \rangle_{\mu_N} = \sum_{\sigma \in \{-1,+1\}^N} f(\sigma) \mu_N(\sigma). \tag{3.3}$$

We are often interested in the *thermodynamic limit* of the Ising model, which we interpret as follows: for a graph sequence $(G_N)_{N \geq 1}$, the behavior in the thermodynamic limit is the behavior of the Ising model on G_N as $N \to \infty$, provided that this is properly defined. Note that this is a bit different from the setting where the Ising model is studied on $\mathbb{Z}^d$. There, it is customary to study the Ising model on a box $[-N, N]^d$ with some specified boundary conditions and then let $N \to \infty$. In our setting there is no such thing as a boundary.

We next extend our definitions to the *random Bethe tree* $\mathcal{T}(D, K, \infty)$, which is the limit $\ell \to \infty$ of the random tree $\mathcal{T}(D, K, \ell)$ as defined in Section 2.5. One has to be very careful in defining a Boltzmann-Gibbs measure on this tree, since trees suffer from the fact that the boundaries of intrinsic (i.e., graph distance) balls in them have a size that is comparable to their volume. We can adapt the construction of the Ising model on the regular tree in [16] to this setting, as we now explain. For $\beta \geq 0, B > 0$, let $\mu_{\beta,B}^{t,+/f}$ be the Ising model on $\mathcal{T}(D, K, t)$ with $+$ respectively free boundary conditions. For a function f that only depends on $\mathcal{T}(D, K, m)$ with $m \leq t$, we let

$$\langle f \rangle_{\mu_{\beta,B}^{+/f}} = \lim_{t \to \infty} \langle f \rangle_{\mu_{\beta,B}^{t,+/f}}. \tag{3.4}$$

In Chapter 4 we show that $\langle f \rangle_{\mu_{\beta,B}^{+}} = \langle f \rangle_{\mu_{\beta,B}^{f}}$, i.e., the behavior of quantities that depend only on spins far away from the boundary is the same for all nonnegative boundary conditions.

3.2 Thermodynamics

To study the Ising model we investigate several thermodynamic quantities. We first define these quantities in finite volume, i.e., for graphs G_N with $N < \infty$. The first quantity of interest is the *pressure*:

Definition 3.1 (Pressure per vertex). *For a graph G_N, the* pressure per vertex *is defined as*

$$\psi_N(\beta, B) = \frac{1}{N} \log Z_N(\beta, B). \tag{3.5}$$

Since the graph G_N might be random, we can take the expectation over the random graph to obtain the *quenched pressure*:

Definition 3.2 (Quenched pressure per vertex). *For a random graph G_N, the* quenched pressure per vertex *is defined as*

$$p_N(\beta,B) = \frac{1}{N}\mathbb{E}\left[\log Z_N(\beta,B)\right]. \tag{3.6}$$

Furthermore, we are interested in the following thermodynamic quantities:

Definition 3.3 (Thermodynamic quantities). *For a graph G_N,*

(a) the magnetization per vertex *equals*

$$M_N(\beta,B) = \frac{1}{N}\sum_{i\in[N]} \langle \sigma_i \rangle_{\mu_N}. \tag{3.7}$$

(b) the susceptibility *equals*

$$\chi_N(\beta,B) = \frac{1}{N}\sum_{i,j\in[N]} \left(\left\langle \sigma_i\sigma_j \right\rangle_{\mu_N} - \langle \sigma_i \rangle_{\mu_N} \left\langle \sigma_j \right\rangle_{\mu_N} \right). \tag{3.8}$$

(c) the internal energy *equals*

$$U_N(\beta,B) = -\frac{1}{N}\sum_{(i,j)\in E_N} \left\langle \sigma_i\sigma_j \right\rangle_{\mu_N}. \tag{3.9}$$

(d) the specific heat *equals*

$$C_N(\beta,B) = -\beta^2 \frac{\partial}{\partial\beta} U_N(\beta,B). \tag{3.10}$$

(e) the entropy *equals*

$$S_N(\beta,B) = -\frac{1}{N}\sum_{\sigma\in\{-1,+1\}^N} \mu_N(\sigma)\log\mu_N(\sigma). \tag{3.11}$$

We are often interested in the *thermodynamic limit* of these quantities, i.e., for the limit $N\to\infty$. For all these quantities, we drop the subscript N for the thermodynamic limit of that quantity, e.g., $M(\beta,B) = \lim_{N\to\infty} M_N(\beta,B)$, provided this limit exists.

When speaking about the magnetization of the random Bethe tree, we mean the expectation of the root magnetization, i.e., for the random Bethe tree with root ϕ,

$$M(\beta,B) = \mathbb{E}\left[\langle\sigma_\phi\rangle_{\mu_{\beta,B}^{+/f}}\right]. \tag{3.12}$$

3.3 Critical behavior

It is well known that when an external magnetic field is applied to a piece of iron the iron becomes magnetic itself. When the temperature is sufficiently low, this iron keeps its magnetic property even when the external magnetic field is removed. When the temperature is too high, however, the magnetism is lost when the external magnetic field is removed. The temperature that separates these two regimes is known as the Curie temperature.

Similar behavior can be proved for the Ising model, when looking at the *spontaneous magnetization*, which is defined as $M(\beta, 0^+)$, where we write $f(0^+)$ for $\lim_{x\searrow 0} f(x)$. The *critical temperature* is then defined as

Definition 3.4 (Critical temperature)**.** *The critical temperature equals*

$$\beta_c \equiv \inf\{\beta : M(\beta, 0^+) > 0\}. \tag{3.13}$$

Note that such a β_c can only exist in the thermodynamic limit, but not for the magnetization of a finite graph, since always $M_N(\beta, 0^+) = 0$. When $0 < \beta_c < \infty$, we say that the system undergoes a *phase transition* at $\beta = \beta_c$ and $B = 0^+$, because the thermodynamic limit of the pressure is non-analytic in this point.

The critical behavior can now be expressed in terms of the following critical exponents. We write $f(x) \asymp g(x)$ if the ratio $f(x)/g(x)$ is bounded away from 0 and infinity for the specified limit.

Definition 3.5 (Critical exponents)**.** *The critical exponents $\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma$, and γ' are defined by:*

$$M(\beta, 0^+) \asymp (\beta - \beta_c)^{\boldsymbol{\beta}}, \qquad \text{for } \beta \searrow \beta_c; \tag{3.14}$$

$$M(\beta_c, B) \asymp B^{1/\boldsymbol{\delta}}, \qquad \text{for } B \searrow 0; \tag{3.15}$$

$$\chi(\beta, 0^+) \asymp (\beta_c - \beta)^{-\gamma}, \qquad \text{for } \beta \nearrow \beta_c; \tag{3.16}$$

$$\chi(\beta, 0^+) \asymp (\beta - \beta_c)^{-\gamma'}, \qquad \text{for } \beta \searrow \beta_c; \tag{3.17}$$

Remark 3.6. Note that there is a difference between the symbol β for the inverse temperature and the bold symbol $\boldsymbol{\beta}$ for the critical exponent in (3.14). Both uses for β are standard in the literature, so we decided to stick to this notation.

Also note that these are stronger definitions than usual. E.g., normally the critical exponent $\boldsymbol{\beta}$ is defined as that value such that

$$M(\beta, 0^+) = (\beta - \beta_c)^{\boldsymbol{\beta} + o(1)}, \tag{3.18}$$

where $o(1)$ is a function tending to zero for $\beta \searrow \beta_c$.

3.4 Correlation inequalities

To analyze the ferromagnetic Ising model we make use of two important correlation inequalities. Note that these inequalities only hold for the ferromagnetic Ising model, but not for the antiferromagnet and spin glasses. The first result on ferromagnetic Ising models we heavily rely on is the *Griffiths, Kelly, Sherman (GKS) inequality*, which gives various monotonicity properties:

Lemma 3.7 (GKS inequality)**.** *Consider two Ising measures μ and μ' on graphs $G = (V, E)$ and $G' = (V, E')$, with inverse temperatures β and β' and external fields $\underline{B}$ and $\underline{B}'$, respectively. If $E \subseteq E'$, $0 \le \beta \le \beta'$ and $0 \le B_i \le B_i'$ for all $i \in V$, then, for any $U \subseteq V$,*

$$0 \le \Big\langle \prod_{i \in U} \sigma_i \Big\rangle_\mu \le \Big\langle \prod_{i \in U} \sigma_i \Big\rangle_{\mu'}. \tag{3.19}$$

A weaker version of this inequality was first proved by Griffiths [57] and later generalized by Kelly and Sherman [71]. The second result on ferromagnetic Ising models is an inequality by Griffiths, Hurst and Sherman [58] which shows the concavity of the magnetization in the external (positive) magnetic fields.

Lemma 3.8 (GHS inequality). *Let $\beta \geq 0$ and $B_i \geq 0$ for all $i \in V$. Denote by*

$$m_j(\underline{B}) = \mu(\{\sigma : \sigma_j = +1\}) - \mu(\{\sigma : \sigma_j = -1\}) \tag{3.20}$$

the magnetization of vertex j when the external fields at the vertices are $\underline{B}$. Then, for any three vertices $j, k, \ell \in V$,

$$\frac{\partial^2}{\partial B_k \partial B_\ell} m_j(\underline{B}) \leq 0. \tag{3.21}$$

We, for example, use these correlation inequalities in Chapter 4 to show that the effect of nonnegative boundary conditions on the root magnetization of a tree diminishes when the boundary converges to infinity.

3.5 Potts model

The Ising model can be generalized to the *Potts model*. For this, fix an integer $q \geq 2$ and assign to each vertex $i \in [N]$ a Potts spin $\sigma_i \in [q]$. The *Potts model on G_N* is then defined by the Boltzmann-Gibbs measure

$$\mu_N(\sigma) = \frac{1}{Z_N(\beta, \underline{B})} \exp\left\{\beta \sum_{(i,j)\in E_N} J_{i,j}\delta(\sigma_i, \sigma_j) + \sum_{i\in[N]} B_i \delta(\sigma_i, 1)\right\}, \tag{3.22}$$

where $\delta(r,s)$ is the Kronecker delta, i.e.,

$$\delta(r,s) = \begin{cases} 1, & \text{when } r = s, \\ 0, & \text{otherwise,} \end{cases} \tag{3.23}$$

and the other definitions are equivalent to the Ising model.

Note that, for $q = 2$, we can alternatively choose $\sigma_i = \pm 1$. Then,

$$\delta(\sigma_i, \sigma_j) = \frac{\sigma_i \sigma_j + 1}{2}. \tag{3.24}$$

Hence, the Potts model with $q = 2$ is equivalent to the Ising model with $\beta_{\text{Ising}} = \beta_{\text{Potts}}/2$, because the constant term $+1/2$ cancels in the partition function.

In Chapters 10–12 we focus on the antiferromagnetic Potts model with $B = 0$. This model is related to the graph coloring problem, where the problem is to find if the vertices of a graph can be colored with q colors in such a way that no two adjacent vertices have the same color. Such a graph coloring would correspond to a ground state of the antiferromagnetic Potts model with $J_{i,j} = -1$. Several phase transitions are predicted in the physics literature for the antiferromagnetic Potts model, both at positive temperature [72] and at zero temperature [102]. We concentrate on the phase transition with the highest temperature. We first compute an analytic expression for the pressure at high temperature, which equals

$$p^{\mathrm{HT}}(\beta) = \frac{c}{2} \log\left(1 - \frac{1 - e^{-\beta}}{q}\right) + \log q, \tag{3.25}$$

and then define

$$\beta_c = \inf\left\{\beta : p(\beta) \neq p^{\mathrm{HT}}(\beta)\right\}. \tag{3.26}$$

Part I

Ferromagnetic Ising model on power-law random graphs

4
TREE RECURSION

We now start the analysis of the Ising model on locally tree-like random graphs. To analyze this model, we first focus on local observables, for example the magnetization of a uniformly selected vertex or the correlation of two spins connected by a uniformly chosen edge. We can bound these quantities by enforcing boundary conditions on a large ball around such a vertex or edge. By our assumption, the graph inside this ball will be a tree a.s. and it remains to compute the quantity of interest on such a tree.

It is therefore of importance to understand the behavior of the Ising model on trees. In this chapter we show that local observables can be computed using an explicit recursion. Furthermore, we show that in a positive field the effect of nonnegative boundary conditions vanishes as the size of the ball tends to infinity.

4.1 Results

Consider the following distributional recursion

$$h^{(t+1)} \stackrel{d}{=} B + \sum_{i=1}^{K_t} \xi(h_i^{(t)}), \tag{4.1}$$

where $h^{(0)} \equiv B$, $(K_t)_{t\geq 1}$, are i.i.d. with distribution ρ, $(h_i^{(t)})_{i\geq 1}$ are i.i.d. copies of $h^{(t)}$ independent of K_t and

$$\xi(h) = \operatorname{atanh}(\hat{\beta} \tanh h). \tag{4.2}$$

with

$$\hat{\beta} = \tanh \beta. \tag{4.3}$$

In Section 4.3 we show that $h^{(t)}$ can be seen as the effective field on the root of the random tree $\mathcal{T}(K,t)$ with 'free' boundary conditions.

We next show that for the limit $t \to \infty$, the distributional recursion above has a unique positive fixed point:

Theorem 4.1 (Tree recursion)**.** *Let $B > 0$ and let $(K_t)_{t\geq 1}$ be i.i.d. according to some distribution ρ and assume that $K_1 < \infty$, a.s. Consider the sequence of random variables $(h^{(t)})_{t\geq 0}$ defined by $h^{(0)} \equiv B$ and, for $t \geq 0$, by (4.1). Then, the distributions of $h^{(t)}$ are* *stochastically increasing and* $h^{(t)}$ *converges in distribution to the* unique *fixed point* h^* *of the recursion* (4.1) *that is supported on* $[0,\infty)$.

This theorem establishes that the recursive relation that gives the effective field acting on the root of the infinite tree $\mathscr{T}(D,K,\infty)$ is well-defined if $B>0$, in the sense that the recursion admits a unique positive fixed point h^*.

Note that we do not prove that the fixed point h^* is the unique fixed point of the recursion, only that it is the unique *positive* fixed point. By choosing negative or mixed boundary conditions other fixed points can be obtained. These are not of interest to us, however, because we can use the correlation inequalities to show that we only need to look at nonnegative boundary conditions.

To prove this theorem, we adapt the proof of Dembo and Montanari [34] by taking the actual forward degrees into account, instead of using Jensen's inequality to replace them by expected forward degrees, which are potentially infinite. This also makes a separate analysis of nodes that have zero offspring superfluous, which considerably simplifies the analysis.

4.2 Discussion

Statistical mechanics models on random trees and random graphs. A key idea to analyze the Ising model on random graphs is to use the fact that expectations of local quantities coincide with the corresponding values for the Ising model on suitable random trees [34]. Statistical mechanics models on deterministic trees have been studied extensively in the literature (see for instance [16, 76] and its relation to "broadcasting on trees" in [48, 79]). The analysis on random trees is more recent and has been triggered by the study of models on random graphs.

Potts model. In [37] a recursion similar to (4.1) is derived for the Potts model on the k-regular tree. Recall that in the definition of the Potts model in Section 3.5 only the color 1 is special, in the sense that the magnetic field only favors this color. Hence, the behavior is symmetric in all other colors and a one-dimensional recursion can again be obtained. A straightforward generalization to Galton-Watson trees yields

$$h \stackrel{d}{=} B + \sum_{i=1}^{K} \log \frac{e^{\beta+h_i}+q-1}{e^{h_i}+e^{\beta}+q-2}. \tag{4.4}$$

The behavior of the Potts model is more complex, since there exists a region in the (β, B) space where this recursion does not have a unique positive fixed point for all $B>0$.

XY **model.** So far, the focus has been on models where the spins can only have a discrete number of values, but also continuous spin models exist. An example is the XY model, where spins can have any value in the interval $[0,2\pi)$ and the Boltzmann-Gibbs measure is given by

$$\mu_N(\sigma) = \frac{1}{Z_N} \exp\left\{\beta \sum_{(i,j)\in E_N} J_{i,j}\cos(\sigma_i-\sigma_j) + \sum_{i\in[N]} B_i \cos(\sigma_i)\right\}. \tag{4.5}$$

Again spins prefer to align, but being approximately the same is also favored over being far from each other. This is different for the Potts model where the spins do or do not align, but there is nothing in between. The continuity of the spin values is making it much harder to obtain a recursion formula. This model on random graphs is briefly studied in [42], but rigorous results are not obtained there.

4.3 Pruning trees

In [34, Lemma 4.1], it is shown that one can compute the marginal Ising measure on a subtree of a tree by 'pruning' the tree. That is, one can leave away parts of the tree, which can be compensated by updating the external magnetic field at the remaining vertices of the tree. We present a slightly different statement of this in the next lemma.

Lemma 4.2 (Pruning trees). *Consider a tree $T=(V,E)$ with a distinguished leaf ℓ and let k be such that $(k,\ell)\in E$. Write $V_{-\ell}=V\setminus\{\ell\}$ and $E_{-\ell}=E\setminus\{(k,\ell)\}$ and the corresponding tree by $T_{-\ell}=(V_{-\ell},E_{-\ell})$. Denote the Ising measure on T with fields $(B_i)_{i\in V}$ by μ_T. Then, for all $\sigma_{-\ell}=(\sigma_i)_{i\in V_{-\ell}}$, the marginal Ising measure on $T_{-\ell}$ satisfies*

$$\sum_{\sigma_\ell=\pm1}\mu_T(\sigma_{-\ell},\sigma_\ell)=\mu'_{T_{-\ell}}(\sigma_{-\ell}), \tag{4.6}$$

where $\mu'_{T_{-\ell}}$ is the Ising measure on $T_{-\ell}$ with magnetic fields,

$$B'_i=\begin{cases} B_i & \text{when } i\neq k,\\ B_k+\xi(B_\ell) & \text{when } i=k,\end{cases} \tag{4.7}$$

where $\xi(h)$ is defined in (4.2).

Proof. We can write

$$\mu(\sigma)=\mu(\sigma_{-\ell},\sigma_\ell)=\frac{e^{H(\sigma_{-\ell})}e^{\beta\sigma_k\sigma_\ell+B_\ell\sigma_\ell}}{\sum_{\sigma_{-\ell}}\sum_{\sigma_\ell}e^{H(\sigma_{-\ell})}e^{\beta\sigma_k\sigma_\ell+B_\ell\sigma_\ell}}, \tag{4.8}$$

where

$$H(\sigma_{-\ell})=\beta\sum_{(i,j)\in E_{-\ell}}\sigma_i\sigma_j+\sum_{i\in V_{-\ell}}B_i\sigma_i. \tag{4.9}$$

Then,

$$\sum_{\sigma_\ell=\pm1}\mu_T(\sigma_{-\ell},\sigma_\ell)=\frac{e^{H(\sigma_{-\ell})}\left(e^{\beta\sigma_k+B_\ell}+e^{-\beta\sigma_k-B_\ell}\right)}{\sum_{\sigma_{-\ell}}e^{H(\sigma_{-\ell})}\left(e^{\beta\sigma_k+B_\ell}+e^{-\beta\sigma_k-B_\ell}\right)}. \tag{4.10}$$

Since σ_k can only take two values we can write, for any function $f(\sigma_k)$,

$$f(\sigma_k)=\mathbb{1}_{\{\sigma_k=+1\}}f(1)+\mathbb{1}_{\{\sigma_k=-1\}}f(-1). \tag{4.11}$$

These indicators can be rewritten as $\mathbb{1}_{\{\sigma_k=\pm1\}}=\frac{1}{2}(1\pm\sigma_k)$, so that

$$\begin{aligned} e^{\beta\sigma_k+B_\ell}+e^{-\beta\sigma_k-B_\ell}&=\exp\left(\log(e^{\beta\sigma_k+B_\ell}+e^{-\beta\sigma_k-B_\ell})\right)\\ &=\exp\left(\frac{1}{2}(1+\sigma_k)\log(e^{\beta+B_\ell}+e^{-\beta-B_\ell})+\frac{1}{2}(1-\sigma_k)\log(e^{-\beta+B_\ell}+e^{\beta-B_\ell})\right)\\ &=\exp\left(\sigma_k\frac{1}{2}\log\left(\frac{e^{\beta+B_\ell}+e^{-\beta-B_\ell}}{e^{-\beta+B_\ell}+e^{\beta-B_\ell}}\right)\right)A(\beta,B_\ell), \end{aligned} \tag{4.12}$$

for some constant $A(\beta,B_\ell)$ independent of σ_k. An elementary calculation shows that

$$\frac{1}{2}\log\left(\frac{e^{\beta+B_\ell}+e^{-\beta-B_\ell}}{e^{-\beta+B_\ell}+e^{\beta-B_\ell}}\right)=\operatorname{atanh}(\tanh\beta\tanh B_\ell)=\xi(B_\ell). \tag{4.13}$$

Hence,

$$\sum_{\sigma_\ell=\pm1}\mu_T(\sigma_{-\ell},\sigma_\ell)=\frac{e^{H(\sigma_{-\ell})}e^{\sigma_k\xi(B_\ell)}A(\beta,B_\ell)}{\sum_{\sigma_{-\ell}}e^{H(\sigma_{-\ell})}e^{\sigma_k\xi(B_\ell)}A(\beta,B_\ell)}=\mu'_{T_{-\ell}}(\sigma_{-\ell}). \tag{4.14}$$

□

This lemma can be applied recursively. When studying the Ising model on the random tree T distributed as $\mathscr{T}(K,\ell)$ this leads to the distributional recursion (4.1).

4.4 Uniqueness of the positive fixed point

To prove Theorem 4.1, we first study the Ising model on a tree with ℓ generations, $\mathscr{T}(\ell)$, with either $+$ or free boundary conditions, where the Ising models on the tree $\mathscr{T}(\ell)$ with $+$/free boundary conditions are defined by the Boltzmann-Gibbs measures

$$\mu^{\ell,+}(\sigma)=\frac{1}{Z^{\ell,+}(\beta,\underline{B})}\exp\left\{\beta\sum_{(i,j)\in\mathscr{T}(\ell)}\sigma_i\sigma_j+\sum_{i\in\mathscr{T}(\ell)}B_i\sigma_i\right\}\mathbb{1}_{\{\sigma_i=+1,\text{ for all }i\in\partial\mathscr{T}(\ell)\}}, \tag{4.15}$$

and

$$\mu^{\ell,f}(\sigma)=\frac{1}{Z^{\ell,f}(\beta,\underline{B})}\exp\left\{\beta\sum_{(i,j)\in\mathscr{T}(\ell)}\sigma_i\sigma_j+\sum_{i\in\mathscr{T}(\ell)}B_i\sigma_i\right\}, \tag{4.16}$$

respectively, where $Z^{\ell,+/f}$ are the proper normalization factors and $\partial\mathscr{T}(\ell)$ denotes the vertices in the ℓ-th generation of $\mathscr{T}(\ell)$. In the next lemma we show that the effect of these boundary conditions vanishes when $\ell\to\infty$. This lemma is a generalization of [34, Lemma 4.3], where this result is proved in expectation for graphs with a finite-variance degree distribution. This generalization is possible by taking the degrees into account more precisely, instead of using Jensen's inequality to replace them by average degrees. This also simplifies the proof.

We then show that the recursion (4.1) has a fixed point and use a coupling with the root magnetization in trees and Lemma 4.3 to show that this fixed point does not depend on the initial distribution $h^{(0)}$, as long as the initial distribution is nonnegative, thus showing that (4.1) has a *unique* positive fixed point.

Lemma 4.3 (Vanishing effect of boundary conditions). *Let $m^{\ell,+/f}(\underline{B})$ denote the root magnetization given $\mathscr{T}(\ell)$ with external field per vertex $B_i\geq B_{\min}>0$ when the tree has $+$/free boundary conditions. Assume that the forward degrees satisfy $\Delta_i<\infty$ a.s., for all $i\in\mathscr{T}(\ell-1)$. Let $0\leq\beta\leq\beta_{\max}<\infty$. Then, there exists an $A=A(\beta_{\max},B_{\min})<\infty$ such that, a.s.,*

$$m^{\ell,+}(\underline{B})-m^{\ell,f}(\underline{B})\leq A/\ell, \tag{4.17}$$

for all $\ell\geq1$.

Remark 4.4. Lemma 4.3 is extremely general. For example, it also applies to trees arising from multitype branching processes.

Proof. The lemma clearly holds for $\beta = 0$, so we assume that $\beta > 0$ in the remainder of the proof.

Denote by $m^\ell(\underline{B}, \underline{H})$ the root magnetization given $\mathscr{T}(\ell)$ with free boundary conditions, when the external field on the vertices $i \in \partial\mathscr{T}(\ell)$ is $B_i + H_i$ and B_i on all other vertices $i \in \mathscr{T}(\ell-1)$. Condition on the tree $\mathscr{T}(\ell)$ and assume that the tree $\mathscr{T}(\ell)$ is finite, which is true a.s., so that we can use Lemma 4.2. Thus, for $1 \le k \le \ell$,

$$m^{k,+}(\underline{B}) \equiv m^k(\underline{B}, \infty) = m^{k-1}(\underline{B}, \{\beta\Delta_i\}), \tag{4.18}$$

where Δ_i is the forward degree of vertex $i \in \partial\mathscr{T}(k-1)$. By the GKS inequality

$$m^{k-1}(\underline{B}, \{\beta\Delta_i\}) \le m^{k-1}(\underline{B}, \infty). \tag{4.19}$$

Since the magnetic field at all vertices in $\partial\mathscr{T}(k)$ is at least $B_{\min}$ we can write, using Lemma 4.2 and the GKS inequality, that

$$m^{k,f}(\underline{B}) \equiv m^k(\underline{B}, 0) \ge m^{k-1}(\underline{B}, \xi(B_{\min})\{\Delta_i\}). \tag{4.20}$$

This inequality holds with equality when $B_i = B_{\min}$ for all $i \in \partial\mathscr{T}(k)$. Using the GKS inequality again, we have that

$$m^{k-1}(\underline{B}, \xi(B_{\min})\{\Delta_i\}) \ge m^{k-1}(\underline{B}, 0). \tag{4.21}$$

Note that $0 \le \xi(B_{\min}) \le \beta$. Since $H \mapsto m^k(\underline{B}, H\{\Delta_i\})$ is concave in H because of the GHS inequality, we have that

$$m^{k-1}(\underline{B}, \beta\{\Delta_i\}) - m^{k-1}(\underline{B}, 0) \le A\left(m^{k-1}(\underline{B}, \xi(B_{\min})\{\Delta_i\}) - m^{k-1}(\underline{B}, 0)\right), \tag{4.22}$$

where

$$A = A(\beta_{\max}, B_{\min}) = \sup_{0<\beta\le\beta_{\max}} \frac{\beta}{\xi(B_{\min})} < \infty. \tag{4.23}$$

Thus, we can rewrite $m^{k,+}(\underline{B})$ using (4.18) and bound $m^{k,f}(\underline{B})$ using (4.20) and (4.21), to obtain

$$m^{k,+}(\underline{B}) - m^{k,f}(\underline{B}) \le m^{k-1}(\underline{B}, \beta\{\Delta_i\}) - m^{k-1}(\underline{B}, 0). \tag{4.24}$$

By (4.22), we then have that

$$\begin{aligned} m^{k,+}(\underline{B}) - m^{k,f}(\underline{B}) &\le A\left(m^{k-1}(\underline{B}, \xi(B_{\min})\{\Delta_i\}) - m^{k-1}(\underline{B}, 0)\right) \\ &\le A\left(m^k(\underline{B}, 0) - m^{k-1}(\underline{B}, 0)\right), \end{aligned} \tag{4.25}$$

where we have used (4.20) in the last inequality.

By (4.18) and (4.19), $m^{k,+}(\underline{B})$ is non-increasing in k and, by (4.20) and (4.21), $m^{k,f}(\underline{B})$ is non-decreasing in k. Thus, by summing the inequality in (4.25) over k, we get that

$$\begin{aligned} \ell\left(m^{\ell,+}(\underline{B}) - m^{\ell,f}(\underline{B})\right) &\le \sum_{k=1}^{\ell}\left(m^{k,+}(\underline{B}) - m^{k,f}(\underline{B})\right) \le A\sum_{k=1}^{\ell}\left(m^k(\underline{B}, 0) - m^{k-1}(\underline{B}, 0)\right) \\ &= A\left(m^\ell(\underline{B}, 0) - m^0(\underline{B}, 0)\right) \le A, \end{aligned} \tag{4.26}$$

since $0 \le m^{\ell/0}(\underline{B}, 0) \le 1$. □

We are now ready to prove Theorem 4.1:

Proof of Theorem 4.1. Condition on the tree $\mathscr{T}(K,\infty)$. Then $h^{(t)} \equiv \operatorname{atanh}(m^{t,f}(B))$ satisfies the recursive distribution (4.1) because of Lemma 4.2. Since, by the GKS inequality, $m^{t,f}(B)$, and hence also $h^{(t)}$, are monotonically increasing in t, we have that $B = h^{(0)} \leq h^{(t)} \leq B + K_0 < \infty$ for all $t \geq 0$, where K_0 is the degree of the root. So, $h^{(t)}$ converges to some limit $\underline{h}$. Since this holds a.s. for any tree $\mathscr{T}(K,\infty)$, the distribution of $\underline{h}$ also exists and one can show that this limit is a fixed point of (4.1) (see [34, Proof of Lemma 2.3]).

In a similar way, $h^{(t,+)} \equiv \operatorname{atanh}(m^{t,+}(B))$ satisfies (4.1) when starting with $h^{(0,+)} = \infty$. Then, $h^{(t,+)}$ is monotonically decreasing and, for $t \geq 1$, $B \leq h^{(t)} \leq B + K_0 < \infty$, so $h^{(t,+)}$ also converges to some limit $\bar{h}$.

Let h be a positive fixed point of (4.1), condition on this h and let $h^{(0,*)} = h$. Then $h^{(t,*)}$ converges as above to a limit h^* say, when applying (4.1). Note that $h^{(0)} \leq h^{(0,*)} \leq h^{(0,+)}$. Coupling so as to have the same $(K_t)_{t\geq 1}$ while applying the recursion (4.1), this order is preserved by the GKS inequality, so that $h^{(t)} \leq h^{(t,*)} \leq h^{(t,+)}$ for all $t \geq 0$. By Lemma 4.3,

$$|\tanh(h^{(t)}) - \tanh(h^{(t,+)})| = |m^{t,f}(B) - m^{t,+}(B)| \to 0, \qquad \text{for } t \to \infty. \tag{4.27}$$

Since the above holds a.s. for any tree $\mathscr{T}(K,\infty)$ and any realization of h^*, the distributions of $\underline{h}, \bar{h}$ and h^* are equal, and, since h is a positive fixed point of (4.1), are all equal in distribution to h. □

5

THERMODYNAMIC LIMIT

Using the results of the previous chapter, we can now compute the thermodynamic limit of various thermodynamic quantities, most importantly the pressure per vertex. We give an explicit formula for this in this chapter, using the fixed point of the recursion (4.1). In the proof we first compute the internal energy which can then be integrated with respect to β to obtain the pressure. The magnetization and susceptibility can then be obtained by differentiating the pressure with respect to B.

5.1 Results

An explicit formula for the thermodynamic limit of the pressure is given in the following theorem:

Theorem 5.1 (Thermodynamic limit of the pressure). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P, where P has strongly finite mean, and is uniformly sparse. Then, for all $0 \leq \beta < \infty$ and all $B \in \mathbb{R}$, the thermodynamic limit of the pressure exists, a.s., and equals*

$$\lim_{N\to\infty} \psi_N(\beta, B) = \varphi(\beta, B), \tag{5.1}$$

where, for $B < 0$, $\varphi(\beta, B) = \varphi(\beta, -B)$, $\varphi(\beta, 0) = \lim_{B\searrow 0} \varphi(\beta, B)$ and, for $B > 0$,

$$\begin{aligned}\varphi(\beta, B) =& \frac{\mathbb{E}[D]}{2} \log\cosh(\beta) - \frac{\mathbb{E}[D]}{2} \mathbb{E}[\log(1 + \hat{\beta}\tanh(h_1)\tanh(h_2))] \\ &+ \mathbb{E}\left[\log\left(e^B \prod_{i=1}^{D} \left(1 + \hat{\beta}\tanh(h_i)\right) + e^{-B} \prod_{i=1}^{D} \left(1 - \hat{\beta}\tanh(h_i)\right)\right)\right],\end{aligned} \tag{5.2}$$

where

(i) D has distribution P;

(ii) $(h_i)_{i\geq 1}$ are i.i.d. copies of the positive fixed point $h = h(\beta, B)$ of the distributional recursion (4.1)*;*

(iii) D and $(h_i)_{i\geq 1}$ are independent.

Various thermodynamic quantities can be computed by taking the proper derivative of the function $\varphi(\beta,B)$ as we show in the next theorem.

Theorem 5.2 (Thermodynamic quantities). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P, where P has strongly finite mean, and is uniformly sparse. Then, for all $\beta \geq 0$ and $B \neq 0$, each of the following statements holds a.s.:*

(a) Magnetization. *The thermodynamic limit of the* magnetization *per vertex exists and is given by*

$$M(\beta,B) = \frac{\partial}{\partial B}\varphi(\beta,B). \tag{5.3}$$

(b) Internal energy. *The thermodynamic limit of the* internal energy *per vertex exists and is given by*

$$U(\beta,B) = -\frac{\partial}{\partial \beta}\varphi(\beta,B). \tag{5.4}$$

(c) Susceptibility. *The thermodynamic limit of the* susceptibility *exists and is given by*

$$\chi(\beta,B) = \frac{\partial^2}{\partial B^2}\varphi(\beta,B). \tag{5.5}$$

Another physical quantity studied in the physics literature is the *specific heat,*

$$C_N(\beta,B) \equiv -\beta^2 \frac{\partial U_N}{\partial \beta}. \tag{5.6}$$

Unfortunately, we were not able to prove that this converges to $\beta^2 \frac{\partial^2}{\partial \beta^2}\varphi(\beta,B)$, because we do not have convexity or concavity of the internal energy in β. We expect, however, that this limit also holds.

Taking the derivatives of Theorem 5.2 we can also give explicit expressions for the magnetization and internal energy which have a physical interpretation:

Corollary 5.3 (Explicit expressions for thermodynamic quantities). *Assume that the graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P, where P has strongly finite mean, and is uniformly sparse. Then, for all $\beta \geq 0$ and $B \in \mathbb{R}$, each of the following statements holds a.s.:*

(a) Magnetization. *Let D have distribution P and let ν_{D+1} be the random Ising measure on a tree with $D+1$ vertices (one root and D leaves) defined by*

$$\nu_{D+1}(\sigma) = \frac{1}{Z_{D+1}(\beta)} \exp\left\{\beta \sum_{i=1}^{D} \sigma_0\sigma_i + B\sigma_0 + \sum_{i=1}^{D} h_i\sigma_i\right\}, \tag{5.7}$$

where $(h_i)_{i\geq 1}$ are i.i.d. copies of h, independent of D. Then, the thermodynamic limit of the magnetization *per vertex is given by*

$$M(\beta,B) = \mathbb{E}\left[\langle \sigma_0 \rangle_{\nu_{D+1}}\right], \tag{5.8}$$

where the expectation is taken over D and $(h_i)_{i\geq 1}$. More explicitly,

$$M(\beta,B)=\mathbb{E}\left[\tanh\left(B+\sum_{i=1}^{D}\operatorname{atanh}(\hat\beta\tanh(h_i))\right)\right]. \tag{5.9}$$

(b) Internal energy. *Let ν_2' be the random Ising measure on one edge, defined by*

$$\nu_2'(\sigma)=\frac{1}{Z_2(\beta,h_1,h_2)}\exp\left\{\beta\sigma_1\sigma_2+h_1\sigma_1+h_2\sigma_2\right\}, \tag{5.10}$$

where h_1 and h_2 are i.i.d. copies of h. Then the thermodynamic limit of the internal energy *per vertex is given by*

$$U(\beta,B)=-\frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\langle\sigma_1\sigma_2\rangle_{\nu_2'}\right], \tag{5.11}$$

where the expectation is taken over h_1 and h_2. More explicitly,

$$U(\beta,B)=-\frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\frac{\hat\beta+\tanh(h_1)\tanh(h_2)}{1+\hat\beta\tanh(h_1)\tanh(h_2)}\right]. \tag{5.12}$$

Note that the magnetization and internal energy are *local* observables, i.e., they are *spin* or *edge* variables averaged out over the graph. This is not true for the susceptibility, which is an average over pairs of spins, making it more difficult to give an explicit expression. We come back to rewriting the susceptibility in Chapter 9.

5.2 Discussion

The objective method. We study the Ising model on a random graph, which gives rise to a model with double randomness. Still, in the thermodynamic limit, the pressure is essentially deterministic. This is possible, because it suffices to study the Ising model on the local neighborhood of a uniformly chosen vertex. This local neighborhood converges by our assumptions to the tree $\mathcal{T}(D,K,\infty)$, and it thus suffices to study the Ising model on this limiting object. An analysis of this kind is therefore known as the *objective method* introduced by Aldous and Steele in [11].

Universality. That our results hold for a wide variety of random graph models does not come as a surprise. It is believed that the behavior of networks shows a great *universality*. Distances in random graph models, for example, also show a remarkably universal behavior. See, e.g., [64] for an overview of results on distances in power-law random graphs. These distances mainly depend on the power-law exponent and not on other details of the graph.

Erdős-Rényi random graphs. The results above, as well as those in [34], include the Erdős-Rényi random graph as a special case. Earlier results for this model have been obtained in [96], where the high-temperature and zero-temperature pressure are computed using interpolation methods.

Non-homogeneous tree-like graphs. Note that the results above only apply to graphs that converge locally to a *homogeneous* tree and thus, for instance, not for many inhomogeneous random graphs studied in [24] where the vertices can be of different types and hence also the local structure is a multi-type Galton-Watson branching process, although certain parts of our proof easily extend to this case.

Also preferential attachment graphs do not fit in our framework. In such models a growing graph is constructed by attaching new vertices to older vertices proportional to their degrees. Preferential attachment graphs are still locally tree-like, but the offspring is again a multi-type branching process, where the type space is continuous [18].

In both models there will not be a single random positive fixed point h, but the effective field at a certain vertex will depend on its type. Also, the expectations in the results should not only be over the degree D and the fields h_i, but should also be over the types of random vertices.

Non-tree-like graphs. The locally tree-likeness assumption is not very realistic for real-world networks. It would therefore be interesting to investigate what happens if this condition is relaxed. An example could be the configuration model with a *household* structure, i.e., besides the edges formed in the configuration model the vertices are partitioned into small groups, the households, which form a complete graph. The SIR model to study epidemics on such a random graph model, where it is assumed that the infection rate in a household is different from the infection rate of the other neighbors in the graph, is studied in [13, 14]. Also for this model, a branching process approximation can be made to describe the evolving process. It seems worthwhile to investigate if this is also possible for the Ising model on such random graphs.

A second example of non-tree-like graphs are *scale-free percolation clusters* as defined in [33]. This model takes geometry into account by starting with a d-dimensional lattice where all vertices are assigned a random weight. The probability that two vertices are connected then depends both on the weights of these two vertices and the distance between them. By choosing the weights and the dependence on the distance appropriately power-law degree distributions can be obtained. It would be interesting to study the Ising model on such percolation clusters, where it would also be a nice possibility to let the interaction strength depend on the distance between two vertices, modeling that close friends are influencing each other more strongly than friends living on opposite sides of the world.

Dynamics. Our results all describe properties of the system in equilibrium. It would also be interesting to study the dynamics of this model, for example by studying Glauber dynamics to answer questions about *metastability*. E.g., the setting in [86] can be used, where the Ising model in $2d$ is studied with a small positive field and very low temperature. The question then is how long it takes for the system to change from the metastable state where all spins are -1 to the stable system where all spins are $+1$. It is to be expected that again the large degree vertices play a crucial role, but it is unclear if this will be the case for all power-law degree distributions as for the contact process [25] or only for τ small.

It would also be interesting to study Gibbs-non-Gibbs transitions [45]. This describes the phenomenon that when the system undergoes a transition under Glauber dynamics from one Gibbs measure μ to another Gibbs measure ν, the system at intermediate times can not be described with a Gibbs measure. This phenomenon has for example been studied on trees in [44].

Absence of magnetic field. In [81], the Ising model on k-regular graphs is studied when there is no magnetic field, i.e., $B = 0$. There it is shown that the Ising measure converges locally weakly to a symmetric mixture of the Ising measure with $+$ and $-$ boundary conditions on the k-regular tree. This is later generalized in [15] to more general locally tree-like graphs under some mild continuity condition.

Potts model. The thermodynamic limit of the pressure for the Potts model can also be computed when the relevant recursion has a unique positive fixed point as is shown in [37] and the result is similar to that of the Ising model. Later, in [36] the pressure is computed for all values of β and B for the Potts model on the k-regular graph, with k even.

5.3 Overview and organization of the proof

In this section, we give an overview of the proof of Theorem 5.1, and reduce it to the proofs of Propositions 5.4 and 5.5 below. Proposition 5.4 is instrumental to control the implicit dependence of the pressure of the random Bethe tree $\varphi(\beta,B)$ on the inverse temperature β via the field h. This is used in Proposition 5.5 which proves that the derivative of the pressure with respect to β, namely minus the *internal energy*, converges in the thermodynamic limit to the derivative of $\varphi(\beta,B)$. We also clearly indicate how our proof deviates from that by Dembo and Montanari in [34].

We first analyze the case where $B > 0$ and deal with $B \le 0$ later. By the fundamental theorem of calculus,

$$\begin{aligned}\lim_{N\to\infty}\psi_N(\beta,B) &= \lim_{N\to\infty}\left[\psi_N(0,B)+\int_0^\beta \frac{\partial}{\partial\beta'}\psi_N(\beta',B)\mathrm{d}\beta'\right] \\ &= \lim_{N\to\infty}\left[\psi_N(0,B)+\int_0^\varepsilon \frac{\partial}{\partial\beta'}\psi_N(\beta',B)\mathrm{d}\beta'+\int_\varepsilon^\beta \frac{\partial}{\partial\beta'}\psi_N(\beta',B)\mathrm{d}\beta'\right],\end{aligned} \tag{5.13}$$

for any $0 < \varepsilon < \beta$. For all $N \ge 1$,

$$\psi_N(0,B) = \log(2\cosh(B)) = \varphi(0,B), \tag{5.14}$$

so this is also true for $N \to \infty$.

By the uniform sparsity of $(G_N)_{N\ge 1}$,

$$\left|\frac{\partial}{\partial\beta}\psi_N(\beta,B)\right| = \left|\frac{1}{N}\sum_{(i,j)\in E_N}\langle\sigma_i\sigma_j\rangle_{\mu_N}\right| \le \frac{|E_N|}{N} \le c, \tag{5.15}$$

for some constant c. Thus, uniformly in N,

$$\left|\int_0^\varepsilon \frac{\partial}{\partial\beta'}\psi_N(\beta',B)\mathrm{d}\beta'\right| \le c\varepsilon. \tag{5.16}$$

Using the boundedness of the derivative for $\beta' \in [\varepsilon, \beta]$ and by dominated convergence, we also have that

$$\lim_{N\to\infty} \int_{\varepsilon}^{\beta} \frac{\partial}{\partial \beta'} \psi_N(\beta', B) \mathrm{d}\beta' = \int_{\varepsilon}^{\beta} \lim_{N\to\infty} \frac{\partial}{\partial \beta'} \psi_N(\beta', B) \mathrm{d}\beta'. \tag{5.17}$$

For $\beta > 0$, we show that the partial derivative with respect to β of $\psi_N(\beta, B)$ converges to the partial derivative with respect to β of $\varphi(\beta, B)$. For this, we need that we can in fact ignore the dependence of h on β when computing the latter derivative as we show first:

Proposition 5.4 (Dependence of φ on (β, B) via h). *Assume that the distribution P has strongly finite mean. Fix $B_1, B_2 > 0$ and $0 < \beta_1, \beta_2 < \infty$. Let h_1 and h_2 be the fixed points of* (4.1) *for (β_1, B_1) and (β_2, B_2), respectively. Let $\varphi_h(\beta, B)$ be defined as in* (5.2) *with $(h_i)_{i\geq 1}$ replaced by i.i.d. copies of the specified h. Then,*

(a) *For $B_1 = B_2$, there exists a $\lambda_1 < \infty$ such that*

$$|\varphi_{h_1}(\beta_1, B_1) - \varphi_{h_2}(\beta_1, B_1)| \leq \lambda_1 |\beta_1 - \beta_2|^{\tau-1}. \tag{5.18}$$

(b) *For $\beta_1 = \beta_2$, there exists a $\lambda_2 < \infty$ such that*

$$|\varphi_{h_1}(\beta_1, B_1) - \varphi_{h_2}(\beta_1, B_1)| \leq \lambda_2 |B_1 - B_2|^{\tau-1}. \tag{5.19}$$

Note that this proposition only holds if $\tau \in (2, 3)$. For $\tau > 3$, the exponent $\tau - 1$ can be improved to 2, as is shown in [34], but this is of no importance to the proof. We need part (b) of the proposition above later in the proof of Corollary 5.3.

Proposition 5.5 (Convergence of the internal energy). *Assume that the graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P, where P has strongly finite mean, and is uniformly sparse. Let $\beta > 0$. Then, a.s.,*

$$\lim_{N\to\infty} \frac{\partial}{\partial \beta} \psi_N(\beta, B) = \frac{\partial}{\partial \beta} \varphi(\beta, B), \tag{5.20}$$

where $\varphi(\beta, B)$ is given in (5.2).

By Proposition 5.5 and bounded convergence,

$$\int_{\varepsilon}^{\beta} \lim_{N\to\infty} \frac{\partial}{\partial \beta'} \psi_N(\beta', B) \mathrm{d}\beta' = \int_{\varepsilon}^{\beta} \frac{\partial}{\partial \beta'} \varphi(\beta', B) \mathrm{d}\beta' = \varphi(\beta, B) - \varphi(\varepsilon, B), \tag{5.21}$$

again by the fundamental theorem of calculus.

Observing that $0 \leq \tanh(h) \leq 1$, one can show that, by dominated convergence, $\varphi(\beta, B)$ is right-continuous in $\beta = 0$. Thus, letting $\varepsilon \searrow 0$,

$$\begin{aligned} \lim_{N\to\infty} \psi_N(\beta, B) &= \lim_{\varepsilon\searrow 0} \lim_{N\to\infty} \left[\psi_N(0, B) + \int_0^{\varepsilon} \frac{\partial}{\partial \beta'} \psi_N(\beta', B) \mathrm{d}\beta' + \int_{\varepsilon}^{\beta} \frac{\partial}{\partial \beta'} \psi_N(\beta', B) \mathrm{d}\beta' \right] \\ &= \varphi(0, B) + \lim_{\varepsilon\searrow 0} \big(\varphi(\beta, B) - \varphi(\varepsilon, B) \big) = \varphi(\beta, B), \end{aligned} \tag{5.22}$$

which completes the proof for $B > 0$.

The Ising model with $B < 0$ is equivalent to the case $B > 0$, because one can multiply all spin variables $(\sigma_i)_{i\in[N]}$ and B with -1 without changing Boltzmann-Gibbs measure (3.1). Furthermore, note that,

$$\left|\frac{\partial}{\partial B}\psi_N(\beta,B)\right| = \left|\frac{1}{N}\sum_{i\in[N]}\langle\sigma_i\rangle_{\mu_N}\right| \leq 1, \tag{5.23}$$

so that $B \mapsto \psi_N(\beta, B)$ is uniformly Lipschitz continuous with Lipschitz constant one. Therefore,

$$\lim_{N\to\infty}\psi_N(\beta,0) = \lim_{N\to\infty}\lim_{B\searrow 0}\psi_N(\beta,B) = \lim_{B\searrow 0}\lim_{N\to\infty}\psi_N(\beta,B) = \lim_{B\searrow 0}\varphi(\beta,B). \tag{5.24}$$

□

The proof given above follows the line of argument in [34], but in order to prove Theorem 4.1 and Propositions 5.4 and 5.5 we have to make substantial changes to generalize the proof to the infinite-variance case.

The proof of Proposition 5.4(a) is somewhat more elaborate, because we have to distinguish between the cases where D in (5.2) is small or large, but the techniques used remain similar. By, again, taking into account the actual degrees more precisely, the analysis is simplified however: we, for example, do not rely on the exponential decay of the correlations. Part (b) of this proposition is new and can be proved with similar techniques. The proof of Proposition 5.5 is proved in a similar way as in [34].

We prove Proposition 5.4 in Section 5.4 and Proposition 5.5 in Section 5.5. In Section 5.6 we study the thermodynamic quantities to prove Corollary 5.3.

5.4 Dependence of φ on (β, B) via h

We now prove Proposition 5.4 by first bounding the dependence of φ on h in Lemma 5.6 and subsequently bounding the dependence of h on β and B in Lemmas 5.7 and 5.8 respectively.

Lemma 5.6 (Dependence of φ on h). *Assume that the distribution P has strongly finite mean. Fix $B_1, B_2 > 0$ and $0 < \beta_1, \beta_2 < \infty$. Let h_1 and h_2 be the fixed points of* (4.1) *for* (β_1, B_1) *and* (β_2, B_2)*, respectively. Let $\varphi_h(\beta, B)$ be defined as in* (5.2) *with* $(h_i)_{i\geq 1}$ *replaced by i.i.d. copies of the specified h. Then, for some $\lambda < \infty$,*

$$|\varphi_{h_1}(\beta_1,B_1) - \varphi_{h_2}(\beta_1,B_1)| \leq \lambda\|\tanh(h_1)-\tanh(h_2)\|_{\text{MK}}^{\tau-1}, \tag{5.25}$$

where $\|X-Y\|_{\text{MK}}$ denotes the Monge-Kantorovich-Wasserstein distance between the random variables X and Y, i.e., $\|X-Y\|_{\text{MK}}$ is the infimum of $\mathbb{E}[|\hat{X}-\hat{Y}|]$ over all couplings $(\hat{X}, \hat{Y})$ of X and Y.

Proof. Let X_i and Y_i be i.i.d. copies of $X = \tanh(h_1)$ and $Y = \tanh(h_2)$ respectively and also independent of D. When $\|X-Y\|_{\text{MK}} = 0$ or $\|X-Y\|_{\text{MK}} = \infty$, the statement in the lemma clearly holds. Thus, without loss of generality, we fix $\gamma > 1$ and assume that (X_i, Y_i) are i.i.d. pairs, independent of D, that are coupled in such a way that $\mathbb{E}|X_i - Y_i| \leq \gamma\|X-Y\|_{\text{MK}} < \infty$.

Let $\hat{\beta} = \tanh(\beta_1)$ and, for $\ell \geq 2$,

$$F_\ell(x_1,\ldots,x_\ell) = \log\left(e^B \prod_{i=1}^{\ell}(1+\hat{\beta}x_i) + e^{-B}\prod_{i=1}^{\ell}(1-\hat{\beta}x_i)\right) - \frac{1}{\ell-1}\sum_{1\leq i<j\leq \ell}\log(1+\hat{\beta}x_ix_j), \tag{5.26}$$

and let

$$F_1(x_1,x_2) = \frac{1}{2}\Big(\log\big(e^B(1+\hat{\beta}x_1)+e^{-B}(1-\hat{\beta}x_1)\big) + \log\big(e^B(1+\hat{\beta}x_2)+e^{-B}(1-\hat{\beta}x_2)\big) - \log(1+\hat{\beta}x_1x_2)\Big). \tag{5.27}$$

Then, with D having distribution P,

$$\varphi_{h_1}(\beta_1,B_1) = F_0 + \mathbb{E}[F_D(X_1,\ldots,X_{\max\{2,D\}})], \tag{5.28}$$

and

$$\varphi_{h_2}(\beta_1,B_1) = F_0 + \mathbb{E}[F_D(Y_1,\ldots,Y_{\max\{2,D\}})], \tag{5.29}$$

for some constant F_0. In the remainder of the proof we assume that F_1 is defined as in (5.26). The proof, however, also works for F_1 as defined in (5.27).

We split the absolute difference between $\varphi_{h_1}(\beta_1,B_1)$ and $\varphi_{h_2}(\beta_1,B_1)$ into two parts depending on whether D is small or large, i.e., for some constant $\theta > 0$ to be chosen later on, we split

$$\Big|\mathbb{E}\big[F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D)\big]\Big| \leq \Big|\mathbb{E}\big[(F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D))\mathbb{1}_{\{D\leq\theta\}}\big]\Big| + \Big|\mathbb{E}\big[(F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D))\mathbb{1}_{\{D>\theta\}}\big]\Big|. \tag{5.30}$$

Note that

$$\begin{aligned} F_\ell(Y_1,\ldots,Y_\ell) - F_\ell(X_1,\ldots,X_\ell) &= \int_0^1 \frac{d}{ds}F_\ell(sY_1+(1-s)X_1,\ldots,sY_\ell+(1-s)X_\ell)\Big|_{s=t}\mathrm{d}t \\ &= \int_0^1 \sum_{i=1}^{\ell}(Y_i-X_i)\frac{\partial F_\ell}{\partial x_i}(tY_1+(1-t)X_1,\ldots,tY_\ell+(1-t)X_\ell)\mathrm{d}t \\ &= \sum_{i=1}^{\ell}(Y_i-X_i)\int_0^1 \frac{\partial F_\ell}{\partial x_i}(tY_1+(1-t)X_1,\ldots,tY_\ell+(1-t)X_\ell)\mathrm{d}t. \end{aligned} \tag{5.31}$$

As observed in [34, Corollary 6.3], $\frac{\partial F_\ell}{\partial x_i}$ is uniformly bounded, so that

$$\Big|F_\ell(Y_1,\ldots,Y_\ell) - F_\ell(X_1,\ldots,X_\ell)\Big| \leq \lambda_1 \sum_{i=1}^{\ell}|Y_i - X_i|, \tag{5.32}$$

where λ_1 is allowed to change from line to line. Hence,

$$\begin{aligned} \Big|\mathbb{E}\big[(F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D))\mathbb{1}_{\{D>\theta\}}\big]\Big| &\leq \mathbb{E}\Big[\sum_{i=1}^{D}|Y_i-X_i|\lambda_1\mathbb{1}_{\{D>\theta\}}\Big] \\ &\leq \lambda_1\|X-Y\|_{\text{MK}}\mathbb{E}[D\mathbb{1}_{\{D>\theta\}}]. \end{aligned} \tag{5.33}$$

By Lemma 2.7,

$$\mathbb{E}[D\mathbb{1}_{\{D>\theta\}}] \leq C_{1,\tau}\theta^{-(\tau-2)}, \tag{5.34}$$

so that

$$\left|\mathbb{E}\left[(F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D))\mathbb{1}_{\{D>\theta\}}\right]\right| \leq \lambda_1 \|X-Y\|_{\text{MK}}\theta^{-(\tau-2)}. \tag{5.35}$$

By the fundamental theorem of calculus, we can also write

$$F_\ell(Y_1,\ldots,Y_\ell) - F_\ell(X_1,\ldots,X_\ell) = \sum_{i=1}^{\ell}\Delta_i F_\ell + \sum_{i\neq j}^{\ell}(Y_i - X_i)(Y_j - X_j) f_{ij}^{(\ell)}, \tag{5.36}$$

with

$$\Delta_i F_\ell = (Y_i - X_i)\int_0^1 \frac{\partial F_\ell}{\partial x_i}(X_1,\ldots,tY_i+(1-t)X_i,\ldots,X_\ell)\mathrm{d}t, \tag{5.37}$$

and

$$f_{ij}^{(\ell)} = \int_0^1\int_0^t \frac{\partial^2 F_\ell}{\partial x_i \partial x_j}(sY_1+(1-s)X_1,\ldots,sY_i+(1-s)X_i,\ldots,sY_\ell+(1-s)X_\ell)\mathrm{d}s\mathrm{d}t. \tag{5.38}$$

Therefore,

$$\begin{aligned}
&\left|\mathbb{E}\left[(F_D(Y_1,\ldots,Y_D) - F_D(X_1,\ldots,X_D))\mathbb{1}_{\{D\leq\theta\}}\right]\right| \qquad (5.39)\\
&\qquad \leq \left|\mathbb{E}\left[\sum_{i=1}^{D}\Delta_i F_D \mathbb{1}_{\{D\leq\theta\}}\right]\right| + \left|\mathbb{E}\left[\sum_{i\neq j}^{D}(Y_i - X_i)(Y_j - X_j) f_{ij}^{(D)}\mathbb{1}_{\{D\leq\theta\}}\right]\right|.
\end{aligned}$$

Since $\frac{\partial^2 F_\ell}{\partial x_i \partial x_j}$ is also uniformly bounded [34, Corollary 6.3],

$$\begin{aligned}
\left|\mathbb{E}\left[\sum_{i\neq j}^{D}(Y_i - X_i)(Y_j - X_j) f_{ij}^{(D)}\mathbb{1}_{\{D\leq\theta\}}\right]\right| &\leq \lambda_2 \mathbb{E}\left[\sum_{i\neq j}^{D}|Y_i - X_i||Y_j - X_j|\mathbb{1}_{\{D\leq\theta\}}\right]\\
&\leq \lambda_2 \|X-Y\|_{\text{MK}}^2 \mathbb{E}[D^2\mathbb{1}_{\{D\leq\theta\}}]\\
&\leq \lambda_2 \|X-Y\|_{\text{MK}}^2 \theta^{-(\tau-3)}, \qquad (5.40)
\end{aligned}$$

by Lemma 2.7, where λ_2 is allowed to change from line to line. We split

$$\left|\mathbb{E}\left[\sum_{i=1}^{D}\Delta_i F_D \mathbb{1}_{\{D\leq\theta\}}\right]\right| \leq \left|\mathbb{E}\left[\sum_{i=1}^{D}\Delta_i F_D\right]\right| + \left|\mathbb{E}\left[\sum_{i=1}^{D}\Delta_i F_D \mathbb{1}_{\{D>\theta\}}\right]\right|. \tag{5.41}$$

By symmetry of the functions F_ℓ with respect to their arguments, for i.i.d. (X_i, Y_i) independent of D,

$$\mathbb{E}\left[\sum_{i=1}^{D}\Delta_i F_D\right] = \mathbb{E}\left[D\Delta_1 F_D\right] = \mathbb{E}\left[D(Y_1 - X_1)\int_0^1 \frac{\partial F_D}{\partial x_1}(tY_1+(1-t)X_1, X_2,\ldots,X_D)\mathrm{d}t\right]. \tag{5.42}$$

Differentiating (5.26) gives, for $\ell \geq 2$,

$$\frac{\partial}{\partial x_1} F_\ell(x_1, \ldots, x_\ell) = \psi(x_1, g_\ell(x_2, \ldots, x_\ell)) - \frac{1}{\ell - 1} \sum_{j=2}^{\ell} \psi(x_1, x_j), \tag{5.43}$$

where $\psi(x, y) = \hat{\beta} y/(1 + \hat{\beta} x y)$ and

$$g_\ell(x_2, \ldots, x_\ell) = \tanh\left(B + \sum_{j=2}^{\ell} \operatorname{atanh}(\hat{\beta} x_j)\right). \tag{5.44}$$

Using that $\ell P_\ell = \mathbb{E}[D]\rho_{\ell-1}$, we have that, with K distributed as ρ,

$$\mathbb{E}[D\psi(X_1, g_D(X_2, \ldots, X_D))] = \mathbb{E}[D]\mathbb{E}[\psi(X_1, g_{K+1}(X_2, \ldots, X_{K+1}))] = \mathbb{E}[D]\mathbb{E}[\psi(X_1, X_2)], \tag{5.45}$$

because $g_{K+1}(X_2, \ldots, X_{K+1})$ is a fixed point of (4.1), so that $g_{K+1}(X_2, \ldots, X_{K+1}) \stackrel{d}{=} X_2$ and is independent of X_1. Therefore, one can show that

$$\mathbb{E}\left[D \frac{\partial F_D}{\partial x_1}(x, X_2, \ldots, X_{\max\{2,D\}})\right] = 0, \qquad \text{for all } x \in [-1, 1]. \tag{5.46}$$

Since $\frac{\partial F_D}{\partial x_1}$ is uniformly bounded, $D\frac{\partial F_D}{\partial x_1}$ is integrable, so that, by Fubini's theorem and (5.46),

$$\begin{aligned} &\mathbb{E}\left[\sum_{i=1}^{D} \Delta_i F_D\right] \\ &= \mathbb{E}\left[(Y_1 - X_1)\int_0^1 \mathbb{E}\left[D\frac{\partial F_D}{\partial x_1}(tY_1 + (1-t)X_1, X_2, \ldots, X_D) \,\Big|\, X_1, Y_1\right] \mathrm{d}t\right] = 0. \end{aligned} \tag{5.47}$$

Furthermore, by (5.37) and the uniform boundedness of $\frac{\partial F_\ell}{\partial x_i}$,

$$\left|\mathbb{E}\left[\sum_{i=1}^{D} \Delta_i F_D \mathbb{1}_{\{D>\theta\}}\right]\right| \leq \mathbb{E}\left[\sum_{i=1}^{D} |Y_i - X_i| \lambda_1 \mathbb{1}_{\{D>\theta\}}\right] \leq \lambda_1 \|X - Y\|_{\text{MK}} \theta^{-(\tau-2)}. \tag{5.48}$$

Therefore, we conclude that

$$\left|\mathbb{E}\left[(F_D(Y_1, \ldots, Y_D) - F_D(X_1, \ldots, X_D))\mathbb{1}_{\{D\leq\theta\}}\right]\right| \leq \lambda_1 \|X-Y\|_{\text{MK}} \theta^{-(\tau-2)} + \lambda_2 \|X-Y\|_{\text{MK}}^2 \theta^{-(\tau-3)}. \tag{5.49}$$

Combining (5.35) and (5.49) and letting $\theta = \|X - Y\|_{\text{MK}}^{-1}$ yields the desired result. □

Lemma 5.7 (Dependence of h on β). *Fix $B > 0$ and $0 < \beta_1, \beta_2 \leq \beta_{\max}$. Let h_{β_1} and h_{β_2}, where we made the dependence of h on β explicit, be the fixed points of* (4.1) *for (β_1, B) and (β_2, B), respectively. Then, there exists a $\lambda < \infty$ such that*

$$\|\tanh(h_{\beta_1}) - \tanh(h_{\beta_2})\|_{\text{MK}} \leq \lambda |\beta_1 - \beta_2|. \tag{5.50}$$

Proof. For a given tree $\mathscr{T}(K,\infty)$ we can, as in the proof of Theorem 4.1, couple $\tanh(h_\beta)$ to the root magnetizations $m_\beta^{\ell,f/+}(B)$ such that, for all $\beta \geq 0$ and $\ell \geq 0$,

$$m_\beta^{\ell,f}(B) \leq \tanh(h_\beta) \leq m_\beta^{\ell,+}(B), \tag{5.51}$$

where we made the dependence of $m^{\ell,f/+}$ on β explicit. Without loss of generality, we assume that $0 < \beta_1 \leq \beta_2 \leq \beta_{\max}$. Then, by the GKS inequality,

$$|\tanh(h_{\beta_2}) - \tanh(h_{\beta_1})| \leq m_{\beta_2}^{\ell,+}(B) - m_{\beta_1}^{\ell,f}(B) = m_{\beta_2}^{\ell,+}(B) - m_{\beta_2}^{\ell,f}(B) + m_{\beta_2}^{\ell,f}(B) - m_{\beta_1}^{\ell,f}(B). \tag{5.52}$$

By Lemma 4.3, a.s.,

$$m_{\beta_2}^{\ell,+}(B) - m_{\beta_2}^{\ell,f}(B) \leq A/\ell, \tag{5.53}$$

for some $A < \infty$. Since $m_\beta^{\ell,f}(B)$ is non-decreasing in β by the GKS inequality,

$$m_{\beta_2}^{\ell,f}(B) - m_{\beta_1}^{\ell,f}(B) \leq (\beta_2 - \beta_1) \sup_{\beta_1 \leq \beta \leq \beta_{\max}} \frac{\partial m^{\ell,f}}{\partial \beta}. \tag{5.54}$$

Letting $\ell \to \infty$, it thus suffices to show that $\partial m^{\ell,f}/\partial\beta$ is, a.s., bounded uniformly in ℓ and $0 < \beta_1 \leq \beta \leq \beta_{\max}$.

From [34, Lemma 4.6] we know that

$$\frac{\partial}{\partial \beta} m^{\ell,f}(\beta, B) \leq \sum_{k=0}^{\ell-1} V_{k,\ell}, \tag{5.55}$$

with

$$V_{k,\ell} = \sum_{i \in \partial \mathscr{T}(k)} \Delta_i \frac{\partial}{\partial B_i} m^\ell(\underline{B}, 0)\Big|_{\underline{B}=B}. \tag{5.56}$$

By Lemma 4.2 and the GHS inequality,

$$\frac{\partial}{\partial B_i} m^\ell(\underline{B}, 0) = \frac{\partial}{\partial B_i} m^{\ell-1}(\underline{B}, \underline{H}) \leq \frac{\partial}{\partial B_i} m^{\ell-1}(\underline{B}, 0), \tag{5.57}$$

for some field $\underline{H}$, so that $V_{k,\ell}$ is non-increasing in ℓ. We may assume that $B_i \geq B_{\min}$ for all $i \in \mathscr{T}(\ell)$ for some $B_{\min}$. Thus, also using Lemma 4.2,

$$\begin{aligned} V_{k,\ell} \leq V_{k,k+1} &= \sum_{i \in \partial \mathscr{T}(k)} \Delta_i \frac{\partial}{\partial B_i} m^{k+1}(\underline{B}, 0)\Big|_{\underline{B}=B} \leq \sum_{i \in \partial \mathscr{T}(k)} \Delta_i \frac{\partial}{\partial B_i} m^k(\underline{B}, \xi\{\Delta_i\})\Big|_{\underline{B}=B} \\ &= \frac{\partial}{\partial H} m^k(B, H\{\Delta_i\})\Big|_{H=\xi(B_{\min})}, \end{aligned} \tag{5.58}$$

where $\xi = \xi(B_{\min})$ is defined in (4.2). By the GHS inequality this derivative is non-increasing in H, so that, by Lemma 4.2, the above is at most

$$\frac{1}{\xi}\left[m^k(B, \xi\{\Delta_i\}) - m^k(B, 0)\right] \leq \frac{1}{\xi}\left[m^{k+1}(B, 0) - m^k(B, 0)\right]. \tag{5.59}$$

Therefore,

$$\frac{\partial}{\partial \beta} m^{\ell,f}(\beta,B) \le \sum_{k=0}^{\ell-1} V_{k,\ell} \le \frac{1}{\xi}\sum_{k=0}^{\ell-1}\left[m^{k+1}(B,0)-m^{k}(B,0)\right] \le \frac{1}{\xi} < \infty, \tag{5.60}$$

for $0<\beta_1\le\beta\le\beta_{\max}$. □

Lemma 5.8 (Dependence of h on B). *Fix $\beta \ge 0$ and $B_1, B_2 \ge B_{\min} > 0$. Let h_{B_1} and h_{B_2}, where we made the dependence of h on B explicit, be the fixed points of* (4.1) *for (β,B_1) and (β,B_2), respectively. Then, there exists a $\lambda<\infty$ such that*

$$\|\tanh(h_{B_1})-\tanh(h_{B_2})\|_{\text{MK}} \le \lambda|B_1-B_2|. \tag{5.61}$$

Proof. This lemma can be proved along the same lines as Lemma 5.7. Therefore, for a given tree $\mathscr{T}(K,\infty)$, we can couple $\tanh(h_B)$ to the root magnetizations $m^{\ell,f/+}(B)$ such that, for all $B>0$ and $\ell\ge0$,

$$m^{\ell,f}(B) \le \tanh(h_B) \le m^{\ell,+}(B). \tag{5.62}$$

Without loss of generality, we assume that $0<B_{\min}\le B_1\le B_2$. Then, by the GKS inequality,

$$|\tanh(h_{B_2})-\tanh(h_{B_1})| \le m^{\ell,+}(B_2)-m^{\ell,f}(B_1) = m^{\ell,+}(B_2)-m^{\ell,f}(B_2)+m^{\ell,f}(B_2)-m^{\ell,f}(B_1). \tag{5.63}$$

By Lemma 4.3, a.s.,

$$m^{\ell,+}(B_2)-m^{\ell,f}(B_2) \le A/\ell, \tag{5.64}$$

for some $A<\infty$. Since $m^{\ell,f}(B)$ is non-decreasing in B by the GKS inequality,

$$m^{\ell,f}(B_2)-m^{\ell,f}(B_1) \le (B_2-B_1)\sup_{B\ge B_{\min}>0}\frac{\partial m^{\ell,f}}{\partial B}. \tag{5.65}$$

Letting $\ell\to\infty$, it thus suffices to show that $\partial m^{\ell,f}/\partial B$ is bounded uniformly in ℓ and $B\ge B_{\min}>0$. This follows from the GHS inequality, i.e., the concavity of the magnetization in B:

$$\sup_{B\ge B_{\min}>0}\frac{\partial m^{\ell,f}}{\partial B} \le \left.\frac{\partial m^{\ell,f}}{\partial B}\right|_{B=B_{\min}} \le \frac{2}{B_{\min}}\left[m^{\ell,f}(B_{\min})-m^{\ell,f}(B_{\min}/2)\right] \le \frac{2}{B_{\min}} < \infty, \tag{5.66}$$

because for a concave function $f: x\mapsto f(x)$ it holds that $f'(x)\le \frac{f(x)-f(x-\varepsilon)}{\varepsilon}$. □

5.5 Convergence of the internal energy

We start by identifying the thermodynamic limit of the internal energy:

Lemma 5.9 (From graphs to trees). *Assume that the graph sequence $(G_N)_{N\ge1}$ is locally tree-like with asymptotic degree distribution P, where P has finite mean, and is uniformly sparse. Then, a.s.,*

$$\lim_{N\to\infty}\frac{\partial}{\partial\beta}\psi_N(\beta,B) = \frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\langle\sigma_1\sigma_2\rangle_{\nu_2'}\right], \tag{5.67}$$

where ν_2' is defined in (5.10).

Lemma 5.9 is proved in Section 5.5.1. Next, we compute the derivative of $\varphi(\beta,B)$ with respect to β in the following lemma and show that it equals the one on the graph:

Lemma 5.10 (Tree analysis). *Assume that distribution P has strongly finite mean. Then,*

$$\frac{\partial}{\partial\beta}\varphi(\beta,B) = \frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\left\langle\sigma_1\sigma_2\right\rangle_{\nu_2'}\right], \tag{5.68}$$

where ν_2' is defined in (5.10).

Lemma 5.10 is proved in Section 5.5.2. Lemmas 5.9 and 5.10 clearly imply Proposition 5.5. □

5.5.1 From graphs to trees: proof of Lemma 5.9

This lemma can be proved as in [34]. The idea is to note that

$$\frac{\partial}{\partial\beta}\psi_N(\beta,B) = \frac{1}{N}\sum_{(i,j)\in E_N}\left\langle\sigma_i\sigma_j\right\rangle_{\mu_N} = \frac{|E_N|}{N}\frac{\sum_{(i,j)\in E_N}\left\langle\sigma_i\sigma_j\right\rangle_{\mu_N}}{|E_N|}. \tag{5.69}$$

By the local convergence and the uniform sparsity, we have that, a.s. (see (2.11)),

$$\lim_{N\to\infty}\frac{|E_N|}{N} = \mathbb{E}[D]/2. \tag{5.70}$$

The second term of the right hand side of (5.69) can be seen as the expectation with respect to a uniformly chosen edge (i,j) of the correlation $\left\langle\sigma_i\sigma_j\right\rangle_{\mu_N}$. For a uniformly chosen edge (i,j), denote by $B_{(i,j)}(t)$ all vertices at distance from either vertex i or j at most t, and let $\partial B_{(i,j)}(t) = B_{(i,j)}(t)\setminus B_{(i,j)}(t-1)$. By the GKS inequality, for any $t\geq 1$,

$$\left\langle\sigma_i\sigma_j\right\rangle^f_{B_{(i,j)}(t)} \leq \left\langle\sigma_i\sigma_j\right\rangle_{\mu_N} \leq \left\langle\sigma_i\sigma_j\right\rangle^+_{B_{(i,j)}(t)}, \tag{5.71}$$

where $\left\langle\sigma_i\sigma_j\right\rangle^{+/f}_{B_{(i,j)}(t)}$ is the correlation in the Ising model on $B_{(i,j)}(t)$ with $+$/free boundary conditions on $\partial B_{(i,j)}(t)$.

Let $\overline{\mathscr{T}}(K,t)$ be the tree formed by joining the roots, ϕ_1 and ϕ_2, of two branching processes with t generations and with offspring being i.i.d. copies of K at each vertex, also at the roots. Then, taking $N\to\infty$, $B_{(i,j)}(t)$ converges to $\overline{\mathscr{T}}(K,t)$, because of the local convergence of the graph sequence. Indeed, observe that a random edge can be chosen, by first picking a vertex with probability proportional to its degree, and then selecting a neighbor uniformly at random. Using this, one can show (see [34, Lemma 6.4]), also using the uniform sparsity, that, for all $t\geq 1$, a.s.,

$$\lim_{N\to\infty}\mathbb{E}_{(i,j)}\left[\left\langle\sigma_i\sigma_j\right\rangle^{+/f}_{B_{(i,j)}(t)}\right] = \mathbb{E}\left[\left\langle\sigma_{\phi_1}\sigma_{\phi_2}\right\rangle^{+/f}_{\overline{\mathscr{T}}(K,t)}\right], \tag{5.72}$$

where the first expectation is with respect to a uniformly at random chosen edge $(i,j)\in E_N$ and the second expectation with respect to the tree $\overline{\mathscr{T}}(K,t)$. By Lemma 4.2 and Theorem 4.1,

$$\lim_{t\to\infty}\mathbb{E}\left[\left\langle\sigma_{\phi_1}\sigma_{\phi_2}\right\rangle^{+/f}_{\overline{\mathscr{T}}(K,t)}\right] = \mathbb{E}\left[\left\langle\sigma_1\sigma_2\right\rangle_{\nu_2'}\right], \tag{5.73}$$

thus proving the lemma. □

5.5.2 Tree analysis: proof of Lemma 5.10

From Proposition 5.4 it follows that we can assume that β is fixed in h when differentiating $\varphi(\beta, B)$ with respect to β. We let $X_i, i \geq 1$, be i.i.d. copies of $\tanh(h)$, also independent of D. Then, differentiating (5.2) gives, using the exchangeability of the X_i,

$$\frac{\partial}{\partial \beta}\varphi(\beta,B)=\frac{\mathbb{E}[D]}{2}\hat{\beta}-\frac{\mathbb{E}[D]}{2}(1-\hat{\beta}^2)\mathbb{E}[\psi(X_1,X_2)]+(1-\hat{\beta}^2)\mathbb{E}[D\psi(X_1,g_D(X_2,\ldots,X_D))], \tag{5.74}$$

where now $\psi(x,y)=\frac{xy}{1+\hat{\beta}xy}$ and $g_\ell(x_1,\ldots,x_\ell)$ is defined as in (5.44). Since (5.45) is valid for any bounded function $\psi(x,y)$

$$\frac{\partial}{\partial \beta}\varphi(\beta,B)=\frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\hat{\beta}+(1-\hat{\beta}^2)\frac{X_1X_2}{1+\hat{\beta}X_1X_2}\right]=\frac{\mathbb{E}[D]}{2}\mathbb{E}\left[\frac{\hat{\beta}+X_1X_2}{1+\hat{\beta}X_1X_2}\right]. \tag{5.75}$$

Since, with h_1, h_2 i.i.d. copies of h,

$$\begin{aligned}\mathbb{E}\left[\frac{\hat{\beta}+X_1X_2}{1+\hat{\beta}X_1X_2}\right]&=\mathbb{E}\left[\frac{\hat{\beta}+\tanh(h_1)\tanh(h_2)}{1+\hat{\beta}\tanh(h_1)\tanh(h_2)}\right]\\&=\mathbb{E}\left[\frac{e^{\beta+h_1+h_2}-e^{-\beta-h_1+h_2}-e^{-\beta+h_1-h_2}+e^{\beta-h_1-h_2}}{e^{\beta+h_1+h_2}+e^{-\beta-h_1+h_2}+e^{-\beta+h_1-h_2}+e^{\beta-h_1-h_2}}\right]=\mathbb{E}\left[\langle\sigma_1\sigma_2\rangle_{\nu_2'}\right],\end{aligned} \tag{5.76}$$

where ν_2' is given in (5.10), we have proved the lemma.

□

5.6 Thermodynamic quantities

To prove the statements in Theorem 5.2 we need to show that we can interchange the limit of $N \to \infty$ and the derivatives of the finite volume pressure. We can do this using the monotonicity properties of the Ising model and the following lemma:

Lemma 5.11 (Interchanging limits and derivatives). *Let $(f_N(x))_{N\geq 1}$ be a sequence of functions that are twice differentiable in x. Assume that*

(i) $\lim_{N\to\infty} f_N(x) = f(x)$ for some function $x \mapsto f(x)$ which is differentiable in x;

(ii) $\frac{d}{dx} f_N(x)$ is monotone in $[x-\delta, x+\delta]$ for all $N \geq 1$ and some $\delta > 0$.

Then,

$$\lim_{N\to\infty}\frac{\mathrm{d}}{\mathrm{d}x}f_N(x)=\frac{\mathrm{d}}{\mathrm{d}x}f(x). \tag{5.77}$$

Proof. First, suppose that $\frac{\mathrm{d}^2}{\mathrm{d}y^2}f_N(y) \geq 0$ for all $y \in [x-\delta, x+\delta]$, all $N \geq 1$ and some $\delta > 0$. Then, for $\delta > 0$ sufficiently small and all $N \geq 1$,

$$\frac{f_N(x-\delta)-f_N(x)}{-\delta}\leq\frac{\mathrm{d}}{\mathrm{d}x}f_N(x)\leq\frac{f_N(x+\delta)-f_N(x)}{\delta}, \tag{5.78}$$

and, by taking $N \to \infty$ and assumption (i),

$$\frac{f(x-\delta)-f(x)}{-\delta}\leq\liminf_{N\to\infty}\frac{\mathrm{d}}{\mathrm{d}x}f_N(x)\leq\limsup_{N\to\infty}\frac{\mathrm{d}}{\mathrm{d}x}f_N(x)\leq\frac{f(x+\delta)-f(x)}{\delta}. \tag{5.79}$$

Taking $\delta \searrow 0$ now proves the result. The proof for $\frac{\mathrm{d}^2}{\mathrm{d}x^2}f_N(x) \leq 0$ is similar. □

We are now ready to prove Theorem 5.2.

Proof of Theorem 5.2. We apply Lemma 5.11 with the role of f_N taken by $B \mapsto \psi_N(\beta,B)$, since

$$M_N(\beta,B) = \frac{1}{N}\sum_{i\in[N]} \langle \sigma_i \rangle_{\mu_N} = \frac{\partial}{\partial B}\psi_N(\beta,B), \tag{5.80}$$

and $\lim_{N\to\infty}\psi_N(\beta,B) = \varphi(\beta,B)$ by Theorem 5.1 and $B \mapsto M_N(\beta,B)$ is non-decreasing by the GKS inequality. Therefore,

$$\lim_{N\to\infty} M_N(\beta,B) = \lim_{N\to\infty}\frac{\partial}{\partial B}\psi_N(\beta,B) = \frac{\partial}{\partial B}\varphi(\beta,B), \tag{5.81}$$

which proves part (a).

Part (b) follows immediately from Proposition 5.5 and the observation that

$$U_N = -\frac{1}{N}\sum_{(i,j)\in E_N} \langle \sigma_i\sigma_j \rangle_{\mu_N} = -\frac{\partial}{\partial \beta}\psi_N(\beta,B). \tag{5.82}$$

Part (c) is proved using Lemma 5.11 by combining part (a) of this theorem and that $B \mapsto \frac{\partial}{\partial B}M_N(\beta,B)$ is non-increasing by the GHS inequality. □

We can now prove each of the statements in Corollary 5.3 by taking the proper derivative of $\varphi(\beta,B)$.

Proof of Corollary 5.3. It follows from Theorem 5.2 (a) that the magnetization per vertex is given by

$$M(\beta,B) = \frac{\partial}{\partial B}\varphi(\beta,B). \tag{5.83}$$

Similar to the proof of Lemma 5.10, we can ignore the dependence of h on B when differentiating $\varphi(\beta,B)$ with respect to B by Proposition 5.4. Therefore,

$$\begin{aligned}
\frac{\partial}{\partial B}\varphi(\beta,B) &= \frac{\partial}{\partial B}\mathbb{E}\left[\log\left(e^B\prod_{i=1}^{D}(1+\hat\beta\tanh(h_i)) + e^{-B}\prod_{i=1}^{D}(1-\hat\beta\tanh(h_i))\right)\right] \\
&= \mathbb{E}\left[\frac{e^B\prod_{i=1}^{D}(1+\hat\beta\tanh(h_i)) - e^{-B}\prod_{i=1}^{D}(1-\hat\beta\tanh(h_i))}{e^B\prod_{i=1}^{D}(1+\hat\beta\tanh(h_i)) + e^{-B}\prod_{i=1}^{D}(1-\hat\beta\tanh(h_i))}\right] \\
&= \mathbb{E}\left[\frac{e^B\prod_{i=1}^{D}\left(\frac{1+\hat\beta\tanh(h_i)}{1-\hat\beta\tanh(h_i)}\right)^{1/2} - e^{-B}\prod_{i=1}^{D}\left(\frac{1-\hat\beta\tanh(h_i)}{1+\hat\beta\tanh(h_i)}\right)^{1/2}}{e^B\prod_{i=1}^{D}\left(\frac{1+\hat\beta\tanh(h_i)}{1-\hat\beta\tanh(h_i)}\right)^{1/2} + e^{-B}\prod_{i=1}^{D}\left(\frac{1-\hat\beta\tanh(h_i)}{1+\hat\beta\tanh(h_i)}\right)^{1/2}}\right],
\end{aligned} \tag{5.84}$$

where D has distribution P and $(h_i)_{i\geq 1}$'s are i.i.d. copies of h, independent of D. Using that $\operatorname{atanh}(x) = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right)$ the above simplifies to

$$\begin{aligned}
&\mathbb{E}\left[\frac{e^B\prod_{i=1}^{D}e^{\operatorname{atanh}(\hat\beta\tanh(h_i))} - e^{-B}\prod_{i=1}^{D}e^{-\operatorname{atanh}(\hat\beta\tanh(h_i))}}{e^B\prod_{i=1}^{D}e^{\operatorname{atanh}(\hat\beta\tanh(h_i))} + e^{-B}\prod_{i=1}^{D}e^{-\operatorname{atanh}(\hat\beta\tanh(h_i))}}\right] \\
&\qquad = \mathbb{E}\left[\tanh\left(B + \sum_{i=1}^{D}\operatorname{atanh}(\hat\beta\tanh(h_i))\right)\right].
\end{aligned} \tag{5.85}$$

By Lemma 4.2, this indeed equals $\mathbb{E}\left[\left\langle\sigma_0\right\rangle_{\nu_{D+1}}\right]$, where ν_{D+1} is given in (5.7), which proves part (a).

Part (b) immediately follows from Theorem 5.2(b) and Lemma 5.10. □

6

Intermezzo: Infinite-mean random graphs

So far, we have studied the Ising model on random graphs with strongly finite mean, that is, there exist constants $\tau > 2$ and $C > 0$ such that

$$p_{\geq k} \leq C k^{-(\tau-1)}. \tag{6.1}$$

The question naturally arises what happens if the degree distribution obeys a power law with exponent $\tau \in (1,2)$ in which case the degrees have *infinite mean*. Such networks are observed in biological networks, see for example references in [97].

In [19, 47] the configuration model (CM) is studied when the degrees satisfy a power-law degree distribution with $\tau \in (1,2)$. There it is shown that the graph does not have a tree-like structure in this case: a part of the vertices, the so-called *super*vertices defined below form a complete subgraph and all other vertices connect only to these supervertices. It therefore does not make sense to use the tree-based approach of the previous chapters. We make heavy use of the results in [19, 47] in this chapter.

The configuration model with infinite-mean degrees will also not give a simple graph with high probability. The behavior might, and will, depend strongly on how multiple edges and self-loops are dealt with. Keeping all edges or removing multiple edges for instance greatly influences first-passage percolation on this model [19]. We therefore investigate both the *original* model, i.e., keeping all the multiple edges, and the *erased* model where multiple edges and self-loops are deleted.

6.1 Model definitions

6.1.1 Model for the original configuration model

The *original* CM is defined as in Section 2.3. For convenience, we make a slightly stronger assumption on the degree sequence, i.e., we assume that there exists a constant $0 < C_p < \infty$ such that

$$p_{\geq k} = C_p k^{-(\tau-1)}(1 + o(1)). \tag{6.2}$$

It is shown in [47, Lemma 2.1] that the total degree $L_N^{\mathrm{or}} = \sum_{i\in[N]} D_i^{\mathrm{or}}$ properly rescaled converges in distribution, i.e.,

$$\frac{L_N^{\mathrm{or}}}{N^{1/(\tau-1)}} \xrightarrow{d} \eta, \tag{6.3}$$

where η is a known $\tau - 1$ stable random variable. Since $L_N = 2|E_N|$ is growing faster than N, we have to scale the temperature with N. With the Ising model defined as

$$\mu(\sigma) = \frac{1}{Z_N} \exp\left\{ \beta J \sum_{(i,j)\in E_N} \sigma_i \sigma_j + B \sum_{i\in[N]} \sigma_i \right\}, \tag{6.4}$$

the internal energy per vertex equals,

$$\frac{J}{N} \sum_{(i,j)\in E_N} \langle \sigma_i \sigma_j \rangle. \tag{6.5}$$

Since we want this to be an *intensive* quantity, i.e., not growing with N, we have to choose $J = J_N$ as a decreasing function of N. Specifically, we choose

$$J_N^{\mathrm{or}} = N/N^{1/(\tau-1)} = 1/N^{(2-\tau)/(\tau-1)}. \tag{6.6}$$

From now on we write $\beta_N = \beta J_N$.

6.1.2 Model for the erased configuration model

The *erased* CM is constructed by starting with a graph generated according to the original CM and after this, erasing all self-loops and merging multiple edges between any pair of vertices into a single edge between these vertices.

The result in [19, Lemma 6.8] suggests that $L_N^{\mathrm{er}} = O(N \log N)$. Hence, we again scale the interaction strength and choose

$$J_N^{\mathrm{er}} = 1/\log N. \tag{6.7}$$

6.2 Results

In the next theorem, we compute the magnetization for the original CM and conclude that the spontaneous magnetization is zero for all positive temperatures.

Theorem 6.1 (Magnetization in original CM). *For all $0 \le \beta < \infty$ and $B > 0$, the magnetization equals*

$$M^{\mathrm{or}}(\beta, B) = \tanh(B), \tag{6.8}$$

and hence, the spontaneous magnetization equals

$$M^{\mathrm{or}}(\beta, 0^+) = 0. \tag{6.9}$$

The idea behind the proof of this theorem is the following. The interaction strength J_N^{or} has to go down to 0 very fast to keep the internal energy per vertex finite a.s. The largest contribution to the total degree, however, comes from only a small number of *super*vertices, which have a very large degree of the order $N^{1/(\tau-1)}$. The degrees of *normal* vertices is much smaller. Hence, the interaction between a normal vertex and the supervertices it is attached to is negligibly small for the magnetization of the normal vertex. Hence, the normal vertices behave more or less independently from the rest of the graph.

For the Ising model on the erased CM, we can be more precise:

Theorem 6.2 (Magnetization in erased CM). *Fix $0 \leq \beta < \infty$ and $B > 0$. Let $I(N)$ be a sequence of random variables with values in $[N]$ such that*

$$\frac{D^{\mathrm{er}}_{I(N)}}{\log N} \xrightarrow{p} A, \tag{6.10}$$

for $N \to \infty$ and some $A \geq 0$. Then,

$$\langle \sigma_{I(N)} \rangle^{\mathrm{er}} \xrightarrow{p} \tanh(B + \beta A). \tag{6.11}$$

From this theorem we can for example conclude that if $I(N)$ is a uniform vertex in $[N]$, then $D_{I(N)}/\log N \xrightarrow{p} 0$ and hence

$$\langle \sigma_{I(N)} \rangle^{\mathrm{er}} \xrightarrow{p} \tanh(B). \tag{6.12}$$

If $I(N)$ is a sequence of supervertices, however, $D_{I(N)}/\log N \xrightarrow{p} \infty$ and hence

$$\langle \sigma_{I(N)} \rangle^{\mathrm{er}} \xrightarrow{p} 1. \tag{6.13}$$

6.3 Discussion

Scaling of the interactions. The scaling of the temperature happens in many models. For example, in the Curie-Weiss model, the interactions have to be scaled by $1/N$, because the number of edges is $N(N-1)/2$ [16]. The Sherrington-Kirkpatrick model is another example. This model is a spin glass on the complete graph, where the interactions $J_{i,j}$ are i.i.d. standard normal distributions. Here, the interactions have to be scaled with $1/\sqrt{N}$ [99].

Internal energy. It would be interesting also to compute the *internal energy* for both models. For the *original model*, we have seen that the role of normal vertices is negligible. Hence, the main contribution to the internal energy will come from the supervertices. These supervertices form a complete graph with multiple edges between each pair of vertices. It is therefore to be expected that the internal energy will behave the same as the Ising model on a complete graph with specific weights on the edges. These weights will depend on the so-called Poisson-Dirichlet distribution.

For the *erased model*, an edge selected uniformly at random will with high probability be an edge between a normal vertex and a supervertex. Since the spin of the supervertex will be $+1$ with high probability, the internal energy will behave like the magnetization of a vertex attached to a randomly selected edge, where the other side of the edge has a $+$ boundary condition.

6.4 Analysis for the original configuration model

In this section we leave out the superscript 'or' from all variables. The lower bound can easily be obtained by deleting all edges from the graph, which does not increase the magnetization by the GKS inequality. Every vertex then behaves independently and has magnetization $\tanh(B)$.

For the upper bound we distinguish between *super*vertices and *normal* vertices. Fix a sequence $\varepsilon_N \searrow 0$ arbitrary slowly. Then, we call a vertex i a supervertex if $D_i \geq \varepsilon_N N^{1/(\tau-1)}$. All other vertices are called normal vertices.

With I denoting a vertex chosen uniformly for $[N]$, we can write

$$M_N(\beta, B) = \frac{1}{N}\sum_{i=1}^{N}\langle\sigma_i\rangle = \mathbb{E}_I[\langle\sigma_I\rangle]. \tag{6.14}$$

We split the analysis into the case where I is a normal vertex and where I is a supervertex:

$$\mathbb{E}_I[\langle\sigma_I\rangle] = \mathbb{E}_I[\langle\sigma_I\rangle\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}] + \mathbb{E}_I[\langle\sigma_I\rangle\mathbb{1}_{\{D_I>\varepsilon_N N^{1/(\tau-1)}\}}]. \tag{6.15}$$

If I is a normal vertex, then we bound the magnetization by forcing all spins of its neighbors to be $+1$, which is the same as letting the magnetic field of its neighbors $B_j \to \infty$. Denote the expectation under this measure by $\langle\cdot\rangle^+$. By the second GKS inequality (Lemma 3.7),

$$\mathbb{E}_I[\langle\sigma_I\rangle\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}] \leq \mathbb{E}_I[\langle\sigma_I\rangle^+\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}]. \tag{6.16}$$

By forcing all spins of the neighbors of vertex I to be $+1$, the behavior of the spin at vertex I only depends on the number of neighbors of I and is independent of the rest of the graph. Hence, by Lemma 4.2,

$$\begin{aligned}\mathbb{E}_I[\langle\sigma_I\rangle^+\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}] &= \mathbb{E}[\tanh(B+\beta_N D_I)\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}] \\ &\leq \tanh(B) + \beta_N\mathbb{E}[D_I\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}],\end{aligned} \tag{6.17}$$

where we used that $\tanh(B+x) \leq \tanh(B)+x$, which is true because $\frac{\mathrm{d}}{\mathrm{d}x}\tanh x = 1 - \tanh^2 x \leq 1$. From Lemma 2.7,

$$\beta_N\mathbb{E}[D_I\mathbb{1}_{\{D_I\leq\varepsilon_N N^{1/(\tau-1)}\}}] \leq C_{1,\tau}\varepsilon_N^{2-\tau}\beta_N N^{(2-\tau)/(\tau-1)} = C_{1,\tau}\beta\varepsilon_N^{2-\tau} = o(1). \tag{6.18}$$

If I is a supervertex, we bound the value of its spin from above by $+1$, which is always true. Hence,

$$\mathbb{E}_I[\langle\sigma_I\rangle\mathbb{1}_{\{D_I>\varepsilon_N N^{1/(\tau-1)}\}}] \leq \mathbb{E}_I[\mathbb{1}_{\{D_I>\varepsilon_N N^{1/(\tau-1)}\}}] \leq C_{0,\tau}\varepsilon_N^{-(\tau-1)}N^{-1} = o(1), \tag{6.19}$$

by Lemma 2.7. This proves the theorem.

6.5 Analysis for the erased configuration model

First of all, we show that the magnetization of the supervertices converges to 1 in probability:

Proposition 6.3. *Let m_N be the number of vertices with degree bigger than $\varepsilon_N N^{1/(\tau-1)}$, where $\varepsilon_N \searrow 0$, such that $m_N \to \infty$ arbitrary slowly. Denote by V_i the vertex with the i-th largest degree. Then, for all $i \in [m_N]$*

$$\langle \sigma_{V_i} \rangle^{\mathrm{er}} \xrightarrow{p} 1, \tag{6.20}$$

for $N \to \infty$.

Proof. The bound

$$\langle \sigma_{V_i} \rangle^{\mathrm{er}} \leq 1, \tag{6.21}$$

holds trivially, a.s. For the lower bound we partition the normal vertices uniformly at random into m_N partitions of equal size. Denote these partitions by $\Pi_1, \ldots, \Pi_{m_N}$. Note that each of these partitions has $\Theta(N/m_N)$ vertices in it, which converges to infinity. We now only keep for all i the edges between vertex V_i and vertices in Π_i. Denote the resulting Ising expectation by $\langle \cdot \rangle^f$.

In [19], it is shown that the number of normal vertices connected to a supervertex is $\Theta(N)$ and hence the number of neighbors of V_i in Π_i is $\Theta(N/m_N)$. Hence, the remaining graph is a forest with m_N trees in it, where each tree consists of a supervertex as root and $\Theta(N/m_N)$ leaves attached to it. It thus follows from the GKS inequality and Lemma 4.2 that the magnetization of vertex V_i

$$\langle \sigma_{V_i} \rangle^{\mathrm{er}} \geq \langle \sigma_{V_i} \rangle^f \geq \tanh\left(B + \frac{cN}{m_N} \xi_{\beta_N}(B) \right). \tag{6.22}$$

It is easy to see that $\frac{cN}{m_N} \xi_{\beta_N}(B) \to \infty$, so that

$$\langle \sigma_{V_i} \rangle^f \xrightarrow{p} 1. \tag{6.23}$$

□

We can now prove the theorem for the erased configuration model:

Proof of Theorem 6.2. If $I(N)$ is a sequence of supervertices the theorem follows from Proposition 6.3, since $I(N)/\log N \to \infty$, a.s. Hence, we can assume that $I(N)$ is a sequence of normal vertices. Again, we use that normal vertices only connect to supervertices. For these supervertices, we have seen in the previous proposition that its external field goes to infinity a.s., and hence it follows from Lemma 4.2 that, conditionally on $D_{I(N)}$,

$$\langle \sigma_{I(N)} \rangle^{\mathrm{er}} = \tanh\left(B + D_{I(N)} \beta_N \right) = \tanh\left(B + \frac{D_{I(N)}}{\log N} \beta \right), \tag{6.24}$$

so that the theorem follows from the assumption on $I(N)$.

Note that to be precise, the behavior of $\sigma_{I(N)}$ is not independent from the behavior of the supervertices, but we can leave out vertex $I(N)$ from the partitioning in the proof of Proposition 6.3 and also keep the edges between vertex $I(N)$ and the supervertices. Then, the analysis in the previous proposition still goes through and the result still holds. □

7

Phase transition

We now return to the Ising model on random graphs with strongly finite mean. In this chapter we start our analysis of the critical behavior by computing the critical temperature. Recall from Chapter 3 that the critical temperature β_c is defined as that value of β where the spontaneous magnetization changes from being positive to being zero, i.e.,

$$\beta_c \equiv \inf\{\beta : M(\beta, 0^+) > 0\}. \tag{7.1}$$

7.1 Results

We first give an expression for the critical temperature:

Theorem 7.1 (Critical temperature). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P and is uniformly sparse. Then, a.s., the critical temperature β_c of $(G_N)_{N\geq 1}$ and of the random Bethe tree $\mathcal{T}(D, K, \infty)$ equals*

$$\beta_c = \operatorname{atanh}(1/\nu). \tag{7.2}$$

Note that this result implies that when $\nu \leq 1$ there is no phase transition at positive temperature. This region corresponds exactly to the regime where there is no giant component in the graph, see Section 2.6. The other extreme is when $\nu = \infty$, which is the case, e.g., if the degree distribution obeys a power law with exponent $\tau \in (2, 3]$. In that case $\beta_c = 0$ and hence the spontaneous magnetization is positive for any finite temperature.

We next show that the phase transition at this critical temperature is *continuous:*

Proposition 7.2 (Continuous phase transition). *It holds that*

$$\lim_{B \searrow 0} \mathbb{E}[\xi(h(\beta_c, B))] = 0, \qquad \text{and} \qquad \lim_{\beta \searrow \beta_c} \mathbb{E}[\xi(h(\beta, 0^+))] = 0. \tag{7.3}$$

7.2 Discussion

Regular trees. On the k-regular tree, finding a fixed point of the recursion (4.1) simplifies to the deterministic relation

$$h = B + (k-1)\xi(h). \tag{7.4}$$

This makes it easy to analyse this case and $\tanh\beta_c = 1/(k-1)$ can be obtained straightforwardly [16].

Deterministic trees. The critical temperature of the Ising model on arbitrary deterministic trees is computed in [76]. An important quantity in this computation is played by the so-called *branching number*, i.e., the average forward degree, which can be properly defined for arbitrary deterministic trees. This branching number replaces ν in the formula for the critical temperature above. The proof in [76] can easily be adapted for our setting of random trees, and hence for random graphs, although certain parts can be simplified as we show below.

Erdős-Rényi random graphs. The critical temperature for the Erdős-Rényi random graph was already obtained in [96], by showing that at $B=0$ the expression for the pressure at high temperature, i.e., for $\beta < \beta_c$, is not correct for $\beta > \beta_c$, thus proving that $\beta = \beta_c$ is a point of non-analyticity.

7.3 Preliminaries

Recall that we have derived an explicit formula for the magnetization in Corollary 5.3(a), which is obtained by differentiating the expression for the thermodynamic limit of the pressure per vertex that was first obtained. We restate this result in the next proposition and present a more intuitive proof of this result.

Proposition 7.3 (Magnetization). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P, where P has strongly finite mean, and is uniformly sparse. Then, a.s., for all $\beta \geq 0$ and $B > 0$, the thermodynamic limit of the* magnetization *per vertex exists and is given by*

$$M(\beta,B) = \mathbb{E}\left[\tanh\left(B + \sum_{i=1}^{D} \xi(h_i)\right)\right], \tag{7.5}$$

where

(i) *D has distribution P;*

(ii) *$(h_i)_{i\geq 1}$ are i.i.d. copies of the positive fixed point of the distributional recursion* (4.1)*;*

(iii) *D and $(h_i)_{i\geq 1}$ are independent.*

The same holds on the random Bethe tree $\mathcal{T}(D,K,\infty)$.

Proof. Let ϕ be a vertex picked uniformly at random from $[N]$ and $\mathbb{E}_N$ be the corresponding expectation. Then,

$$M_N(\beta,B) = \frac{1}{N}\sum_{i=1}^{N}\langle\sigma_i\rangle = \mathbb{E}_N[\langle\sigma_\phi\rangle]. \tag{7.6}$$

Denote by $\langle\cdot\rangle^{\ell,+/f}$ the expectations with respect to the Ising measure with +/free boundary conditions on vertices at graph distance ℓ from ϕ. Note that $\langle\sigma_\phi\rangle^{\ell,+/f}$ only depends on the spins of vertices in $B_\phi(\ell)$. By the GKS inequality [71],

$$\langle\sigma_\phi\rangle^{\ell,f} \le \langle\sigma_\phi\rangle \le \langle\sigma_\phi\rangle^{\ell,+}. \tag{7.7}$$

Taking the limit $N \to \infty$, the ball $B_\phi(\ell)$ has the same distribution as the random tree $\mathscr{T}(D,K,\ell)$, because of the locally tree-like nature of the graph sequence. Conditioned on the tree $\mathscr{T}$, we can prune the tree, see Lemma 4.2, to obtain that

$$\langle\sigma_\phi\rangle^{\ell,f} = \tanh\left(B + \sum_{i=1}^{D} \xi\left(h_i^{(\ell-1)}\right)\right). \tag{7.8}$$

Similarly,

$$\langle\sigma_\phi\rangle^{\ell,+} = \tanh\left(B + \sum_{i=1}^{D} \xi\left(h_i^{'(\ell-1)}\right)\right), \tag{7.9}$$

where $h_i^{'(t+1)}$ also satisfies (4.1), but has initial value $h^{'(0)} = \infty$. Since this recursion has a unique positive fixed point, see Theorem 4.1, we prove the proposition by taking the limit $\ell \to \infty$ and taking the expectation over the tree $\mathscr{T}(D,K,\infty)$. □

To study the critical behavior we investigate the function $\xi(x) = \operatorname{atanh}(\hat\beta \tanh x)$ and prove two important bounds that play a crucial role throughout the rest of Part I in this thesis:

Lemma 7.4 (Properties of $x \mapsto \xi(x)$). *For all $x, \beta \ge 0$,*

$$\hat\beta x - \frac{\hat\beta}{3(1-\hat\beta^2)} x^3 \le \xi(x) \le \hat\beta x. \tag{7.10}$$

The upper bound holds with strict inequality if $x, \beta > 0$.

Proof. By Taylor's theorem,

$$\xi(x) = \xi(0) + \xi'(0)x + \xi''(\zeta)\frac{x^2}{2}, \tag{7.11}$$

for some $\zeta \in (0,x)$. It is easily verified that $\xi(0) = 0$,

$$\xi'(0) = \left.\frac{\hat\beta(1-\tanh^2 x)}{1-\hat\beta^2\tanh^2 x}\right|_{x=0} = \hat\beta, \tag{7.12}$$

and

$$\xi''(\zeta) = -\frac{2\hat\beta(1-\hat\beta^2)(\tanh\zeta)(1-\tanh^2\zeta)}{(1-\hat\beta^2\tanh^2\zeta)^2} \le 0, \tag{7.13}$$

thus proving the upper bound. If $x, \beta > 0$ then also $\zeta > 0$ and hence the above holds with strict inequality.

For the lower bound, note that $\xi''(0) = 0$ and

$$\begin{aligned}\xi'''(\zeta) &= -\frac{2\hat\beta(1-\hat\beta^2)(1-\tanh^2\zeta)}{(1-\hat\beta^2\tanh^2\zeta)^3}\Big(1-3(1-\hat\beta^2)\tanh^2\zeta-\hat\beta^2\tanh^4\zeta\Big)\\ &\geq -\frac{2\hat\beta(1-\hat\beta^2)(1-\tanh^2\zeta)}{(1-\hat\beta^2)^2(1-\tanh^2\zeta)} = -\frac{2\hat\beta}{1-\hat\beta^2}.\end{aligned} \tag{7.14}$$

Thus, for some $\zeta \in (0,x)$,

$$\xi(x) = \xi(0)+\xi'(0)x+\xi''(0)\frac{x^2}{2}+\xi'''(\zeta)\frac{x^3}{3!} \geq \hat\beta x - \frac{2\hat\beta}{1-\hat\beta^2}\frac{x^3}{3!}. \tag{7.15}$$

□

7.4 Critical temperature

In this section we compute the critical temperature.

Proof of Theorem 7.1. Let $\beta^* = \operatorname{atanh}(1/\nu)$. We first show that if $\beta < \beta^*$, then

$$\lim_{B\searrow 0} M(\beta,B) = 0, \tag{7.16}$$

which implies that $\beta_c \geq \beta^*$. Later, we show that if $\lim_{B\searrow 0} M(\beta,B) = 0$ then $\beta \leq \beta^*$, implying that $\beta_c \leq \beta^*$.

Proof of $\beta_c \geq \beta^*$. Suppose that $\beta < \beta^*$. Then, by the fact that $\tanh x \leq x$ and Wald's identity,

$$M(\beta,B) = \mathbb{E}\left[\tanh\left(B+\sum_{i=1}^{D}\xi(h_i)\right)\right] \leq B + \mathbb{E}[D]\mathbb{E}[\xi(h)]. \tag{7.17}$$

We use the upper bound in Lemma 7.4 to get

$$\mathbb{E}[\xi(h)] = \mathbb{E}[\operatorname{atanh}(\hat\beta\tanh h)] \leq \hat\beta\mathbb{E}[h] = \hat\beta\left(B+\nu\mathbb{E}[\xi(h)]\right). \tag{7.18}$$

Further, note that

$$\mathbb{E}[\xi(h)] = \mathbb{E}[\operatorname{atanh}(\hat\beta\tanh h)] \leq \beta, \tag{7.19}$$

because $\tanh h \leq 1$. Applying inequality (7.18) ℓ times to (7.17) and subsequently using inequality (7.19) once gives

$$M(\beta,B) \leq B + B\hat\beta\mathbb{E}[D]\frac{1-(\hat\beta\nu)^\ell}{1-\hat\beta\nu} + \beta\mathbb{E}[D](\hat\beta\nu)^\ell. \tag{7.20}$$

Hence,

$$\begin{aligned}M(\beta,B) &\leq \limsup_{\ell\to\infty}\left(B + B\hat\beta\mathbb{E}[D]\frac{1-(\hat\beta\nu)^\ell}{1-\hat\beta\nu} + \beta\mathbb{E}[D](\hat\beta\nu)^\ell\right)\\ &= B\Big(1+\hat\beta\mathbb{E}[D]\frac{1}{1-\hat\beta\nu}\Big),\end{aligned} \tag{7.21}$$

because $\hat{\beta} < \hat{\beta}^* = 1/\nu$. Therefore,

$$\lim_{B\searrow 0} M(\beta,B) \leq \lim_{B\searrow 0} B\Big(1+\hat{\beta}\mathbb{E}[D]\frac{1}{1-\hat{\beta}\nu}\Big) = 0. \tag{7.22}$$

This proves the lower bound on β_c.

Proof of $\beta_c \leq \beta^*$. We adapt Lyons' proof in [76] for the critical temperature of deterministic trees to the random tree to show that $\beta_c \leq \beta^*$. Assume that $\lim_{B\searrow 0} M(\beta,B) = 0$. Note that Proposition 7.3 shows that the magnetization $M(\beta,B)$ is equal to the expectation over the random tree $\mathcal{T}(D,K,\infty)$ of the root magnetization. Hence, if we denote the root of the tree $\mathcal{T}(D,K,\infty)$ by ϕ, then it follows from our assumption on $M(\beta,B)$ that, a.s., $\lim_{B\searrow 0}\langle\sigma_\phi\rangle = 0$.

We therefore condition on the tree $T = \mathcal{T}(D,K,\infty)$ and assuming $\lim_{B\searrow 0}\langle\sigma_\phi\rangle = 0$, implies that also $\lim_{B\searrow 0} h(\phi) = 0$. Because of (4.1), we must then have, for all $v \in T$,

$$\lim_{B\searrow 0}\lim_{\ell\to\infty} h^{\ell,+}(v) = 0, \tag{7.23}$$

where we can take $+$ boundary conditions, since the recursion converges to a unique positive fixed point (Theorem 4.1). Now, fix $0 < \beta_0 < \beta$ and choose ℓ large enough and B small enough such that, for some $\varepsilon = \varepsilon(\beta_0,\beta) > 0$ that we choose later,

$$h^{\ell,+}(v) \leq \varepsilon, \tag{7.24}$$

for all $v \in T$ with $|v| = 1$, where $|v|$ denotes the graph distance from ϕ to v. Note that $h^{\ell,+}(v) = \infty > \varepsilon$ for $v \in T$ with $|v| = \ell$.

As in [76], we say that Π is a *cutset* if Π is a finite subset of $T\setminus\{\phi\}$ and every path from ϕ to infinity intersects Π at exactly one vertex $v \in \Pi$. We write $v \leq \Pi$ if every infinite path from v intersects Π and write $\sigma < \Pi$ if $\sigma \leq \Pi$ and $\sigma \notin \Pi$. Furthermore, we say that $w \leftarrow v$ if $\{w,v\}$ is an edge in T and $|w| = |v|+1$. Then, since $h^{\ell,+}(v) = \infty > \varepsilon$ for $v \in T$ with $|v| = \ell$, there is a unique cutset Π, such that $h^{\ell,+}(v) \leq \varepsilon$ for all $v \leq \Pi$, and for all $v \in \Pi$ there is at least one $w \leftarrow v$ such that $h^{\ell,+}(w) > \varepsilon$.

It follows from the lower bound in Lemma 7.4 that, for $v < \Pi$,

$$\begin{aligned} h^{\ell,+}(v) = B + \sum_{w\leftarrow v}\xi(h^{\ell,+}(w)) &\geq \sum_{w\leftarrow v}\hat{\beta}h^{\ell,+}(w) - \frac{\hat{\beta}h^{\ell,+}(w)^3}{3(1-\hat{\beta}^2)} \\ &\geq \sum_{w\leftarrow v}\hat{\beta}h^{\ell,+}(w)\Big(1-\frac{\varepsilon^2}{3(1-\hat{\beta}^2)}\Big), \end{aligned} \tag{7.25}$$

while, for $v \in \Pi$,

$$h^{\ell,+}(v) = B + \sum_{w\leftarrow v}\xi(\tanh h(w)) > \xi(\varepsilon). \tag{7.26}$$

If we now choose $\varepsilon > 0$ such that

$$\hat{\beta}\Big(1-\frac{\varepsilon^2}{3(1-\hat{\beta}^2)^2}\Big) = \hat{\beta}_0, \tag{7.27}$$

which is possible because $\beta_0 < \beta$, then,

$$h^{\ell,+}(\phi) \geq \sum_{v\in\Pi} \hat{\beta}_0^{|v|}\xi(\varepsilon). \tag{7.28}$$

Since $\xi(\varepsilon) > 0$ and $\lim_{B\searrow 0}\lim_{\ell\to\infty} h^{\ell,+}(\phi) = 0$,

$$\inf_{\Pi}\sum_{v\in\Pi} \hat{\beta}_0^{|v|} = 0. \tag{7.29}$$

From [77, Proposition 6.4] it follows that $\hat{\beta}_0 \leq 1/\nu$. This holds for all $\beta_0 < \beta$, so

$$\beta \leq \operatorname{atanh}(1/\nu) = \beta^*. \tag{7.30}$$

This proves the upper bound on β_c, thus concluding the proof. □

7.5 Continuity of the phase transition

Proof of Proposition 7.2. Note that $\lim_{B\searrow 0}\mathbb{E}[\xi(h(\beta_c,B))] = c$ exists, because $B \mapsto \mathbb{E}[\xi(h(\beta_c,B))]$ is non-decreasing and non-negative. Assume, by contradiction, that $c > 0$. By the recursion in (4.1), for $B > 0$,

$$\mathbb{E}[\xi(h(\beta,B))] = \mathbb{E}\left[\xi\left(B+\sum_{i=1}^{K}\xi(h_i(\beta,B))\right)\right] \leq \xi\left(B+\nu\mathbb{E}[\xi(h(\beta,B))]\right), \tag{7.31}$$

where the inequality holds because of Jensen's inequality and the concavity of $h \mapsto \xi(h)$. Hence,

$$c = \lim_{B\searrow 0}\mathbb{E}[\xi(h(\beta_c,B))] \leq \lim_{B\searrow 0}\xi\Big(B+\nu\mathbb{E}[\xi(h(\beta_c,B))]\Big) = \xi(\nu c). \tag{7.32}$$

Since $\xi(x) < \hat{\beta}_c x$ for $x > 0$ by Lemma 7.4 and using $\hat{\beta}_c = 1/\nu$, we obtain

$$\xi(\nu c) < \hat{\beta}_c \nu c = c, \tag{7.33}$$

leading to a contradiction.

An adaptation of this argument shows the second statement of the lemma. Again $\beta \mapsto \mathbb{E}[\xi(h(\beta,0^+))]$ is non-decreasing and non-negative and we assume that

$$\lim_{\beta\searrow\beta_c}\mathbb{E}[\xi(h(\beta,0^+))] = c > 0. \tag{7.34}$$

Then,

$$\begin{aligned} c = \lim_{\beta\searrow\beta_c}\mathbb{E}[\xi(h(\beta,0^+))] &= \lim_{\beta\searrow\beta_c}\lim_{B\searrow 0}\mathbb{E}\left[\xi\left(B+\sum_{i=1}^{K}\xi(h_i(\beta,B))\right)\right] \\ &\leq \lim_{\beta\searrow\beta_c}\lim_{B\searrow 0}\xi\Big(B+\nu\mathbb{E}[\xi(h(\beta,B))]\Big) = \xi(\nu c), \end{aligned} \tag{7.35}$$

leading again to a contradiction when $c > 0$. □

8

Critical behavior of the magnetization

In the previous chapter we have computed the value of the critical temperature and showed that this phase transition is continuous for $B = 0^+$ in the limit $\beta \searrow \beta_c$ and for $\beta = \beta_c$ in the limit $B \searrow 0$. The question remains how fast this convergence is close to the critical point. This can be described by the critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ respectively, defined in Definition 3.5. We compute the values of these critical exponents in this chapter.

8.1 Results

The values of the critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ for different values of τ are stated in the following theorem:

Theorem 8.1 (Critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P that obeys $\mathbb{E}[K^3] < \infty$ or a power law with exponent $\tau \in (3,5]$, and is uniformly sparse, or that the random Bethe tree obeys $\mathbb{E}[K^3] < \infty$ or a power law with exponent $\tau \in (3,5]$. Then, the critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ defined in Definition 3.5 exist and satisfy*

	$\tau \in (3,5)$	$\mathbb{E}[K^3] < \infty$
$\boldsymbol{\beta}$	$1/(\tau-3)$	$1/2$
$\boldsymbol{\delta}$	$\tau - 2$	3

For the boundary case $\tau = 5$ there are logarithmic corrections for $\boldsymbol{\beta} = 1/2$ and $\boldsymbol{\delta} = 3$. Indeed,

$$M(\beta, 0^+) \asymp \Big(\frac{\beta - \beta_c}{\log 1/(\beta-\beta_c)}\Big)^{1/2} \quad \text{for } \beta \searrow \beta_c, \tag{8.1}$$

and

$$M(\beta_c, B) \asymp \Big(\frac{B}{\log(1/B)}\Big)^{1/3} \quad \text{for } B \searrow 0. \tag{8.2}$$

The proof of this theorem relies on Taylor expansions performed up to the right order. For the k-regular random graph we can heuristically compute the critical exponents as follows. In this case the recursion (4.1) is deterministic and becomes

$$h = B + (k-1)\xi(h), \tag{8.3}$$

and $\hat{\beta}_c = 1/(k-1)$. Hence, by Taylor expanding $\xi(h)$ as in Lemma 7.4,

$$\xi(h) \approx \hat{\beta}h - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)h^3 = \hat{\beta}B + \hat{\beta}(k-1)\xi(h) - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)(B+(k-1)\xi(h))^3$$
$$\approx \hat{\beta}B + \hat{\beta}(k-1)\xi(h) - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)\xi(h)^3, \tag{8.4}$$

where we ignored terms that converge faster to 0 for the limits of interest. When $\beta > \beta_c$ by definition $\lim_{B\searrow 0} h = h_0 > 0$, so after taking the limit $B \searrow 0$ in (8.4) we can divide by $\xi(h_0)$ to obtain

$$1 \approx \hat{\beta}(k-1) - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)\xi(h_0)^2, \tag{8.5}$$

so that

$$\xi(h_0) \approx C(\hat{\beta}(k-1)-1)^{1/2} \asymp (\beta-\beta_c)^{1/2}. \tag{8.6}$$

When we put $\beta = \beta_c$ in (8.4) for $B > 0$, we get

$$\xi(h_c) \approx \hat{\beta}_c B + \hat{\beta}_c(k-1)\xi(h_c) - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)\xi(h_c)^3 = \hat{\beta}_c B + \xi(h_c) - \frac{1}{3}\hat{\beta}(1-\hat{\beta}^2)\xi(h_c)^3, \tag{8.7}$$

so that

$$\xi(h_c) \approx CB^{1/3}. \tag{8.8}$$

The critical exponents are now obtained by observing that $\tanh(x) \approx x$ for x small.

This explains where the values for the critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ come from, at least in the case $\mathbb{E}[K^3] < \infty$. For the critical exponents for the random graph we follow the same strategy, but then h is a random variable and higher moments of $\xi(h)$ appear by the Taylor expansions. Therefore, we first bound these higher moments of $\xi(h)$ in terms of its first moment in Section 8.3. In Section 8.4 we use these bounds to give appropriate bounds on $\mathbb{E}[\xi(h)]$ which finally allow us to compute the critical exponents $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ in Section 8.5.

For $\tau \in (2,3]$, we have that $\beta_c = 0$ as we have seen in the previous chapter. Therefore, there is not really a phase transition in the sense that there is no point of non-analyticity of the pressure and speaking about the critical behavior seems odd. Still, we can study the behavior of $M(\beta, 0^+)$ as $\beta \searrow 0$, now describing the behavior of the magnetization in the infinite temperature limit.

Theorem 8.2 (Infinite temperature limit for $\tau \in (2,3]$)**.** *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P that obeys a power law with exponent $\tau \in (2,3]$, and is uniformly sparse, or that the random Bethe tree obeys a power law with exponent $\tau \in (2,3]$. Then, for $\tau \in (2,3)$,*

$$M(\beta, 0^+) \asymp \beta^{1/(3-\tau)}, \tag{8.9}$$

while for $\tau = 3$,

$$c_0 e^{-1/(c_e\beta)} \leq M(\beta, 0^+) \leq C_0 e^{-1/(C_e\beta)}, \tag{8.10}$$

for some constants $0 < c_0, c_e, C_0, C_e < \infty$.

This theorem is proved in Section 8.6.

8.2 Discussion

Relation to the physics literature. Theorems 8.1 and 8.2 confirm the predictions in [40, 74]. For $\tau = 5$, in [40], also the logarithmic correction for $\boldsymbol{\beta} = 1/2$ in (8.1) is computed, but not that of $\boldsymbol{\delta} = 3$.

Light tails. The case $\mathbb{E}[K^3] < \infty$ includes all power-law degree distributions with $\tau > 5$, but also cases where P does *not* obey a power law. This means, e.g., that Theorem 8.1 also identifies the critical exponents for the Erdős-Rényi random graph where the degrees have an asymptotic Poisson distribution.

Inclusion of slowly varying functions. In Definition 2.2, we have assumed that the asymptotic degree distribution obeys a perfect power law. Alternatively, one could assume that $\sum_{\ell \geq k} p_\ell \asymp L(k) k^{-(\tau-1)}$ for some function $k \mapsto L(k)$ that is slowly varying at $k = \infty$. For $\tau > 5$ and any slowly varying function, we still have $\mathbb{E}[K^3] < \infty$, so the results do not change and Theorem 8.1 still holds. For $\tau \in (3, 5]$, we expect slowly varying corrections to the critical behavior in Theorem 8.1. For example, $\mathbb{E}[K^3] < \infty$ for $\tau = 5$ and $L(k) = (\log k)^{-2}$, so that the logarithmic corrections present for $\tau = 5$ disappear.

Beyond the root magnetization for the random Bethe tree. We have identified the critical value and some critical exponents for the root magnetization on the random Bethe tree. The random Bethe tree is a so-called *unimodular* graph, which is a rooted graph that often arises as the local weak limit of a sequence of graphs (in this case, the random graphs $(G_N)_{N \geq 1}$). See [10, 17] for more background on unimodular graphs and trees, in particular, $\mathcal{T}(D, K, \infty)$ is the so-called *unimodular Galton-Watson tree* as proved by Lyons, Pemantle and Peres in [78]. One would expect that the magnetization of the graph, which can be defined by

$$M_T(\beta, B) = \lim_{t \to \infty} \frac{1}{|B_\phi(t)|} \sum_{v \in B_\phi(t)} \sigma_v, \tag{8.11}$$

where $B_\phi(t)$ is the graph induced by vertices at graph distance at most t from the root ϕ and $|B_\phi(t)|$ is the number of elements in it, also converges a.s. to a limit. However, we expect that $M_T(\beta, B) \neq M(\beta, B)$ due to the special role of the root ϕ, which vanishes in the above limit. Thus one would expect that $M_T(\beta, B)$ equals the root magnetization of the tree where each vertex has degree distribution $K + 1$. Our results show that also $M_T(\beta, B)$ has the same critical temperature and critical exponents as $M(\beta, B)$.

Mean-field values. Our results show that locally tree-like random graphs with finite fourth moment of the degree distribution are in the same universality class as the mean-field model on the complete graph, which is the Curie-Weiss model [16]. We further believe that the Curie-Weiss model should enter as the limit of $k \to \infty$ for the k-regular random graph, in the sense that these have the same critical exponents (as we already know), as well as that all constants arising in asymptotics match up nicely (cf. the discussion at the end of Section 9.4). These values are also the same for the Ising model on $\mathbb{Z}^d$ for $d > 4$ with possible logarithmic corrections at $d = 4$ [3].

Further, our results show that for $\tau \in (3,5]$, the Ising model has *different* critical exponents than the ones for the Curie-Weiss model, so these constitute a set of different universality classes.

Potts model. In [41], a mean-field analysis of the Potts model on complex networks predicts that the system undergoes a first-order phase transition for $q \geq 3$ and $\tau > 3$. Hence, critical exponents for this model are not expected to exist. When $\tau \in (2,3]$ the critical temperature is again expected to be infinite so that a phase transition is absent.

8.3 Bounds on higher moments of $\xi(h)$

Throughout the rest of this chapter we assume that B is sufficiently close to zero and $\beta_c < \beta < \beta_c + \varepsilon$ for ε sufficiently small. We write $c_i, C_i, i \geq 1$ for constants that only depend on β and moments of K, and satisfy

$$0 < \liminf_{\beta \searrow \beta_c} C_i(\beta) \leq \limsup_{\beta \searrow \beta_c} C_i(\beta) < \infty. \tag{8.12}$$

Here C_i appears in upper bounds, while c_i appears in lower bounds. Furthermore, we write $e_i, i \geq 1$, for error functions that only depend on $\beta, B, \mathbb{E}[\xi(h)]$ and moments of K, and satisfy

$$\limsup_{B \searrow 0} e_i(\beta, B) < \infty \qquad \text{and} \qquad \lim_{B \searrow 0} e_i(\beta_c, B) = 0. \tag{8.13}$$

Finally, we write $\nu_k = \mathbb{E}[K(K-1)\cdots(K-k+1)]$ for the kth factorial moment of K, so that $\nu_1 = \nu$.

Lemma 8.3 (Bounds on second moment of $\xi(h)$). *Let $\beta \geq \beta_c$ and $B > 0$. Then,*

$$\mathbb{E}[\xi(h)^2] \leq \begin{cases} C_2 \mathbb{E}[\xi(h)]^2 + Be_2 & \text{when } \mathbb{E}[K^2] < \infty, \\ C_2 \mathbb{E}[\xi(h)]^2 \log(1/\mathbb{E}[\xi(h)]) + Be_2 & \text{when } \tau = 4, \\ C_2 \mathbb{E}[\xi(h)]^{\tau-2} + Be_2 & \text{when } \tau \in (3,4). \end{cases} \tag{8.14}$$

Proof. We first treat the case $\mathbb{E}[K^2] < \infty$. We use Lemma 7.4 and the recursion in (4.1) to obtain

$$\begin{aligned} \mathbb{E}[\xi(h)^2] \leq \hat{\beta}^2 \mathbb{E}[h^2] &= \hat{\beta}^2 \mathbb{E}\left[\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2\right] \\ &= \hat{\beta}^2 \left(B^2 + 2B\nu \mathbb{E}[\xi(h)] + \nu_2 \mathbb{E}[\xi(h)]^2 + \nu \mathbb{E}[\xi(h)^2]\right). \end{aligned} \tag{8.15}$$

Since $1 - \hat{\beta}^2 \nu > 0$, because β is sufficiently close to β_c and $\hat{\beta}_c = 1/\nu < 1$, the lemma holds with

$$C_2 = \frac{\hat{\beta}^2 \nu_2}{1 - \hat{\beta}^2 \nu}, \qquad \text{and} \qquad e_2 = \frac{B\hat{\beta}^2 + 2\hat{\beta}^2 \nu \mathbb{E}[\xi(h)]}{1 - \hat{\beta}^2 \nu}. \tag{8.16}$$

It is not hard to see that (8.12) holds. For e_2 the first property of (8.13) can also easily be seen. The second property in (8.13) follows from Proposition 7.2.

If $\tau \leq 4$, then $\mathbb{E}[K^2] = \infty$ and the above does not work. To analyze this case, we apply the recursion (4.1) and split the expectation over K in small and large degrees:

$$\mathbb{E}[\xi(h)^2] = \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2 \mathbb{1}_{\{K\leq \ell\}}\right] + \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2 \mathbb{1}_{\{K> \ell\}}\right]. \tag{8.17}$$

We use Lemma 7.4 to bound the first term as follows:

$$\begin{aligned}\mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2 \mathbb{1}_{\{K\leq \ell\}}\right] &\leq \hat{\beta}^2 \mathbb{E}\left[\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2 \mathbb{1}_{\{K\leq \ell\}}\right] \\ &\leq \hat{\beta}^2 \left(B^2 + 2B\nu B\mathbb{E}[\xi(h)] + \mathbb{E}[K^2 \mathbb{1}_{\{K\leq \ell\}}]\mathbb{E}[\xi(h)]^2 + \nu \mathbb{E}[\xi(h)^2]\right).\end{aligned} \tag{8.18}$$

For $\tau \in (3,4)$,

$$\mathbb{E}[K^2 \mathbb{1}_{\{K\leq \ell\}}] \leq C_{2,\tau} \ell^{4-\tau}, \tag{8.19}$$

by Lemma 2.7, while for $\tau = 4$,

$$\mathbb{E}[K^2 \mathbb{1}_{\{K\leq \ell\}}] \leq C_{2,4} \log \ell. \tag{8.20}$$

To bound the second sum in (8.17), note that $\xi(h) \leq \beta$. Hence,

$$\mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^2 \mathbb{1}_{\{K> \ell\}}\right] \leq \beta^2 \mathbb{E}[\mathbb{1}_{\{K> \ell\}}] \leq C_{0,\tau} \beta^2 \ell^{2-\tau}. \tag{8.21}$$

The optimal bound (up to a constant) can be achieved by choosing ℓ such that $\ell^{4-\tau}\mathbb{E}[\xi(h)]^2$ and $\ell^{2-\tau}$ are of the same order of magnitude. Hence, we choose $\ell = 1/\mathbb{E}[\xi(h)]$. Combining the two upper bounds then gives the desired result with

$$C_2 = \frac{1}{1-\hat{\beta}^2\nu}\left(C_{2,\tau}\hat{\beta}^2 + C_{0,\tau}\beta^2\right), \tag{8.22}$$

where we have also used that $\mathbb{E}[\xi(h)]^2 \leq \mathbb{E}[\xi(h)]^2 \log(1/\mathbb{E}[\xi(h)])$, and

$$e_2 = \frac{B\hat{\beta}^2 + 2\hat{\beta}^2\nu\mathbb{E}[\xi(h)]}{1-\hat{\beta}^2\nu}. \tag{8.23}$$

□

Remark 8.4. In the previous lemma, we showed that the correct value for the truncation is $\ell = 1/\mathbb{E}[\xi(h)]$. It turns out that this is always the correct order of magnitude, but we sometimes need to truncate at $\ell = \varepsilon/\mathbb{E}[\xi(h)]$ for ε small.

We next derive upper bounds on the third moment of $\xi(h)$:

Lemma 8.5 (Bounds on third moment of $\xi(h)$). *Let $\beta \geq \beta_c$ and $B > 0$. Then,*

$$\mathbb{E}[\xi(h)^3] \leq \begin{cases} C_3\mathbb{E}[\xi(h)]^3 + Be_3 & \text{when } \mathbb{E}[K^3] < \infty, \\ C_3\mathbb{E}[\xi(h)]^3 \log(1/\mathbb{E}[\xi(h)]) + Be_3 & \text{when } \tau = 5, \\ C_3\mathbb{E}[\xi(h)]^{\tau-2} + Be_3 & \text{when } \tau \in (3,5). \end{cases} \tag{8.24}$$

Proof. For $\mathbb{E}[K^3] < \infty$ we bound, in a similar way as in Lemma 8.3,

$$\mathbb{E}[\xi(h)^3] \leq \hat{\beta}^3 \Big(B^3 + 3B^2 \nu \mathbb{E}[\xi(h)] + 3B\nu_2 \mathbb{E}[\xi(h)]^2 + 3B\nu \mathbb{E}[\xi(h)^2] \\ + \nu_3 \mathbb{E}[\xi(h)]^3 + 3\nu_2 \mathbb{E}[\xi(h)]\mathbb{E}[\xi(h)^2] + \hat{\beta}^3 \nu \mathbb{E}[\xi(h)^3] \Big). \tag{8.25}$$

Using (8.14), we indeed get the bound

$$\mathbb{E}[\xi(h)^3] \leq C_3 \mathbb{E}[\xi(h)]^3 + Be_3, \tag{8.26}$$

where

$$C_3 = \frac{\hat{\beta}^3}{1 - \hat{\beta}^3 \nu} \left(\nu_3 + 3\nu_2 C_2\right), \tag{8.27}$$

and

$$e_3 = \frac{\hat{\beta}^3}{1 - \hat{\beta}^3 \nu} \left(B^2 + 3B\nu e_2 + 3\left(B\nu + \nu_2 e_2\right) \mathbb{E}[\xi(h)] + 3\left(\nu_2 + \nu C_2\right) \mathbb{E}[\xi(h)]^2 \right). \tag{8.28}$$

For $\tau \in (3,5]$, we use the recursion (4.1) and the expectation split in small and large values of K to obtain

$$\mathbb{E}[\xi(h)^3] = \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^3 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}} \right] + \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^3 \mathbb{1}_{\{K > \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}} \right]. \tag{8.29}$$

We bound the first sum from above by

$$\hat{\beta}^3 \mathbb{E}\left[\left(B + \sum_{i=1}^{K} \xi(h_i)\right)^3 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}} \right] \\ = \hat{\beta}^3 \Big(B^3 + 3B^2 \mathbb{E}[K \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)] + 3B \mathbb{E}[K(K-1) \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)^2] \\ + 3B \mathbb{E}[K \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)^2] + \mathbb{E}[K(K-1)(K-2) \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)]^3 \\ + 3\mathbb{E}[K(K-1) \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)] \mathbb{E}[\xi(h)^2] + \mathbb{E}[K \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \mathbb{E}[\xi(h)^3] \Big).$$

By Lemma 2.7, for $\tau \in (3,5)$,

$$\mathbb{E}[K^3 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \leq C_{3,\tau} \mathbb{E}[\xi(h)]^{\tau - 5}, \tag{8.30}$$

while, for $\tau = 5$,

$$\mathbb{E}[K^3 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \leq C_{3,5} \left(1 + \log(1/\mathbb{E}[\xi(h)])\right). \tag{8.31}$$

Similarly, by Lemma 2.7, for $\tau \in (3,4)$,

$$\mathbb{E}[K^2 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \leq C_{2,\tau} \mathbb{E}[\xi(h)]^{\tau - 4}, \tag{8.32}$$

while, for $\tau = 4$,

$$\mathbb{E}[K^2 \mathbb{1}_{\{K \leq \lfloor 1/\mathbb{E}[\xi(h)] \rfloor\}}] \leq C_{2,4} \left(1 + \log(1/\mathbb{E}[\xi(h)])\right). \tag{8.33}$$

For the other terms we can replace the upper bound in the sum by infinity and use the upper bound on $\mathbb{E}[\xi(h)^2]$ of Lemma 8.3. For the second sum in (8.29) we bound $\xi(x) \leq \beta$, so that this sum is bounded from above by $C_{0,\tau} \mathbb{E}[\xi(h)]^{\tau-2}$. Combining these bounds gives the desired result. □

8.4 Bounds on the first moment of $\xi(h)$

Proposition 8.6 (Upper bound on first moment of $\xi(h)$)**.** *Let $\beta \geq \beta_c$ and $B > 0$. Then, there exists a $C_1 > 0$ such that*

$$\mathbb{E}[\xi(h)] \leq \beta B + \hat{\beta}\nu\mathbb{E}[\xi(h)] - C_1\mathbb{E}[\xi(h)]^{\delta}, \tag{8.34}$$

where

$$\delta = \begin{cases} 3 & \text{when } \mathbb{E}[K^3] < \infty, \\ \tau - 2 & \text{when } \tau \in (3,5]. \end{cases} \tag{8.35}$$

For $\tau = 5$,

$$\mathbb{E}[\xi(h)] \leq \beta B + \hat{\beta}\nu\mathbb{E}[\xi(h)] - C_1\mathbb{E}[\xi(h)]^3 \log(1/\mathbb{E}[\xi(h)]). \tag{8.36}$$

Proof. We first use recursion (4.1) and rewrite it as

$$\mathbb{E}[\xi(h)] = \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K}\xi(h_i)\right)\right] = \hat{\beta}B + \hat{\beta}\nu\mathbb{E}[\xi(h)] + T_1 + T_2, \tag{8.37}$$

where

$$T_1 = \mathbb{E}\Big[\xi\Big(B + K\mathbb{E}[\xi(h)]\Big) - \hat{\beta}\left(B + K\mathbb{E}[\xi(h)]\right)\Big], \tag{8.38}$$

and

$$T_2 = \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K}\xi(h_i)\right) - \xi\left(B + K\mathbb{E}[\xi(h)]\right)\right]. \tag{8.39}$$

Here, T_1 can be seen as the error of a first-order Taylor series approximation around 0 of $\xi(B + K\mathbb{E}[\xi(h)])$, whereas T_2 is the error made by replacing $\xi(h_i)$ by its expected value in the sum. By concavity of $x \mapsto \xi(x)$, both random variables in the expectation of T_1 and T_2 are non-positive. In particular, $T_2 \leq 0$, which is enough for our purposes. We next bound T_1 in the cases where $\mathbb{E}[K^3] < \infty$, $\tau \in (3,5)$, and $\tau = 5$ separately.

Bound on T_1 when $\mathbb{E}[K^3] < \infty$. To bound T_1 for $\mathbb{E}[K^3] < \infty$ we use that, a.s.,

$$\xi\left(B + K\mathbb{E}[\xi(h)]\right) - \hat{\beta}\left(B + K\mathbb{E}[\xi(h)]\right) \leq 0, \tag{8.40}$$

which follows from Lemma 7.4. Hence,

$$T_1 \leq \mathbb{E}\left[\Big(\xi\left(B + K\mathbb{E}[\xi(h)]\right) - \hat{\beta}\left(B + K\mathbb{E}[\xi(h)]\right)\Big)\mathbb{1}_{\{B+K\mathbb{E}[\xi(h)]\leq \mathrm{atanh}\frac{1}{2}\}}\right]. \tag{8.41}$$

Since $\xi''(0) = 0$, it follows from Taylor's theorem that, a.s.,

$$\xi\left(B + K\mathbb{E}[\xi(h)]\right) - \hat{\beta}\left(B + K\mathbb{E}[\xi(h)]\right) = \frac{\xi'''(\zeta)}{6}\left(B + K\mathbb{E}[\xi(h)]\right)^3, \tag{8.42}$$

for some $\zeta \in (0, B + K\mathbb{E}[\xi(h)])$. If $B + K\mathbb{E}[\xi(h)] \leq \mathrm{atanh}\frac{1}{2}$, then

$$\begin{aligned}\xi'''(\zeta) &= -\frac{2\hat{\beta}(1-\hat{\beta}^2)(1-\tanh^2\zeta)}{(1-\hat{\beta}^2\tanh^2\zeta)^3}\left(1 - 3(1-\hat{\beta}^2)\tanh^2\zeta - \hat{\beta}^2\tanh^4\zeta\right) \\ &\leq -\frac{9}{32}\hat{\beta}(1-\hat{\beta}^2).\end{aligned} \tag{8.43}$$

Hence,

$$T_1 \le -\frac{3}{64}\hat\beta(1-\hat\beta^2)\mathbb{E}\left[(B+K\mathbb{E}[\xi(h)])^3\,\mathbb{1}_{\{B+K\mathbb{E}[\xi(h)]\le \mathrm{atanh}\frac{1}{2}\}}\right]$$
$$\le -\frac{3}{64}\hat\beta(1-\hat\beta^2)\mathbb{E}[K^3\mathbb{1}_{\{K\mathbb{E}[\xi(h)]\le \mathrm{atanh}\frac{1}{2}-B\}}]\mathbb{E}[\xi(h)]^3. \tag{8.44}$$

Bound on T_1 when $\tau\in(3,5]$. For $\tau\in(3,5]$, we make the expectation over K explicit:

$$T_1=\sum_{k=0}^{\infty}\rho_k\left(\xi\left(B+k\mathbb{E}[\xi(h)]\right)-\hat\beta\left(B+k\mathbb{E}[\xi(h)]\right)\right), \tag{8.45}$$

where it should be noted that all terms in this sum are negative because of Lemma 7.4. Define $f(k)=\xi\left(B+k\mathbb{E}[\xi(h)]\right)-\hat\beta\left(B+k\mathbb{E}[\xi(h)]\right)$ and note that $f(k)$ is non-increasing. We use (2.15) and Lemma 2.6 to rewrite

$$T_1=\sum_{k=0}^{\infty}f(k)\rho_k=f(0)+\sum_{k\ge1}[f(k)-f(k-1)]\rho_{\ge k}$$
$$\le f(0)+c_\rho\sum_{k\ge1}[f(k)-f(k-1)](k+1)^{-(\tau-2)}. \tag{8.46}$$

Then, use (2.15) in reverse to rewrite this as

$$T_1\le f(0)+c_\rho\sum_{k\ge0}f(k)[k^{-(\tau-2)}-(k+1)^{-(\tau-2)}]$$
$$\le f(0)(1-c_\rho\sum_{k\ge1}k^{-\tau})+(\tau-1)c_\rho\sum_{k\ge0}f(k)(k+1)^{-(\tau-1)}. \tag{8.47}$$

Hence, with $e=f(0)(1-c_\rho\sum_{k\ge1}k^{-\tau})/B$,

$$T_1\le eB+(\tau-1)c_\rho\left(\mathbb{E}[\xi(h)]\right)^{\tau-1}$$
$$\times\sum_{k=0}^{\infty}((k+1)\mathbb{E}[\xi(h)])^{-(\tau-1)}\left(\xi\left(B+k\mathbb{E}[\xi(h)]\right)-\hat\beta\left(B+k\mathbb{E}[\xi(h)]\right)\right)$$
$$\le eB+(\tau-1)c_\rho\left(\mathbb{E}[\xi(h)]\right)^{\tau-1} \tag{8.48}$$
$$\times\sum_{k=a/\mathbb{E}[\xi(h)]}^{b/\mathbb{E}[\xi(h)]}(k\mathbb{E}[\xi(h)])^{-(\tau-1)}\left(\xi\left(B+k\mathbb{E}[\xi(h)]\right)-\hat\beta\left(B+k\mathbb{E}[\xi(h)]\right)\right),$$

where we choose a and b such that $0<a<b<\infty$. We use dominated convergence on the above sum. The summands are uniformly bounded, and $\mathbb{E}[\xi(h)]\to0$ for both limits of interest. Further, when $k\mathbb{E}[\xi(h)]=y$, the summand converges pointwise to $y^{-(\tau-1)}\left(\xi\left(B+y\right)-\hat\beta\left(B+y\right)\right)$. Hence, we can write the sum above as

$$\mathbb{E}[\xi(h)]^{-1}\left(\int_a^b y^{-(\tau-1)}\left(\xi\left(B+y\right)-\hat\beta\left(B+y\right)\right)\mathrm{d}y+o(1)\right), \tag{8.49}$$

where $o(1)$ is a function tending to zero for both limits of interest [67, 216 A]. The integrand is uniformly bounded for $y\in[a,b]$ and hence we can bound the integral from above by a (negative) constant $-I$ for B sufficiently small and β sufficiently close to β_c. Hence,

$$\mathbb{E}[\xi(h)] \leq \hat{\beta}B + \hat{\beta}\nu\mathbb{E}[\xi(h)] - (\tau-1)c_\rho I\mathbb{E}[\xi(h)]^{\tau-2}. \tag{8.50}$$

Logarithmic corrections in the bound for $\tau = 5$. We complete the proof by identifying the logarithmic correction for $\tau = 5$. Since the random variable in the expectation in T_1 is non-positive, we can bound

$$T_1 \leq \mathbb{E}\left[\xi\left(B + K\mathbb{E}[\xi(h)]\right) - \hat{\beta}\left(B + K\mathbb{E}[\xi(h)]\right)\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \tag{8.51}$$

Taylor expanding $h \mapsto \xi(h)$ to third order, using that $\xi(0) = \xi''(0) = 0$, while the linear term cancels, leads to

$$T_1 \leq \mathbb{E}\left[\frac{\xi'''(\zeta)}{6}\left(B + K\mathbb{E}[\xi(h)]\right)^3\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right], \tag{8.52}$$

for some $\zeta \in (0, K\mathbb{E}[\xi(h)])$. On the event that $K \leq \varepsilon/\mathbb{E}[\xi(h)]$, we thus have that $\zeta \in (0,\varepsilon)$, and $\xi'''(\zeta) \geq c_\varepsilon \equiv \inf_{x\in(0,\varepsilon)}\xi'''(x)$ when ε is sufficiently small. Thus,

$$\begin{aligned} T_1 &\leq \frac{c_\varepsilon}{6}\mathbb{E}\left[\left(B + K\mathbb{E}[\xi(h)]\right)^3\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right] \\ &\leq \frac{c_\varepsilon}{6}\mathbb{E}[\xi(h)]^3\mathbb{E}\left[K(K-1)(K-2)\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \end{aligned} \tag{8.53}$$

When $\tau = 5$, it holds that $\mathbb{E}\left[K(K-1)(K-2)\mathbb{1}_{\{K\leq\ell\}}\right] \geq c_{3,5}\log\ell$, as we show now. We compute, using (2.15) with $f(k) = k(k-1)(k-2)$,

$$\mathbb{E}\left[K(K-1)(K-2)\mathbb{1}_{\{K\leq\ell\}}\right] = \sum_{k=1}^{\infty}[f(k) - f(k-1)]\sum_{i=k}^{\ell}\rho_i = \sum_{k=3}^{\infty}3(k-1)(k-2)\sum_{i=k}^{\ell}\rho_i. \tag{8.54}$$

We bound this from below by

$$\mathbb{E}\left[K(K-1)(K-2)\mathbb{1}_{\{K\leq\ell\}}\right] \geq \sum_{k=0}^{\sqrt{\ell}}3(k-1)(k-2)[\rho_{\geq k} - \rho_{\geq\ell}]. \tag{8.55}$$

By Lemma 2.6, for $\tau = 5$, the contribution due to $\rho_{\geq\ell}$ is at most

$$\ell^{3/2}\rho_{\geq\ell} \leq C_\rho\ell^{-3/2} = o(1), \tag{8.56}$$

while the contribution due to $\rho_{\geq k}$ and using $3(k-1)(k-2) \geq k^2$ for every $k \geq 4$, is at least

$$c_\rho\sum_{k=4}^{\sqrt{\ell}}k^{-1} \geq c_\rho\int_4^{\sqrt{\ell}+1}\frac{\mathrm{d}x}{x} = c_\rho[\log(\sqrt{\ell}+1) - \log 4], \tag{8.57}$$

which proves the claim by choosing the constant $c_{3,5}$ correctly. □

Proposition 8.7 (Lower bound on first moment of $\xi(h)$)**.** *Let $\beta \geq \beta_c$ and $B > 0$. Then, there exists a constant $C_2 > 0$ such that*

$$\mathbb{E}[\xi(h)] \geq \beta B + \hat{\beta}\nu\mathbb{E}[\xi(h)] - c_1\mathbb{E}[\xi(h)]^\delta - Be_1, \tag{8.58}$$

where

$$\delta = \begin{cases} 3 & \text{when } \mathbb{E}[K^3] < \infty, \\ \tau - 2 & \text{when } \tau \in (3,5). \end{cases} \tag{8.59}$$

For $\tau = 5$*,*

$$\mathbb{E}[\xi(h)] \geq \beta B + \hat{\beta}\nu\mathbb{E}[\xi(h)] - C_2\mathbb{E}[\xi(h)]^3\log(1/\mathbb{E}[\xi(h)]) - Be_1. \tag{8.60}$$

Proof. We again use the split in (8.37) and we bound T_1 and T_2.

The lower bound on T_1. For $\mathbb{E}[K^3] < \infty$, we use the lower bound of Lemma 7.4 to get

$$T_1 \geq -\frac{\hat{\beta}}{3(1-\hat{\beta}^2)}\mathbb{E}\left[(B + K\mathbb{E}[\xi(h)])^3\right]. \tag{8.61}$$

By expanding, this can be rewritten as

$$T_1 \geq -\frac{\hat{\beta}}{3(1-\hat{\beta}^2)}\mathbb{E}[K^3]\mathbb{E}[\xi(h)]^3 - Be_4. \tag{8.62}$$

For $\tau \in (3,5]$, we first split T_1 in a small K and a large K part. For this, write

$$t_1(k) = \xi\left(B + k\mathbb{E}[\xi(h)]\right) - \hat{\beta}\left(B + k\mathbb{E}[\xi(h)]\right). \tag{8.63}$$

Then,

$$T_1 = \mathbb{E}[t_1(K)] = \mathbb{E}\left[t_1(K)\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right] + \mathbb{E}\left[t_1(K)\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \tag{8.64}$$

To bound the first term, we again use (8.61):

$$\mathbb{E}\left[t_1(K)\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right] \geq -\frac{\hat{\beta}}{3(1-\hat{\beta}^2)}\mathbb{E}\left[(B + K\mathbb{E}[\xi(h)])^3\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \tag{8.65}$$

It is easy to see that the terms $B^3\mathbb{E}\left[\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}\right]$ and $3B^2\mathbb{E}[\xi(h)]\mathbb{E}\left[K\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]$ that we get by expanding the above are of the form Be. To bound the other two terms, we use Lemma 2.7 to obtain, for $\varepsilon \leq 1$,

$$3B\mathbb{E}[\xi(h)]^2\mathbb{E}\left[K^2\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right] \leq \begin{cases} 3B\mathbb{E}[\xi(h)]^2\mathbb{E}\left[K^2\right] & \text{when } \tau \in (4,5], \\ 3BC_{2,4}\mathbb{E}[\xi(h)]^2\log(1/\mathbb{E}[\xi(h)]) & \text{when } \tau = 4, \\ 3BC_{2,\tau}\mathbb{E}[\xi(h)]^{\tau-2} & \text{when } \tau \in (3,4), \end{cases} \tag{8.66}$$

which are all of the form Be, and

$$\mathbb{E}\left[K^3\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]\mathbb{E}[\xi(h)]^3 \leq \begin{cases} C_{3,5}\mathbb{E}[\xi(h)]^3\log(1/\mathbb{E}[\xi(h)]) & \text{when } \tau = 5, \\ C_{3,\tau}\mathbb{E}[\xi(h)]^{\tau-2} & \text{when } \tau \in (3,5). \end{cases} \tag{8.67}$$

To bound T_1 for large K, we observe that

$$\mathbb{E}\left[t_1(K)\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}\right] \geq -\hat{\beta}B\mathbb{E}[\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}] - \hat{\beta}\mathbb{E}[\xi(h)]\mathbb{E}[K\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}]. \quad (8.68)$$

Applying Lemma 2.7 now gives, for $\tau \in (3,5]$

$$\begin{aligned}\mathbb{E}\left[t_1(K)\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}\right] &\geq -\hat{\beta}BC_{0,\tau}\mathbb{E}[\xi(h)]^{\tau-2} - \hat{\beta}C_{1,\tau}\mathbb{E}[\xi(h)]^{\tau-2} \\ &= -C_4\mathbb{E}[\xi(h)]^{\tau-2} - Be_4. \end{aligned} \quad (8.69)$$

The lower bound on T_2. To bound T_2, we split in a small and a large K contribution:

$$T_2 = \mathbb{E}[t_2(K)\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}] + \mathbb{E}[t_2(K)\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}] \equiv T_2^{\leq} + T_2^{>}, \quad (8.70)$$

where

$$t_2(k) = \xi\left(B + \sum_{i=1}^{k}\xi(h_i)\right) - \xi\left(B + k\mathbb{E}[\xi(h)]\right). \quad (8.71)$$

To bound $T_2^{>}$, we note that

$$t_2(k) \geq -\beta, \quad (8.72)$$

so that

$$T_2^{>} \geq -\beta\mathbb{E}[\mathbb{1}_{\{K>\varepsilon/\mathbb{E}[\xi(h)]\}}] \geq -C_5\mathbb{E}[\xi(h)]^{(\tau-2)\wedge 3}, \quad (8.73)$$

where we have used Lemma 2.7 in the last inequality and the Markov inequality when $\mathbb{E}[K^3] < \infty$.

It remains to bound $T_2^{\leq}$. This can also be seen as a Taylor expansion of $\xi(B + \sum_{i=1}^{K}\xi(h_i))$ around $B + K\mathbb{E}[\xi(h)]$. Note that, a.s.,

$$\mathbb{E}\left[\xi'(B + K\mathbb{E}[\xi(h)])\left(\sum_{i=1}^{K}\xi(h_i) - K\mathbb{E}[\xi(h)]\right) \,\Big|\, K\right] = 0, \quad (8.74)$$

and hence also the expectation over K of the above equals 0. Thus, for some $\zeta \in \left(B + \sum_{i=1}^{K}\xi(h_i), B + K\mathbb{E}[\xi(h)]\right)$,

$$T_2^{\leq} = \mathbb{E}\left[\frac{\xi''(\zeta)}{2}\left(\sum_{i=1}^{K}\xi(h_i) - K\mathbb{E}[\xi(h)]\right)^2 \mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \quad (8.75)$$

We use that, for some ζ in between $B + \sum_{i=1}^{K}\xi(h_i)$ and $B + K\mathbb{E}[\xi(h)]$,

$$\xi''(\zeta) \geq -\frac{2\hat{\beta}}{1-\hat{\beta}^2}\left(B + \sum_{i=1}^{K}\xi(h_i) + K\mathbb{E}[\xi(h)]\right), \quad (8.76)$$

to obtain

$$\begin{aligned}
T_2^{\leq} &\geq -\frac{\hat\beta}{1-\hat\beta^2}\mathbb{E}\Big[\Big(B+\sum_{i=1}^K \xi(h_i)+K\mathbb{E}[\xi(h)]\Big)\Big(\sum_{i=1}^K \xi(h_i)-K\mathbb{E}[\xi(h)]\Big)^2 \mathbb{1}_{\{K\leq \varepsilon/\mathbb{E}[\xi(h)]\}}\Big]\\
&\geq -\frac{\hat\beta}{1-\hat\beta^2}\Big(B\nu\mathbb{E}\left[(\xi(h)-\mathbb{E}[\xi(h)])^2\right]+\mathbb{E}[K\mathbb{1}_{\{K\leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}\left[(\xi(h)-\mathbb{E}[\xi(h)])^3\right]\\
&\quad +2\mathbb{E}[K^2\mathbb{1}_{\{K\leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)]\mathbb{E}\left[(\xi(h)-\mathbb{E}[\xi(h)])^2\right]\Big)\\
&\geq -\frac{\hat\beta}{1-\hat\beta^2}\Big(B\nu\mathbb{E}[\xi(h)^2]+2\mathbb{E}[K^2\mathbb{1}_{\{K\leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)]\mathbb{E}[\xi(h)^2]+\nu\mathbb{E}[\xi(h)^3]\Big).
\end{aligned} \tag{8.77}$$

Using the bounds of Lemmas 8.3 and 8.5 we get,

$$T_2^{\leq} \geq \begin{cases} -\frac{\hat\beta}{1-\hat\beta^2}\left(2\mathbb{E}[K^2]C_2+C_3\nu\right)\mathbb{E}[\xi(h)]^3-Be_5 & \text{when } \mathbb{E}[K^3]<\infty,\\ -\frac{\hat\beta}{1-\hat\beta^2}\left(2\mathbb{E}[K^2]C_2+C_3\nu\right)\mathbb{E}[\xi(h)]^3\log(1/\mathbb{E}[\xi(h)])-Be_5 & \text{when } \tau=5,\\ -\frac{\hat\beta}{1-\hat\beta^2}\left(C'_{2,\tau}+C_3\nu\right)\mathbb{E}[\xi(h)]^{\tau-2}-Be_5 & \text{when } \tau\in(3,5), \end{cases} \tag{8.78}$$

where $C'_{2,\tau}=\mathbb{E}[K^2]C_2$ for $\tau\in(4,5)$ and $C'_{2,\tau}=C_2$ for $\tau\in(3,4]$. Here, we have also used that (a) $\mathbb{E}[\xi(h)]^3\leq \mathbb{E}[\xi(h)]^3\log(1/\mathbb{E}[\xi(h)])$ for $\tau=5$; (b) $\mathbb{E}[\xi(h)]^3\leq\mathbb{E}[\xi(h)]^{\tau-2}$ for $\tau\in(4,5]$; and (c) $\mathbb{E}[K^2\mathbb{1}_{\{K\leq\varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)]\leq\varepsilon\nu\leq\nu$ for $\tau\in(3,4]$.

Combining the bounds on T_1 and T_2 gives the desired lower bound on $\mathbb{E}[\xi(h)]$. □

8.5 The critical exponents β and δ

It remains to show that the bounds on $\mathbb{E}[\xi(h)]$ give us the desired result:

Theorem 8.8 (Values of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$). *The critical exponent $\boldsymbol{\beta}$ equals*

$$\boldsymbol{\beta}=\begin{cases} 1/2 & \text{when } \mathbb{E}[K^3]<\infty,\\ 1/(\tau-3) & \text{when } \tau\in(3,5), \end{cases} \tag{8.79}$$

and the critical exponent $\boldsymbol{\delta}$ equals

$$\boldsymbol{\delta}=\begin{cases} 3 & \text{when } \mathbb{E}[K^3]<\infty,\\ \tau-2 & \text{when } \tau\in(3,5). \end{cases} \tag{8.80}$$

For $\tau=5$,

$$M(\beta,0^+)\asymp\left(\frac{\beta-\beta_c}{\log(1/(\beta-\beta_c))}\right)^{1/2} \quad \text{for } \beta\searrow\beta_c, \tag{8.81}$$

and

$$M(\beta_c,B)\asymp\left(\frac{B}{\log(1/B)}\right)^{1/3} \quad \text{for } B\searrow 0. \tag{8.82}$$

Proof. We prove the upper and the lower bounds separately, starting with the upper bound.

The upper bounds on the magnetization. We start by bounding the magnetization from above:

$$M(\beta,B)=\mathbb{E}\left[\tanh\left(B+\sum_{i=1}^{D}\xi(h_i)\right)\right]\leq B+\mathbb{E}[D]\mathbb{E}[\xi(h)]. \tag{8.83}$$

We first perform the analysis for $\boldsymbol{\beta}$. Taking the limit $B\searrow 0$ in (8.34) in Proposition 8.6 yields

$$\mathbb{E}[\xi(h_0)]\leq\hat{\beta}\nu\mathbb{E}[\xi(h_0)]-C_1\mathbb{E}[\xi(h_0)]^{\boldsymbol{\delta}}, \tag{8.84}$$

where $h_0=h(\beta,0^+)$. For $\beta>\beta_c$, by definition, $\mathbb{E}[\xi(h_0)]>0$ and thus we can divide by $\mathbb{E}[\xi(h_0)]$ to obtain

$$\mathbb{E}[\xi(h_0)]^{\boldsymbol{\delta}-1}\leq\frac{\hat{\beta}\nu-1}{C_1}. \tag{8.85}$$

By Taylor's theorem,

$$\hat{\beta}\nu-1\leq\nu(1-\hat{\beta}_c^2)(\beta-\beta_c). \tag{8.86}$$

Hence,

$$\mathbb{E}[\xi(h_0)]\leq\left(\frac{\nu(1-\hat{\beta}_c^2)}{C_1}\right)^{1/(\boldsymbol{\delta}-1)}(\beta-\beta_c)^{1/(\boldsymbol{\delta}-1)}. \tag{8.87}$$

Using that $\boldsymbol{\beta}=1/(\boldsymbol{\delta}-1)$,

$$M(\beta,0^+)\leq\mathbb{E}[D]\left(\frac{\nu(1-\hat{\beta}_c^2)}{C_1}\right)^{\boldsymbol{\beta}}(\beta-\beta_c)^{\boldsymbol{\beta}}, \tag{8.88}$$

from which it easily follows that

$$\limsup_{\beta\searrow\beta_c}\frac{M(\beta,0^+)}{(\beta-\beta_c)^{\boldsymbol{\beta}}}<\infty. \tag{8.89}$$

We complete the analysis for $\boldsymbol{\beta}$ by analyzing $\tau=5$. Since (8.34) also applies to $\tau=5$, (8.89) holds as well. We now improve upon this using (8.36) in Proposition 8.6, which yields in a similar way as above that

$$\mathbb{E}[\xi(h_0)]^2\leq\frac{\hat{\beta}\nu-1}{C_1\log(1/\mathbb{E}[\xi(h_0)])}. \tag{8.90}$$

Since $x\mapsto 1/\log(1/x)$ is increasing on $(0,1)$ and $\mathbb{E}[\xi(h_0)]\leq C(\beta-\beta_c)^{1/2}$ for some $C>0$, we immediately obtain that

$$\mathbb{E}[\xi(h_0)]^2\leq\frac{\hat{\beta}\nu-1}{C_1\log(1/\mathbb{E}[\xi(h_0)])}\leq\frac{\hat{\beta}\nu-1}{C_1\log(1/[C(\beta-\beta_c)^{1/2}])}. \tag{8.91}$$

Taking the limit of $\beta\searrow\beta_c$ as above then completes the proof.

We continue with the analysis for $\boldsymbol{\delta}$. Setting $\beta=\beta_c$ in (8.34) and rewriting gives

$$\mathbb{E}[\xi(h_c)]\leq\left(\frac{\hat{\beta}_c}{C_1}\right)^{1/\boldsymbol{\delta}}B^{1/\boldsymbol{\delta}}, \tag{8.92}$$

with $h_c = h(\beta_c, B)$. Hence,

$$M(\beta_c, B) \leq B + \mathbb{E}[D] \left(\frac{\hat{\beta}_c}{C_1} \right)^{1/\delta} B^{1/\delta}, \tag{8.93}$$

so that, using $1/\delta < 1$,

$$\limsup_{B \searrow 0} \frac{M(\beta_c, B)}{B^{1/\delta}} < \infty. \tag{8.94}$$

The analysis for δ for $\tau = 5$ can be performed in an identical way as for β.

The lower bounds on the magnetization. For the lower bound on the magnetization we use that

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} \tanh x = -2 \tanh x (1 - \tanh^2 x) \geq -2, \tag{8.95}$$

so that

$$\tanh x \geq x - x^2. \tag{8.96}$$

Hence,

$$\begin{aligned} M(\beta, B) &\geq B + \mathbb{E}[D]\mathbb{E}[\xi(h)] - \mathbb{E}\left[\left(B + \sum_{i=1}^{D} \xi(h_i) \right)^2 \right] \\ &\geq B + \mathbb{E}[D]\mathbb{E}[\xi(h)] - Be_6 - \mathbb{E}[D(D-1)]\mathbb{E}[\xi(h)]^2 - \mathbb{E}[D]C_2\mathbb{E}[\xi(h)]^{2\wedge(\tau-2)} \\ &= B + (\mathbb{E}[D] - e_7)\mathbb{E}[\xi(h)] - Be_6, \end{aligned} \tag{8.97}$$

with $a \wedge b$ denoting the minimum of a and b, because $\mathbb{E}[\xi(h)]$ converges to zero for both limits of interest.

We again first perform the analysis for β and $\tau \neq 5$. We get from (8.58) in Proposition 8.7 that

$$\mathbb{E}[\xi(h_0)] \geq \left(\frac{\hat{\beta}\nu - 1}{c_1} \right)^{1/(\delta-1)} \geq \left(\frac{\nu(1 - \hat{\beta}^2)}{c_1} \right)^{\beta} (\beta - \beta_c)^{\beta}, \tag{8.98}$$

where the last inequality holds because, by Taylor's theorem,

$$\hat{\beta}\nu - 1 \geq \nu(1 - \hat{\beta}^2)(\beta - \beta_c). \tag{8.99}$$

Hence,

$$\liminf_{\beta \searrow \beta_c} \frac{M(\beta, 0^+)}{(\beta - \beta_c)^{\beta}} \geq \mathbb{E}[D] \left(\frac{\nu(1 - \hat{\beta}^2)}{c_1} \right)^{\beta} > 0. \tag{8.100}$$

For $\tau = 5$, we note that (8.60) as well as the fact that $\log 1/x \leq A_\varepsilon x^{-\varepsilon}$ for all $x \in (0,1)$ and some $A_\varepsilon > 0$, yields that

$$\mathbb{E}[\xi(h_0)] \geq \left(\frac{\hat{\beta}\nu - 1}{A_\varepsilon c_1} \right)^{1/(2+\varepsilon)} \geq \left(\frac{\nu(1 - \hat{\beta}^2)}{A_\varepsilon c_1} \right)^{1/(2+\varepsilon)} (\beta - \beta_c)^{1/(2+\varepsilon)}. \tag{8.101}$$

Then again using (8.60) yields, for some constant $c > 0$,

$$\mathbb{E}[\xi(h_0)] \geq \left(\frac{\hat{\beta}\nu - 1}{c_1 \log(1/\mathbb{E}[\xi(h_0)])} \right)^{1/2} \geq c \left(\frac{\beta - \beta_c}{\log(1/(\beta - \beta_c))} \right)^{1/2}, \tag{8.102}$$

once more since $x \mapsto 1/(\log(1/x))$ is increasing.

We continue with the analysis for $\boldsymbol{\delta}$. Again, setting $\beta = \beta_c$ in (8.58), we get

$$\mathbb{E}[\xi(h_c)] \geq \left(\frac{\hat{\beta}_c - e_1}{c_1} \right)^{1/\boldsymbol{\delta}} B^{1/\boldsymbol{\delta}}, \tag{8.103}$$

from which it follows that

$$\liminf_{B \searrow 0} \frac{M(\beta_c, B)}{B^{1/\boldsymbol{\delta}}} \geq \mathbb{E}[D] \left(\frac{\hat{\beta}_c}{c_1} \right)^{1/\boldsymbol{\delta}} > 0, \tag{8.104}$$

as required. The extension to $\tau = 5$ can be dealt with in an identical way as in (8.101)–(8.102). This proves the theorem. □

8.6 Infinite temperature behavior for $\tau \in (2,3]$

We start with proving the upper bound:

Proposition 8.9 (Upper bound for $\tau \in (2,3]$). *For $\tau \in (2,3)$,*

$$M(\beta, 0^+) \leq \mathbb{E}[D](C_{1,\tau} + C_{0,\tau})^{1/(3-\tau)} \beta^{1/(3-\tau)}, \tag{8.105}$$

while for $\tau = 3$,

$$M(\beta, 0^+) \leq \mathbb{E}[D] e^{-1/((C_{1,3}+C_{0,3})\beta)}. \tag{8.106}$$

Proof. We bound the first moment of $\xi(h)$ by splitting the analysis for K small and K large:

$$\begin{aligned} \mathbb{E}[\xi(h)] &= \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i) \right) \right] \qquad (8.107) \\ &= \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i) \right) \mathbb{1}_{\{K \leq 1/\mathbb{E}[\xi(h)]\}} \right] + \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i) \right) \mathbb{1}_{\{K > 1/\mathbb{E}[\xi(h)]\}} \right]. \end{aligned}$$

For the small K term, we use the upper bound of Lemma 7.4:

$$\mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i) \right) \mathbb{1}_{\{K \leq 1/\mathbb{E}[\xi(h)]\}} \right] \leq \hat{\beta} B + \hat{\beta} \mathbb{E}[K \mathbb{1}_{\{K \leq 1/\mathbb{E}[\xi(h)]\}}] \mathbb{E}[\xi(h)] \tag{8.108}$$

and for the large K term we use $\xi(h) \leq \beta$ a.s.:

$$\mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i) \right) \mathbb{1}_{\{K > 1/\mathbb{E}[\xi(h)]\}} \right] \leq \beta \mathbb{E}[\mathbb{1}_{\{K > 1/\mathbb{E}[\xi(h)]\}}]. \tag{8.109}$$

Combining the above two equations and applying Lemma 2.7 gives, for $\tau \in (2,3)$,

$$\mathbb{E}[\xi(h)] \leq \hat{\beta} B + \beta(C_{1,\tau} + C_{0,\tau})\mathbb{E}[\xi(h)]^{\tau-2}, \tag{8.110}$$

where we also used that $\hat{\beta} \leq \beta$. Taking the limit $B \searrow 0$ and dividing by $\mathbb{E}[\xi(h_0)]$ then gives

$$\mathbb{E}[\xi(h_0)] \leq (C_{1,\tau} + C_{0,\tau})^{1/(3-\tau)} \beta^{1/(3-\tau)}. \tag{8.111}$$

Observing that

$$\lim_{B\searrow 0} M(\beta,B) = \lim_{B\searrow 0} \mathbb{E}\left[\tanh\left(B + \sum_{i=1}^{D} \xi(h_i)\right)\right] \leq \lim_{B\searrow 0} B + \mathbb{E}[D]\mathbb{E}[\xi(h)] = \mathbb{E}[D]\mathbb{E}[\xi(h_0)], \tag{8.112}$$

now gives the desired result for $\tau \in (2,3)$.

Applying Lemma 2.7 for $\tau = 3$ gives

$$\begin{aligned}\mathbb{E}[\xi(h)] &\leq \hat{\beta} B + \hat{\beta} C_{1,3}\mathbb{E}[\xi(h)] \log(1/\mathbb{E}[\xi(h)]) + \beta C_{0,3}\mathbb{E}[\xi(h)] \\ &\leq \hat{\beta} B + \beta(C_{1,3} + C_{0,3})\mathbb{E}[\xi(h)] \log(1/\mathbb{E}[\xi(h)]).\end{aligned} \tag{8.113}$$

Again taking the limit $B \searrow 0$ and dividing by $\mathbb{E}[\xi(h_0)]$ gives

$$\mathbb{E}[\xi(h_0)] \leq e^{-1/((C_{1,3}+C_{0,3})\beta)}, \tag{8.114}$$

which gives the desired result after plugging this into (8.112). □

It remains to compute the lower bound:

Proposition 8.10 (Lower bound for $\tau \in (2,3]$)**.** *For $\tau \in (2,3)$,*

$$M(\beta, 0^+) \geq c_0 \beta^{1/(3-\tau)}, \tag{8.115}$$

for some constant $0 < c_0 < \infty$, while for $\tau = 3$,

$$M(\beta, 0^+) \geq c_0 e^{-1/(c_e \beta)}, \tag{8.116}$$

for some constants $0 < c_0, c_e < \infty$.

Proof. We can bound

$$\mathbb{E}[\xi(h)] = \mathbb{E}\left[\xi\left(B + \sum_{i=1}^{K} \xi(h_i)\right)\right] \geq \mathbb{E}\left[\xi\left(\sum_{i=1}^{K} \xi(h_i)\right)\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}\right]. \tag{8.117}$$

We use a Taylor expansion up to the second order and use that $\xi''(\zeta) \geq -2\beta/(1-\hat{\beta}^2)$ to obtain

$$\begin{aligned}&\mathbb{E}\left[\xi\left(\sum_{i=1}^{K} \xi(h_i)\right)\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}\right] \\ &\quad\geq \mathbb{E}\left[\left(\hat{\beta}\sum_{i=1}^{K} \xi(h_i) - \frac{\hat{\beta}}{1-\hat{\beta}^2}\left(\sum_{i=1}^{K} \xi(h_i)\right)^2\right)\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}\right] \\ &\quad= \hat{\beta}\mathbb{E}[K\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)] - \frac{\hat{\beta}}{1-\hat{\beta}^2}\mathbb{E}[K(K-1)\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)]^2 \\ &\qquad\quad - \mathbb{E}[K\mathbb{1}_{\{K \leq \varepsilon/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)^2].\end{aligned} \tag{8.118}$$

For a lower bound on the truncated first moment of K note that

$$\mathbb{E}[K\mathbb{1}_{\{K\le \ell\}}] = \sum_{k=1}^{\lfloor \ell \rfloor} k\rho_k = \sum_{k=1}^{\lfloor \ell \rfloor}\sum_{i=k}^{\lfloor \ell \rfloor} \rho_k = \sum_{k=1}^{\lfloor \ell \rfloor}[\rho_{\ge k} - \rho_{>\ell}] \ge \sum_{k=1}^{\lfloor \delta\ell \rfloor}[\rho_{\ge k} - \rho_{\ge \ell}], \tag{8.119}$$

for some $0<\delta\le 1$. From Lemma 2.6 we know that

$$c_\rho k^{-(\tau-2)} \le \rho_{\ge k} \le C_\rho k^{-(\tau-2)}. \tag{8.120}$$

Hence, for $\tau\in(2,3)$,

$$\begin{aligned}\mathbb{E}[K\mathbb{1}_{\{K\le \ell\}}] &\ge c_\rho \sum_{k=1}^{\lfloor \delta\ell \rfloor} k^{-(\tau-2)} - \delta C_\rho \ell^{3-\tau} \ge c_\rho \int_1^{\delta\ell} k^{-(\tau-2)}\mathrm{d}k - \delta C_\rho \ell^{3-\tau}\\ &= \frac{c_\rho}{3-\tau}\left((\delta\ell)^{3-\tau}-1\right) - \delta C_\rho \ell^{3-\tau},\end{aligned} \tag{8.121}$$

which gives, by choosing δ small enough,

$$\mathbb{E}[K\mathbb{1}_{\{K\le \ell\}}] \ge c_{1,\tau}\ell^{3-\tau}, \tag{8.122}$$

for some constant $c_{1,\tau}>0$.

The other truncated moments in (8.118) can be bounded using Lemma 2.7. Furthermore, we use that $\xi(h)\le\beta$ a.s., so that $\mathbb{E}[\xi(h)^2]\le\beta\mathbb{E}[\xi(h)]$.

Using all these bounds on (8.118) gives, for $\tau\in(2,3)$,

$$\begin{aligned}\mathbb{E}[\xi(h)] &\ge c_{1,\tau}\varepsilon^{3-\tau}\hat\beta\mathbb{E}[\xi(h)]^{\tau-2} - \frac{C_{2,\tau}\varepsilon^{4-\tau}}{1-\hat\beta^2}\hat\beta\mathbb{E}[\xi(h)]^{\tau-2} - \frac{\beta\varepsilon^{3-\tau}}{1-\hat\beta^2}\hat\beta\mathbb{E}[\xi(h)]^{\tau-2}\\ &= \left(c_{1,\tau} - \frac{C_{2,\tau}\varepsilon}{1-\hat\beta^2} - \frac{\beta}{1-\hat\beta^2}\right)\varepsilon^{3-\tau}\hat\beta\mathbb{E}[\xi(h)]^{\tau-2} \ge c_1\beta\mathbb{E}[\xi(h)]^{\tau-2},\end{aligned} \tag{8.123}$$

for some $c_1>0$ if ε and β are small enough. Taking the limit $B\searrow 0$ and dividing by $\mathbb{E}[\xi(h_0)]$ then gives

$$\mathbb{E}[\xi(h_0)] \ge c_1^{3-\tau}\beta^{1/(3-\tau)}. \tag{8.124}$$

To show that the magnetization has the same behavior, we truncate and use that $\tanh x\ge x-x^2$,

$$\begin{aligned}M(\beta,B) &= \mathbb{E}\left[\tanh\left(B+\sum_{i=1}^{D}\xi(h_i)\right)\right] \ge \mathbb{E}\left[\tanh\left(\sum_{i=1}^{D}\xi(h_i)\right)\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}\right]\\ &\ge \mathbb{E}\left[\left(\sum_{i=1}^{D}\xi(h_i) - \left(\sum_{i=1}^{D}\xi(h_i)\right)^2\right)\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}\right]\\ &= \mathbb{E}[D\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)] - \mathbb{E}[D(D-1)\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)]^2\\ &\qquad - \mathbb{E}[D\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)^2].\end{aligned} \tag{8.125}$$

Using Lemma 2.7 and $\xi(h)\le\beta$, we get

$$\begin{aligned}M(\beta,B) &\ge \mathbb{E}[D\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)] - C_{2,\tau}\mathbb{E}[\xi(h)]^{\tau-1} - \mathbb{E}[D\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}]\mathbb{E}[\xi(h)^2]\\ &\ge \left((1-\beta)\mathbb{E}[D\mathbb{1}_{\{D\le 1/\mathbb{E}[\xi(h)]\}}] - C_{2,\tau}\beta^{\tau-2}\right)\mathbb{E}[\xi(h)].\end{aligned} \tag{8.126}$$

Taking the limit $B \searrow 0$ and plugging in (8.124) gives the result.

For $\tau = 3$,

$$\begin{aligned}\mathbb{E}[K\mathbb{1}_{\{K\leq\ell\}}] &= \sum_{k=1}^{\lfloor \ell \rfloor}[\rho_{\geq k} - \rho_{>\ell}] \geq c_\rho \sum_{k=1}^{\lfloor \ell \rfloor} k^{-(\tau-2)} - C_\rho \\ &\geq c_\rho \int_1^\ell k^{-1}\mathrm{d}k - C_\rho = c_\rho \log \ell - C_\rho \geq c_{1,3} \log \ell, \end{aligned} \tag{8.127}$$

for ℓ large enough and some $c_{1,3} > 0$. The rest of the analysis is similar to that of the case where $\tau \in (2,3)$. □

9

CRITICAL BEHAVIOR OF THE SUSCEPTIBILITY

We now turn to the critical behavior of the susceptibility. It is expected that the susceptibility $\chi(\beta, 0^+)$ blows up as $\beta \to \beta_c$. We show in this chapter that this is indeed the case, proving that the phase transition is of *second order*, because the second derivative of the pressure with respect to B is non-analytic in $(\beta_c, 0^+)$. We also quantify how fast the convergence to infinity is in the high-temperature regime, i.e., for $\beta \nearrow \beta_c$, by identifying the critical exponent γ. For the low temperature limit $\beta \searrow \beta_c$ we are not able to compute the critical exponent γ', but we do give a lower bound and present a heuristic argument why the corresponding upper bound should also hold.

9.1 Results

For the susceptibility in the *subcritical* phase, i.e., in the high-temperature region $\beta < \beta_c$, we can not only identify the critical exponent γ, but we can also identify the constant:

Theorem 9.1 (Critical exponent γ). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P that obeys $\mathbb{E}[K] < \infty$. Then, for $\beta < \beta_c$,*

$$\chi(\beta, 0^+) = 1 + \frac{\mathbb{E}[D]\hat{\beta}}{1 - \nu\hat{\beta}}. \tag{9.1}$$

In particular,

$$\lim_{\beta \nearrow \beta_c} \chi(\beta, 0^+)(\beta_c - \beta) = \frac{\mathbb{E}[D]\hat{\beta}_c}{1 - \hat{\beta}_c^2}, \tag{9.2}$$

and hence

$$\gamma = 1. \tag{9.3}$$

For the supercritical susceptibility, we prove the following lower bound on γ':

Proposition 9.2 (Critical exponent γ'). *Assume that the random graph sequence $(G_N)_{N\geq 1}$ is locally tree-like with asymptotic degree distribution P that obeys $\mathbb{E}[K^3] < \infty$ or a power law with exponent $\tau \in (3, 5]$. Then,*

$$\gamma' \geq 1. \tag{9.4}$$

9.2 Discussion

Mean-field values. The above result shows that the critical exponent γ has the same value as in the Curie-Weiss model [16] for all locally tree-like random graphs with finite second moment of the degree distribution. Again, these values are also the same for the Ising model on $\mathbb{Z}^d$ for $d > 4$ [2] with possible logarithmic corrections at $d = 4$ [4]. The mean-field value for γ' is also equal to 1 [16], but we are not able to prove this for our model.

The critical exponents γ' and other critical exponents. Proposition 9.2 only gives a lower bound on the critical exponent γ'. It is predicted that $\gamma' = 1$ for all $\tau > 3$, while there are also predictions for other critical exponents. For instance the critical exponent α' for the specific heat in the low-temperature phase satisfies $\alpha' = 0$ when $\mathbb{E}[K^3] < \infty$ and $\alpha' = (\tau - 5)/(\tau - 3)$ in the power-law case with $\tau \in (3, 5)$ (see [40, 74]). We prove the lower bound $\gamma' \geq 1$ in Section 9.4 below, and we also present a heuristic argument that $\gamma' \leq 1$ holds. The critical exponent α' for the specific heat is beyond our current methods, partly since we are not able to relate the specific heat on a random graph to that on the random Bethe tree.

9.3 The critical exponent γ

The proof is divided into three steps. We first reduce the susceptibility on the random graph to the one on the random Bethe tree. Secondly, we rewrite the susceptibility on the tree using transfer matrix techniques. Finally, we use this rewrite (which applies to *all* β and $B > 0$) to prove that $\gamma = 1$.

Reduction to the random tree. Let ϕ denote a vertex selected uniformly at random from $[N]$ and let $\mathbb{E}_\phi$ denote its expectation. Then we can write the susceptibility as

$$\chi_N \equiv \frac{1}{N} \sum_{i,j=1}^{N} \Big(\langle \sigma_i \sigma_j \rangle_{\mu_N} - \langle \sigma_i \rangle_{\mu_N} \langle \sigma_j \rangle_{\mu_N} \Big) = \mathbb{E}_\phi \left[\sum_{j=1}^{N} \Big(\langle \sigma_\phi \sigma_j \rangle_{\mu_N} - \langle \sigma_\phi \rangle_{\mu_N} \langle \sigma_j \rangle_{\mu_N} \Big) \right]. \tag{9.5}$$

Note that

$$\langle \sigma_i \sigma_j \rangle_{\mu_N} - \langle \sigma_i \rangle_{\mu_N} \langle \sigma_j \rangle_{\mu_N} = \frac{\partial \langle \sigma_i \rangle_{\mu_N}}{\partial B_j}, \tag{9.6}$$

which is, by the GHS inequality [58], decreasing in external fields at all other vertices $k \in [N]$. Denote by $\langle \cdot \rangle^{t,+/f}$ the Ising model with $+$/free boundary conditions, respectively, at all vertices at graph distance t from ϕ. Then, for all $t \geq 1$,

$$\chi_N \geq \mathbb{E}_\phi \left[\sum_{j=1}^{N} \Big(\langle \sigma_\phi \sigma_j \rangle_{\mu_N}^{t,+} - \langle \sigma_\phi \rangle_{\mu_N}^{t,+} \langle \sigma_j \rangle_{\mu_N}^{t,+} \Big) \right]. \tag{9.7}$$

By introducing boundary conditions, only vertices in the ball $B_\phi(t)$ contribute to the sum. Hence, by taking the limit $N \to \infty$ and using that the graph is locally tree-like,

$$\chi \geq \mathbb{E}\left[\sum_{j \in T_t} \Big(\langle \sigma_\phi \sigma_j \rangle^{t,+} - \langle \sigma_\phi \rangle^{t,+} \langle \sigma_j \rangle^{t,+} \Big) \right], \tag{9.8}$$

where the expectation now is over the random tree $T_t \sim \mathscr{T}(D,K,t)$ with root ϕ.

For an upper bound on χ_N we use a trick similar to one used in the proof of [34, Corollary 4.5]: Let $B'_j = B$ if $j \in B_t(\phi)$ and $B'_j = B + H$ if $j \notin B_t(\phi)$ for some $H > -B$. Denote by $\langle\cdot\rangle_H$ the associated Ising expectation. Then, because of (9.6),

$$\mathbb{E}_\phi\left[\sum_{j\notin B_t(\phi)}\Big(\langle\sigma_\phi\sigma_j\rangle - \langle\sigma_\phi\rangle\langle\sigma_j\rangle\Big)\right] = \mathbb{E}_\phi\left[\frac{\partial}{\partial H}\langle\sigma_\phi\rangle_H\bigg|_{H=0}\right]. \tag{9.9}$$

By the GHS inequality, $\langle\sigma_\phi\rangle_H$ is a concave function of H and hence,

$$\mathbb{E}_\phi\left[\frac{\partial}{\partial H}\langle\sigma_\phi\rangle_H\bigg|_{H=0}\right] \le \mathbb{E}_\phi\left[\frac{2}{B}\Big(\langle\sigma_\phi\rangle_{H=0} - \langle\sigma_\phi\rangle_{H=-B/2}\Big)\right]. \tag{9.10}$$

Using the GKS inequality this can be bounded from above by

$$\mathbb{E}_\phi\left[\frac{2}{B}\Big(\langle\sigma_\phi\rangle^{t,+}_{H=0} - \langle\sigma_\phi\rangle^{t,f}_{H=-B/2}\Big)\right] = \mathbb{E}_\phi\left[\frac{2}{B}\Big(\langle\sigma_\phi\rangle^{t,+} - \langle\sigma_\phi\rangle^{t,f}\Big)\right], \tag{9.11}$$

where the equality holds because the terms depend only on the system in the ball $B_t(\phi)$ and hence not on H. By letting $N\to\infty$, by the locally tree-likeness, this is equal to

$$\frac{2}{B}\mathbb{E}\left[\Big(\langle\sigma_\phi\rangle^{t,+} - \langle\sigma_\phi\rangle^{t,f}\Big)\right], \tag{9.12}$$

where the expectation and the Ising model now is over the random tree $T_t \sim \mathscr{T}(D,K,t)$ with root ϕ. From Lemma 4.3 we know that this expectation can be bounded from above by A/t for some constant $A = A(\beta,B) < \infty$. Hence, if $t\to\infty$,

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}\left[\sum_{j\in T_t}\Big(\langle\sigma_\phi\sigma_j\rangle^{t,+} - \langle\sigma_\phi\rangle^{t,+}\langle\sigma_j\rangle^{t,+}\Big)\right] &\le \chi \\ &\le \lim_{t\to\infty}\mathbb{E}\left[\sum_{j\in T_t}\Big(\langle\sigma_\phi\sigma_j\rangle^{t,f} - \langle\sigma_\phi\rangle^{t,f}\langle\sigma_j\rangle^{t,f}\Big)\right].\end{aligned} \tag{9.13}$$

Rewrite of the susceptibility on trees. It remains to study the susceptibility on trees. For this, condition on the tree T_∞. Then, for some vertex j at height $\ell \le t$ in the tree, denote the vertices on the unique path from ϕ to j by $\phi = v_0, v_1, \ldots, v_\ell = j$ and let, for $0 \le i \le \ell$, $S_{\le i} = (\sigma_{v_0},\ldots,\sigma_{v_i})$. We first compute the expected value of a spin σ_{v_i} on this path, conditioned on the spin values $S_{\le i-1}$. Note that under this conditioning the expected spin value only depends on the spin value $\sigma_{v_{i-1}}$ and the effective field $h_{v_i} = h_{v_i}^{t,+/f}$ obtained by pruning the tree at vertex v_i, i.e., by removing all edges at vertex v_i going away from the root and replacing the external magnetic field at vertex v_i by h_{v_i} which can be exactly computed using Lemma 4.2. Hence,

$$\langle\sigma_{v_i}|S_{\le i-1}\rangle^{t,+/f} = \frac{e^{\beta\sigma_{v_{i-1}}+h_{v_i}} - e^{-\beta\sigma_{v_{i-1}}-h_{v_i}}}{e^{\beta\sigma_{v_{i-1}}+h_{v_i}} + e^{-\beta\sigma_{v_{i-1}}-h_{v_i}}}. \tag{9.14}$$

We can write the indicators $\mathbb{1}_{\{\sigma_{v_{i-1}}=\pm 1\}} = \frac{1}{2}(1 \pm \sigma_{v_{i-1}})$, so that the above equals

$$\begin{aligned} &\frac{1}{2}(1+\sigma_{v_{i-1}})\frac{e^{\beta+h_{v_i}} - e^{-\beta-h_{v_i}}}{e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}}} + \frac{1}{2}(1-\sigma_{v_{i-1}})\frac{e^{-\beta+h_{v_i}} - e^{\beta-h_{v_i}}}{e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}}} \\ &\qquad = \sigma_{v_{i-1}}\frac{1}{2}\left(\frac{e^{\beta+h_{v_i}} - e^{-\beta-h_{v_i}}}{e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}}} - \frac{e^{-\beta+h_{v_i}} - e^{\beta-h_{v_i}}}{e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}}}\right) \\ &\qquad\qquad + \frac{1}{2}\left(\frac{e^{\beta+h_{v_i}} - e^{-\beta-h_{v_i}}}{e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}}} + \frac{e^{-\beta+h_{v_i}} - e^{\beta-h_{v_i}}}{e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}}}\right). \end{aligned} \tag{9.15}$$

By pairwise combining the terms over a common denominator the above equals

$$\begin{aligned} &\sigma_{v_{i-1}}\frac{1}{2}\frac{(e^{\beta+h_{v_i}} - e^{-\beta-h_{v_i}})(e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}}) - (e^{-\beta+h_{v_i}} - e^{\beta-h_{v_i}})(e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}})}{(e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}})(e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}})} \\ &\quad + \frac{1}{2}\frac{(e^{\beta+h_{v_i}} - e^{-\beta-h_{v_i}})(e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}}) + (e^{-\beta+h_{v_i}} - e^{\beta-h_{v_i}})(e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}})}{(e^{\beta+h_{v_i}} + e^{-\beta-h_{v_i}})(e^{-\beta+h_{v_i}} + e^{\beta-h_{v_i}})}. \end{aligned} \tag{9.16}$$

By expanding all products, this equals, after cancellations,

$$\begin{aligned} &\sigma_{v_{i-1}}\frac{e^{2\beta} + e^{-2\beta}}{e^{2\beta} + e^{-2\beta} + e^{2h_{v_i}} + e^{-2h_{v_i}}} + \frac{e^{2h_{v_i}} + e^{-2h_{v_i}}}{e^{2\beta} + e^{-2\beta} + e^{2h_{v_i}} + e^{-2h_{v_i}}} \\ &\qquad = \sigma_{v_{i-1}}\frac{\sinh(2\beta)}{\cosh(2\beta) + \cosh(2h_{v_i})} + \frac{\sinh(2h_{v_i})}{\cosh(2\beta) + \cosh(2h_{v_i})}. \end{aligned} \tag{9.17}$$

Using this, we have that

$$\begin{aligned} \langle \sigma_{v_\ell} \rangle^{t,+/f} &= \langle \langle \sigma_{v_\ell} | S_{\leq \ell-1} \rangle^{t,+/f} \rangle^{t,+/f} \\ &= \langle \sigma_{v_{\ell-1}} \rangle^{t,+/f} \frac{\sinh(2\beta)}{\cosh(2\beta) + \cosh(2h_{v_\ell})} + \frac{\sinh(2h_{v_\ell})}{\cosh(2\beta) + \cosh(2h_{v_\ell})}. \end{aligned} \tag{9.18}$$

Applying this recursively, we get

$$\begin{aligned} \langle \sigma_{v_\ell} \rangle^{t,+/f} = \langle \sigma_{v_0} \rangle^{t,+/f} &\prod_{i=1}^{\ell} \frac{\sinh(2\beta)}{\cosh(2\beta) + \cosh(2h_{v_i})} \\ &+ \sum_{i=1}^{\ell}\left(\frac{\sinh(2h_{v_i})}{\cosh(2\beta) + \cosh(2h_{v_i})} \prod_{k=i+1}^{\ell} \frac{\sinh(2\beta)}{\cosh(2\beta) + \cosh(2h_{v_k})}\right). \end{aligned} \tag{9.19}$$

Similarly,

$$\begin{aligned}
\langle \sigma_{v_0}\sigma_{v_\ell}\rangle^{t,+/f} &= \Bigg\langle \sigma_{v_0}\Bigg(\sigma_{v_0}\prod_{i=1}^{\ell}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_i})} \\
&\quad + \sum_{i=1}^{\ell}\Bigg(\frac{\sinh(2h_{v_i})}{\cosh(2\beta)+\cosh(2h_{v_i})}\prod_{k=i+1}^{\ell}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_k})}\Bigg)\Bigg)\Bigg\rangle^{t,+/f} \\
&= \prod_{i=1}^{\ell}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_i})} \\
&\quad + \langle \sigma_{v_0}\rangle^{t,+/f}\sum_{i=1}^{\ell}\Bigg(\frac{\sinh(2h_{v_i})}{\cosh(2\beta)+\cosh(2h_{v_i})}\prod_{k=i+1}^{\ell}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_k})}\Bigg).
\end{aligned} \tag{9.20}$$

Combining the above yields

$$\langle \sigma_{v_0}\sigma_{v_\ell}\rangle^{t,+/f} - \langle \sigma_{v_0}\rangle^{t,+/f}\langle \sigma_{v_\ell}\rangle^{t,+/f} = \Big(1-\big(\langle \sigma_{v_0}\rangle^{t,+/f}\big)^2\Big)\prod_{i=1}^{\ell}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_i})}. \tag{9.21}$$

By taking the limit $t\to\infty$, we obtain

$$\chi = \mathbb{E}\Bigg[\sum_{j\in T_\infty}\big(1-\langle \sigma_{v_0}\rangle^2\big)\prod_{i=1}^{|j|}\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_i})}\Bigg]. \tag{9.22}$$

Finally, we can rewrite

$$\frac{\sinh(2\beta)}{\cosh(2\beta)+\cosh(2h_{v_i})} = \frac{2\sinh(\beta)\cosh(\beta)}{2\cosh(\beta)^2-1+\cosh(2h_{v_i})} = \frac{\hat\beta}{1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}}, \tag{9.23}$$

so that

$$\chi(\beta,B) = \mathbb{E}\Bigg[\big(1-\langle \sigma_{v_0}\rangle^2\big)\sum_{j\in T_\infty}\hat\beta^{|j|}\prod_{i=1}^{|j|}\Big(1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}\Big)^{-1}\Bigg]. \tag{9.24}$$

The rewrite in (9.24) is valid for all β and $B>0$, and provides the starting point for all our results on the susceptibility.

Identification of the susceptibility for $\beta<\beta_c$. We take the limit $B\searrow 0$, for $\beta<\beta_c$, and apply dominated convergence. First of all, all fields h_i converge to zero by the definition of β_c, so we have pointwise convergence. Secondly, $1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}\geq 1$, so that the random variable in the expectation is bounded from above by $\sum_{j\in T_\infty}\hat\beta^{|j|}$, which has finite expectation as we show below. Thus, by dominated convergence, the above converges to

$$\lim_{B\searrow 0}\chi(\beta,B) = \mathbb{E}\Bigg[\sum_{j\in T_\infty}\hat\beta^{|j|}\Bigg]. \tag{9.25}$$

Denote by Z_ℓ the number of vertices at distance ℓ from the root. Then,

$$\mathbb{E}\Big[\sum_{j\in T_\infty}\hat{\beta}^{|j|}\Big]=\mathbb{E}\Big[\sum_{\ell=0}^{\infty}Z_\ell\hat{\beta}^{\ell}\Big]=\sum_{\ell=0}^{\infty}\mathbb{E}[Z_\ell]\hat{\beta}^{\ell}, \tag{9.26}$$

because $Z_\ell \geq 0$, a.s. Note that $Z_\ell/(\mathbb{E}[D]\nu^{\ell-1})$ is a martingale, because the offspring of the root has expectation $\mathbb{E}[D]$ and all other vertices have expected offspring ν. Hence,

$$\lim_{B\searrow 0}\chi(\beta,B)=\sum_{\ell=0}^{\infty}\mathbb{E}[Z_\ell]\hat{\beta}^{\ell}=1+\sum_{\ell=1}^{\infty}\mathbb{E}[D]\nu^{\ell-1}\hat{\beta}^{\ell}=1+\frac{\mathbb{E}[D]\hat{\beta}}{1-\hat{\beta}\nu}. \tag{9.27}$$

This proves (9.1). We continue to prove (9.2), which follows by using (8.86) and (8.99):

$$\frac{\mathbb{E}[D]\hat{\beta}}{1-\hat{\beta}^2}(\beta_c-\beta)^{-1}+1\leq\frac{\mathbb{E}[D]}{\nu}\frac{1}{1-\hat{\beta}\nu}\leq\frac{\mathbb{E}[D]\hat{\beta}}{1-\hat{\beta}_c^2}(\beta_c-\beta)^{-1}+1. \tag{9.28}$$

9.4 Partial results for the critical exponent γ'

Proof of Proposition 9.2. We start by rewriting the susceptibility in a form that is convenient in the low-temperature phase.

A rewrite of the susceptibility in terms of i.i.d. random variables. For $\beta > \beta_c$ we start from (9.24). We further rewrite

$$\chi(\beta,B)=\sum_{\ell=0}^{\infty}\hat{\beta}^{\ell}\mathbb{E}\Big[(1-\langle\sigma_{v_0}\rangle^2)\sum_{v_\ell\in T_\infty}\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.29}$$

Here, and in the sequel, we use the convention that empty products, arising when $\ell=0$, equal 1, while empty sums equal 0. Thus, the contribution due to $\ell=0$ in the above sum equals 1. We write $v_0=\phi$ and $v_i=a_0\cdots a_{i-1}\in\mathbb{N}^i$ for $i\geq 1$, so that v_i is the a_{i-1}st child of v_{i-1}. Then,

$$\chi(\beta,B)=\sum_{\ell=0}^{\infty}\hat{\beta}^{\ell}\sum_{a_0,\ldots,a_{\ell-1}}\mathbb{E}\Big[(1-\langle\sigma_{v_0}\rangle^2)\mathbb{1}_{\{v_\ell\in T_\infty\}}\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.30}$$

Let K_{v_i} be the number of children of v_i, and condition on $K_{v_i}=k_i$ for every $i\in[0,\ell-1]$, where we abuse notation to write $[0,m]=\{0,\ldots,m\}$. As a result, we obtain that

$$\begin{aligned}\chi(\beta,B)=\sum_{\ell=0}^{\infty}\hat{\beta}^{\ell}\sum_{a_0,\ldots,a_{\ell-1}}\sum_{k_0,\ldots,k_{\ell-1}}&\mathbb{P}(v_\ell\in T_\infty,K_{v_i}=k_i\ \forall i\in[0,\ell-1])\\ &\times\mathbb{E}\Big[(1-\langle\sigma_{v_0}\rangle^2)\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_{v_i})-1}{2\cosh(\beta)^2}\Big)\Big\}\\ &\qquad\mid v_\ell\in T_\infty,K_{v_i}=k_i\ \forall i\in[0,\ell-1]\Big].\end{aligned} \tag{9.31}$$

Note that

$$\mathbb{P}(K_{v_i} = k_i \ \forall i \in [0,\ell-1], v_\ell \in T_\infty) = \mathbb{P}(D = k_0)\mathbb{1}_{\{a_0 \le k_0\}} \prod_{i=1}^{\ell-1} \mathbb{P}(K = k_i)\mathbb{1}_{\{a_i \le k_i\}}. \tag{9.32}$$

Let $T_{i,j}$ be the tree that descibes all descendants of the jth child of v_i, with the a_ith child removed, and T_ℓ the offspring of v_ℓ. When $v_\ell \in T_\infty$, all information of the tree T_∞ can be encoded in the collection of trees $(T_{i,j})_{j\in[0,K_{v_i}-1], i\in[0,\ell-1]}$ and T_ℓ, together with the sequence $(a_i)_{i=0}^{\ell-1}$. Denote $\vec{T} = \big((T_{i,j})_{j\in[0,K_{v_i}-1], i\in[0,\ell-1]}, T_\ell\big)$. Then, for any collection of trees $\vec{t} = \big((t_{i,j})_{j\in[0,k_i-1], i\in[0,\ell-1]}, t_\ell\big)$,

$$\mathbb{P}(\vec{T} = \vec{t} \mid K_{v_i} = k_i \ \forall i \in [0,\ell-1], v_\ell \in T_\infty) = \mathbb{P}(T = t_\ell) \prod_{(i,j)\in[0,k_i-1]\times[0,\ell-1]} \mathbb{P}(T = t_{i,j}), \tag{9.33}$$

where the law of T is that of a Galton-Watson tree with offspring distribution K. We conclude that

$$\chi(\beta,B) = \sum_{\ell=0}^{\infty} \hat\beta^\ell \sum_{a_0,\dots,a_{\ell-1}} \sum_{k_0,\dots,k_{\ell-1}} \mathbb{P}(D = k_0)\mathbb{1}_{\{a_0 \le k_0\}} \prod_{i=1}^{\ell-1} \mathbb{P}(K = k_i)\mathbb{1}_{\{a_i \le k_i\}}$$
$$\times \mathbb{E}\Big[(1 - \langle \sigma_{v_0} \rangle^2) \exp\Big\{ -\sum_{i=1}^{\ell} \log\Big(1 + \frac{\cosh(2h_i^\star(\vec{k})) - 1}{2\cosh(\beta)^2}\Big)\Big\}\Big], \tag{9.34}$$

where $(h_i^\star(\vec{k}))_{i=0}^{\ell}$ satisfy the recursion relations $h_\ell^\star = h_{\ell,1}$

$$h_i^\star(\vec{k}) = B + \xi(h_{i+1}^\star(\vec{k})) + \sum_{j=1}^{k_i-1} \xi(h_{i,j}), \tag{9.35}$$

and where $(h_{i,j})_{i\in[0,\ell],j\ge1}$ are i.i.d. copies of the random variable $h(\beta,B)$. We note that the law of $(h_i^\star(\vec{k}))_{i=0}^{\ell}$ does not depend on $(a_i)_{i\in[0,\ell-1]}$, so that the summation over $(a_i)_{i\in[0,\ell-1]}$ yields

$$\chi(\beta,B) = \sum_{\ell=0}^{\infty} \hat\beta^\ell \sum_{k_0,\dots,k_{\ell-1}} k_0\mathbb{P}(D = k_0) \prod_{i=1}^{\ell-1} k_i\mathbb{P}(K = k_i)$$
$$\times \mathbb{E}\Big[(1 - \langle \sigma_{v_0} \rangle^2) \exp\Big\{ -\sum_{i=1}^{\ell} \log\Big(1 + \frac{\cosh(2h_i^\star(\vec{k})) - 1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.36}$$

For a random variable X on the non-negative integers with $\mathbb{E}[X] > 0$, we let $X^\star$ be the size-biased distribution of X given by

$$\mathbb{P}(X^\star = k) = \frac{k}{\mathbb{E}[X]}\mathbb{P}(X = k). \tag{9.37}$$

Then

$$\chi(\beta,B) = \frac{\mathbb{E}[D]}{\nu} \sum_{\ell=0}^{\infty} (\hat\beta\nu)^\ell \sum_{k_0,\dots,k_{\ell-1}} \mathbb{P}(D^\star = k_0) \prod_{i=1}^{\ell-1} \mathbb{P}(K^\star = k_i)$$
$$\times \mathbb{E}\Big[(1 - \langle \sigma_{v_0} \rangle^2) \exp\Big\{ -\sum_{i=1}^{\ell} \log\Big(1 + \frac{\cosh(2h_i^\star(\vec{k})) - 1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.38}$$

Define $(h_i^\star)_{i=0}^{\ell} = \big(h_i^\star(D^\star, K_1^\star, \ldots, K_{\ell-1}^\star, K_\ell)\big)_{i=0}^{\ell}$, where the random variables $(D^\star, K_1^\star, \ldots, K_{\ell-1}^\star, K_\ell)$ are independent. Then we finally arrive at

$$\chi(\beta,B) = \frac{\mathbb{E}[D]}{\nu}\sum_{\ell=0}^{\infty}(\hat\beta\nu)^{\ell}\mathbb{E}\Big[(1-\langle\sigma_{v_0}\rangle^2)\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_i^\star)-1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.39}$$

Reduction to second moments. We now proceed towards the lower bound on γ'. Note that, a.s.,

$$\langle\sigma_{v_0}\rangle = \tanh(h_{v_0}^\star), \tag{9.40}$$

where

$$h_{v_0^\star} = B + \xi(h_{v_1}^\star) + \sum_{j=1}^{D^\star-1}\xi(h_{0,j}) \le B + \beta + \sum_{j=1}^{D^\star-1}\xi(h_{0,j}). \tag{9.41}$$

Therefore,

$$\langle\sigma_{v_0}\rangle \le \tanh\Big(B+\beta+\sum_{j=1}^{D^\star-1}\xi(h_{0,j})\Big). \tag{9.42}$$

The right hand side is independent of $(h_i^\star)_{i=1}^{\ell}$, so that the expectation factorizes. Further,

$$\mathbb{E}\Big[\tanh\Big(B+\beta+\sum_{j=1}^{D^\star-1}\xi(h_{0,j})\Big)\Big] \to \tanh(\beta) = \hat\beta < 1, \tag{9.43}$$

as $B \searrow 0, \beta \searrow \beta_c$.

Further, we restrict the sum over all ℓ to $\ell \le m$, where we take $m = (\beta-\beta_c)^{-1}$. This leads to

$$\chi(\beta,B) \ge \frac{(1-\hat\beta^2)\mathbb{E}[D]}{\nu}\sum_{\ell=0}^{m}(\hat\beta\nu)^{\ell}\mathbb{E}\Big[\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_i^\star)-1}{2\cosh(\beta)^2}\Big)\Big\}\Big]. \tag{9.44}$$

We condition on all coordinates of $(D^\star, K_1^\star, \ldots, K_{\ell-1}^\star, K_\ell)$ being at most $b = (\beta-\beta_c)^{-1/(\tau-3)}$, which has probability

$$\begin{aligned}\mathbb{P}(D^\star \le b, K_1^\star \le b, \ldots, K_{\ell-1}^\star \le b, K_\ell \le b) &\ge (1-o(1))\mathbb{P}(K^\star \le b)^m \\ &\ge (1-o(1))\big(1-C_{K^\star}b^{-(\tau-3)}\big)^m,\end{aligned} \tag{9.45}$$

which is uniformly bounded from below by a constant for the choices $m = (\beta-\beta_c)^{-1}$ and $b = (\beta-\beta_c)^{-1/(\tau-3)}$. Also, we use that $\hat\beta\nu \ge 1$, since $\beta > \beta_c$. This leads us to

$$\chi(\beta,B) \ge c_\chi\sum_{\ell=0}^{m}\overline{\mathbb{E}}_b\Big[\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_i^\star)-1}{2\cosh(\beta)^2}\Big)\Big\}\Big], \tag{9.46}$$

where $\overline{\mathbb{E}}_b$ denotes the conditional expectation given that $D^\star \le b, K_1^\star \le b, \ldots, K_{\ell-1}^\star \le b, K_\ell \le b$. Using that $\mathbb{E}[e^X] \ge e^{\mathbb{E}[X]}$, which follows from Jensen's inequality, this leads us to

$$\chi(\beta,B) \ge c_\chi\sum_{\ell=0}^{m}\exp\Big\{-\sum_{i=1}^{\ell}\overline{\mathbb{E}}_b\Big[\log\Big(1+\frac{\cosh(2h_i^\star)-1}{2\cosh(\beta)^2}\Big)\Big]\Big\}. \tag{9.47}$$

Define, for $a > 0$ and $x \geq 0$, the function $q(x) = \log\big(1+a(\cosh(x)-1)\big)$. Differentiating leads to

$$q'(x) = \frac{a\sinh(x)}{1+a(\cosh(x)-1)}, \tag{9.48}$$

so that $q'(x) \leq C_q x/2$ for some constant C_q and all $x \geq 0$. As a result, $q(x) \leq C_q x^2/4$, so that

$$\chi(\beta,B) \geq c_\chi \sum_{\ell=0}^{m} \exp\Big\{-C_q \sum_{i=1}^{\ell} \overline{\mathbb{E}}_b\left[(h_i^\star)^2\right]\Big\}. \tag{9.49}$$

Second-moment analysis of $h_i^\star$. As a result, it suffices to investigate second moments of $h_i^\star$, which we proceed with now. We note that

$$h_i^\star = \xi(h_{i+1}^\star) + B + \sum_{j=1}^{K_i^\star-1} \xi(h_{i,j}). \tag{9.50}$$

Taking expectations and using that $\xi(h) \leq \hat{\beta} h$ leads to

$$\overline{\mathbb{E}}_b\left[h_i^\star\right] \leq \hat{\beta}\overline{\mathbb{E}}_b\left[h_{i+1}^\star\right] + B + \mathbb{E}[K^\star - 1 \mid K^\star \leq b]\mathbb{E}[\xi(h)]. \tag{9.51}$$

Iterating this inequality until $\ell - i$ and using that $\overline{\mathbb{E}}_b\left[h_\ell^\star\right] \leq B + \nu\mathbb{E}[\xi(h)]$ (since $\overline{\mathbb{E}}_b[K] \leq \mathbb{E}[K]$) leads to

$$\begin{aligned}\overline{\mathbb{E}}_b\left[h_i^\star\right] &\leq \hat{\beta}^{\ell-i}(B+\nu\mathbb{E}[\xi(h)]) + \sum_{s=0}^{\ell-i-1}\hat{\beta}^s\big(B + \mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)]\big)\\ &\leq \hat{\beta}^{\ell-i}(B+\nu\mathbb{E}[\xi(h)]) + \frac{B+\mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)]}{1-\hat{\beta}}.\end{aligned} \tag{9.52}$$

Similarly,

$$\begin{aligned}\overline{\mathbb{E}}_b\left[(h_i^\star)^2\right] \leq{}& \hat{\beta}^2\overline{\mathbb{E}}_b\left[(h_{i+1}^\star)^2\right] + 2\hat{\beta}\overline{\mathbb{E}}_b\left[h_{i+1}^\star\right]\big(B+\mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)]\big)\\ &+ B^2 + 2B\mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)]\\ &+ \mathbb{E}[(K^\star-1)(K^\star-2)\mid K^\star\leq b]\mathbb{E}[\xi(h)]^2\\ &+ \mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)^2].\end{aligned} \tag{9.53}$$

Taking the limit $B \searrow 0$ we thus obtain

$$\begin{aligned}\overline{\mathbb{E}}_b\left[(h_i^\star)^2\right] \leq{}& \hat{\beta}^2\overline{\mathbb{E}}_b\left[(h_{i+1}^\star)^2\right] + 2\hat{\beta}\overline{\mathbb{E}}_b\left[h_{i+1}^\star\right]\mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)] \qquad (9.54)\\ &+ \mathbb{E}[(K^\star-1)(K^\star-2)\mid K^\star\leq b]\mathbb{E}[\xi(h)]^2 + \mathbb{E}[K^\star-1\mid K^\star\leq b]\mathbb{E}[\xi(h)^2].\end{aligned}$$

We start analysing the case where $\mathbb{E}[K^3] < \infty$. By Theorem 8.1, for $\mathbb{E}[K^3] < \infty$,

$$\mathbb{E}[\xi(h)] \leq C_0(\beta-\beta_c)^{1/2}, \tag{9.55}$$

for some constant C_0. Substituting (9.52), and iterating in a similar fashion as in the proof of (9.52), we obtain that, for $\mathbb{E}[K^3] < \infty$,

$$\overline{\mathbb{E}}_b\left[(h_i^\star)^2\right] \leq C(\beta-\beta_c). \tag{9.56}$$

We next extend this analysis to $\tau \in (3,5)$. Note that, for every $a > 0$,

$$\mathbb{E}[(K^\star)^a \mid K^\star \le b] = \frac{\mathbb{E}[K^{a+1}\mathbb{1}_{\{K\le b\}}]}{\mathbb{E}[K\mathbb{1}_{\{K\le b\}}]}, \tag{9.57}$$

so that, for $\tau \in (3,5)$,

$$\mathbb{E}[(K^\star)^2 \mid K^\star \le b] \le \frac{C_{3,\tau}}{\mathbb{E}[K\mathbb{1}_{\{K\le b\}}]} b^{5-\tau}, \tag{9.58}$$

Further, for $\tau \in (3,5)$,

$$\mathbb{E}[\xi(h)] \le C_0(\beta-\beta_c)^{1/(3-\tau)}, \tag{9.59}$$

and thus

$$\mathbb{E}[(K^\star)^2 \mid K^\star \le b]\mathbb{E}[\xi(h)]^2 C \le b^{5-\tau}\mathbb{E}[\xi(h)]^2 \le C(\beta-\beta_c)^{-(5-\tau)/(3-\tau)+2/(3-\tau)} = C(\beta-\beta_c). \tag{9.60}$$

It can readily be seen that all other contributions to $\overline{\mathbb{E}}_b\left[(h_i^\star)^2\right]$ are of the same or smaller order. For example, when $\mathbb{E}[K^2] < \infty$ and using that $1/(\tau-3) \ge 1/2$ for all $\tau \in (3,5)$,

$$\mathbb{E}[K^\star - 1 \mid K^\star \le b]\mathbb{E}[\xi(h)^2] \le C\mathbb{E}[\xi(h)]^2 = O(\beta-\beta_c), \tag{9.61}$$

while, when $\tau \in (3,4)$,

$$\begin{aligned}\mathbb{E}[K^\star - 1 \mid K^\star \le b]\mathbb{E}[\xi(h)^2] &\le C b^{4-\tau}\mathbb{E}[\xi(h)]^{\tau-2} \\ &= C(\beta-\beta_c)^{-(4-\tau)/(3-\tau)+(\tau-2)/(3-\tau)} = C(\beta-\beta_c)^2.\end{aligned} \tag{9.62}$$

We conclude that

$$\overline{\mathbb{E}}_b\left[(h_i^\star)^2\right] \le C(\beta-\beta_c). \tag{9.63}$$

Therefore,

$$\chi(\beta,B) \ge c_\chi \sum_{\ell=0}^{m} \exp\Big\{-C\ell(\beta-\beta_c)\Big\} = O((\beta-\beta_c)^{-1}), \tag{9.64}$$

as required.

The proof for $\tau = 5$ is similar when noting that the logarithmic corrections present in $\mathbb{E}[\xi(h)]^2$ and in $\mathbb{E}[(K^\star)^2 \mid K^\star \le b]$ precisely cancel. □

We close this section by performing a heuristic argument to determine the upper bound on γ'. Unfortunately, as we discuss in more detail following the heuristics, we are currently not able to turn this analysis into a rigorous proof.

The upper bound on γ': heuristics for $\mathbb{E}[K^3] < \infty$. We can bound from above

$$\chi(\beta,B) \le \frac{\mathbb{E}[D]}{\nu}\sum_{\ell=0}^{\infty}(\hat{\beta}\nu)^\ell \mathbb{E}\left[\exp\Big\{-\sum_{i=1}^{\ell}\log\Big(1+\frac{\cosh(2h_i^\star)-1}{2\cosh(\beta)^2}\Big)\Big\}\right]. \tag{9.65}$$

Now, the problem is that $\hat{\beta}\nu > 1$ when $\beta > \beta_c$, so that we need to extract extra decay from the exponential term, which is technically demanding, and requires us to know various constants rather precisely. Let us show this heuristically. It suffices to study large values of ℓ, since small values can be bounded in a simple way.

We blindly put the expectation in the exponential, and Taylor expand to obtain that

$$\chi(\beta,B) \approx \frac{\mathbb{E}[D]}{\nu}\sum_{\ell=0}^{\infty}(\hat{\beta}\nu)^{\ell}\exp\Big\{-\sum_{i=1}^{\ell}\frac{\mathbb{E}\big[(h_i^{\star})^2\big]}{\cosh(\beta)^2}\Big\}. \tag{9.66}$$

We compute that

$$\cosh(\beta)^2 = \frac{1}{1-\hat{\beta}^2}. \tag{9.67}$$

Since

$$h_i^{\star} \approx \hat{\beta}h_{i+1}^{\star} + \sum_{j=1}^{K_i^{\star}-1}\xi(h_{i,j}), \tag{9.68}$$

we have

$$\mathbb{E}\big[h_i^{\star}\big] \approx \frac{\mathbb{E}[K^{\star}-1]}{1-\hat{\beta}}\mathbb{E}[\xi(h)], \tag{9.69}$$

and

$$\begin{aligned}\mathbb{E}\big[(h_i^{\star})^2\big] \approx{}& \frac{2\hat{\beta}\mathbb{E}[K^{\star}-1]^2+\mathbb{E}[(K^{\star}-1)(K^{\star}-2)](1-\hat{\beta})}{(1-\hat{\beta}^2)(1-\hat{\beta})}\mathbb{E}[\xi(h)]^2\\ &+\frac{\mathbb{E}[K^{\star}-1]}{1-\hat{\beta}^2}\mathbb{E}[\xi(h)^2].\end{aligned} \tag{9.70}$$

Ignoring all error terms in the proof of Lemma 8.3 shows that

$$\mathbb{E}[\xi(h)^2] \approx \frac{\nu_2\hat{\beta}^2}{1-\hat{\beta}}\mathbb{E}[\xi(h)]^2 = C_2\mathbb{E}[\xi(h)]^2, \tag{9.71}$$

so in total we arrive at (also using that $\hat{\beta} \approx 1/\nu$)

$$\mathbb{E}\big[(h_i^{\star})^2\big] \approx \frac{\nu_3(1-\hat{\beta})/\nu+3\nu_2^2/\nu^3}{(1-\hat{\beta}^2)(1-\hat{\beta})}\mathbb{E}[\xi(h)]^2. \tag{9.72}$$

As a result,

$$\frac{\mathbb{E}\big[(h_i^{\star})^2\big]}{\cosh(\beta)^2} \approx \frac{\nu_3(1-\hat{\beta})/\nu+3\nu_2^2/\nu^3}{1-\hat{\beta}}\mathbb{E}[\xi(h)]^2. \tag{9.73}$$

Ignoring error terms in the computation in Lemma 8.5 shows that

$$\mathbb{E}[\xi(h)^3] \approx C_3\mathbb{E}[\xi(h)]^3, \tag{9.74}$$

where

$$\begin{aligned}C_3 &= \frac{\hat{\beta}^3}{1-\hat{\beta}^3\nu}\big(\nu_3+3\nu_2C_2\big) \approx \frac{\hat{\beta}^3}{1-\hat{\beta}^2}\big(\nu_3+3\nu_2C_2\big)\\ &= \frac{\hat{\beta}^3}{(1-\hat{\beta}^2)(1-\hat{\beta})}\Big(\nu_3(1-\hat{\beta})+3(\nu_2/\nu)^2\Big),\end{aligned} \tag{9.75}$$

since $\hat{\beta} \approx 1/\nu$. Further, again ignoring error terms in (8.34) and Taylor expanding to third order shows that

$$\mathbb{E}[\xi(h)] \approx \hat{\beta}\nu\mathbb{E}[\xi(h)] - C_1\mathbb{E}[\xi(h)]^3, \tag{9.76}$$

where

$$C_1 = -\frac{\xi'''(0)}{6}\big(\nu C_3 + 3\nu_2 C_2 + \nu_3\big), \tag{9.77}$$

and $\xi'''(0) = -2\hat{\beta}(1-\hat{\beta}^2)$. Substituting the definitions for C_2 and C_3 yields

$$\begin{aligned} C_1 &= \frac{\hat{\beta}(1-\hat{\beta}^2)}{3}\big(\nu C_3 + 3\nu_2 C_2 + \nu_3\big) \\ &= \frac{\hat{\beta}}{3(1-\hat{\beta})}\big(\nu\hat{\beta}^3\nu_3(1-\hat{\beta}) + 3\nu\hat{\beta}^3(\nu_2/\nu)^2 + 3\nu_2^2\hat{\beta}^2(1-\hat{\beta}^2) + \nu_3(1-\hat{\beta})(1-\hat{\beta}^2)\big) \\ &= \frac{\hat{\beta}}{3(1-\hat{\beta})}\big(\nu_3(1-\hat{\beta}) + 3\nu_2^2\hat{\beta}^2\big). \end{aligned} \tag{9.78}$$

Thus, we arrive at

$$\mathbb{E}[\xi(h)]^2 \approx \frac{\hat{\beta}\nu - 1}{C_1}, \tag{9.79}$$

so that substitution into (9.73) leads to

$$\frac{\mathbb{E}\left[(h_i^\star)^2\right]}{\cosh(\beta)^2} \approx (\hat{\beta}\nu - 1)\frac{3\big(\nu_3(1-\hat{\beta})/\nu + 3\nu_2^2/\nu^3\big)}{\hat{\beta}\big(\nu_3(1-\hat{\beta}) + 3\nu_2^2\hat{\beta}^2\big)} = 3(\hat{\beta}\nu - 1). \tag{9.80}$$

We conclude that

$$(\hat{\beta}\nu)\exp\Big\{-\frac{\mathbb{E}\left[(h_i^\star)^2\right]}{\cosh(\beta)^2}\Big\} \le \big(1 + (\hat{\beta}\nu - 1)\big)e^{-3(\hat{\beta}\nu-1)} \le e^{-2(\hat{\beta}\nu-1)}. \tag{9.81}$$

This suggests that

$$\chi(\beta, B) \le \frac{\mathbb{E}[D]}{\nu}\sum_{\ell=0}^{\infty} e^{-2\ell(\hat{\beta}\nu-1)} = O((\hat{\beta}\nu - 1)^{-1}), \tag{9.82}$$

as required. Also, using (9.66), this suggests that

$$\lim_{\beta \searrow \beta_c}(\hat{\beta}\nu - 1)\chi(\beta, 0^+) = \mathbb{E}[D]/(2\nu), \tag{9.83}$$

where the constant is precisely half the one for the subcritical susceptibility (see (9.1)). It can be seen by an explicit computation that the same factor $1/2$ is also present in the same form for the Curie-Weiss model [16]. Indeed for the Boltzmann-Gibbs measure with Hamiltonian $H_N(\sigma) = -\frac{1}{2N}\sum_{i,j}\sigma_i\sigma_j$ one has $\beta_c = 1$ and a susceptibility $\chi(\beta, 0^+) = 1/(1-\beta)$ for $\beta < \beta_c$, $\chi(\beta, 0^+) = (1-m^2)/(1-\beta(1-m^2))$ with m the non-zero solution of $m = \tanh(\beta m)$ for $\beta > \beta_c$. Expanding this gives $m^2 = 3(\beta - 1)(1 + o(1))$ for $\beta \searrow 1$ and hence $\chi(\beta, 0^+) = (1 + o(1))/(1 - \beta(1 - 3(\beta - 1))) = (1 + o(1))/(2(\beta - 1))$.

It is a non-trivial task to turn the heuristic of this section into a proof because of several reasons: (a) We need to be able to justify the step where we put expectations in the exponential. While we are dealing with random variables with small means, they are not independent, so this is demanding; (b) We need to know the constants very precisely, as we are using the fact that a positive and negative term cancel in (9.81). The analysis performed in the previous sections does not give optimal control over these constants, so this step also requires substantial work.

The above heuristic does not apply to $\tau \in (3,5]$. However, the constant in (9.80) is *always* equal to 3, irrespective of the degree distribution. This suggests that also for $\tau \in (3,5]$, we should have $\gamma' \leq 1$.

Part II

Antiferromagnetic Potts model on Erdős-Rényi random graphs

10

EXISTENCE OF THERMODYNAMIC LIMIT OF THE PRESSURE

In this chapter we use superadditivity to prove that the thermodynamic limit of the quenched pressure exists for the antiferromagnetic Potts model.

Recall that the antiferromagnetic Potts model without external field is defined by the Boltzmann-Gibbs measure as in (3.22), which can alternatively be written as

$$\mu(\sigma) = \frac{1}{Z_N(\beta)} e^{-\beta H(\sigma)}, \tag{10.1}$$

where

$$H(\sigma) = \sum_{i,j=1}^{N} J_{i,j} \delta(\sigma_i, \sigma_j), \tag{10.2}$$

with $J_{i,j} \geq 0$, and

$$Z_N(\beta) = \sum_{\sigma \in [q]^N} e^{-\beta H(\sigma)}. \tag{10.3}$$

We study this model on a Poissonian Erdős-Rényi random graph, and hence let $(J_{i,j})_{i,j \in [N]}$ be i.i.d. Poisson random variables with

$$\mathbb{E}[J_{i,j}] = \frac{c}{2N}, \tag{10.4}$$

so that all vertices in the underlying graph will have expected degree c. Note that due to this construction we study the Potts model on a complete graph with *disorder*, because the interaction strength between two vertices is random. As mentioned in Chapter 2, self-loops and multiple edges might occur, but there will only be a small number of them and hence the effect of these self-loops and multiple edges will be negligible.

10.1 Results

We now state the main result in this chapter, namely the existence of the pressure:

Theorem 10.1. *The thermodynamic limit of the quenched pressure per vertex exists and is equal to*

$$p(\beta) = \sup_N p_N(\beta) < \infty. \tag{10.5}$$

We prove this theorem by using an interpolation scheme to show that the sequence $p_N(\beta)$ is superadditive in Section 10.3 and then use this to prove the existence in Section 10.4.

10.2 Discussion

Relation to the ferromagnet. The techniques as used in Part I of this thesis do not work for the antiferromagnetic Potts model. One reason is that in this case correlation inequalities like the GKS inequality do not hold. Also, the locally treelike behavior of the Erdős-Rényi random graph does not simplify things, since long loops present in the graph cause frustration and thus cannot be ignored.

Interpolation. The interpolation scheme to prove existence of the thermodynamic limit of the pressure was introduced in [60]. There it was used to prove that the pressure of the Sherrington-Kirkpatrick model is superadditive. After this, it has been employed for many other models, for example for the Viana-Bray model, a spin glass on the Erdős-Rényi random graph with interactions $J_{i,j}$ that have a symmetric distribution [61], and the p-spin model where the interactions depend on p Ising spins [54].

10.3 Superadditivity

As usual in disordered systems, it is convenient to study multiple copies, often called *replicas*, of the system. For this, consider for $n \in \mathbb{N}$ the product Boltzmann-Gibbs measure and define the expectation with respect to this measure as

$$\left\langle f(\sigma^{(1)},\ldots,\sigma^{(n)})\right\rangle = \frac{1}{Z^n(\beta)} \sum_{\sigma^{(1)},\ldots,\sigma^{(n)}\in[q]^N} f(\sigma^{(1)},\ldots,\sigma^{(n)}) e^{-\beta\left(H_N(\sigma^{(1)})+\ldots+H_N(\sigma^{(n)})\right)}. \tag{10.6}$$

An important observable that appears later is the sequence $q_N(r_1, r_2, \ldots, r_n)$, for $n \in \mathbb{N}$, with $(r_1, r_2, \ldots, r_n) \in [q]^n$, which represents a generalized multi-overlap between the n spin configurations $\sigma^{(1)}, \ldots, \sigma^{(n)}$, and is defined as

$$q_N(r_1,\ldots,r_n) = \frac{1}{N}\sum_{i=1}^N \delta(\sigma_i^{(1)}, r_1)\cdots\delta(\sigma_i^{(n)}, r_n). \tag{10.7}$$

Using these replicas we can prove that the pressure is superadditive.

Proposition 10.2 (Superadditivity)**.** *The quenched pressure per vertex is a superadditive sequence, i.e., for all $N_1, N_2, N \in \mathbb{N}$ with $N_1 + N_2 = N$,*

$$N p_N(\beta) \geq N_1 p_{N_1}(\beta) + N_2 p_{N_2}(\beta). \tag{10.8}$$

Proof. The proof is obtained by interpolation. For a partition of a system of size N into two subsystems of sizes N_1 and N_2 and for a $t \in [0,1]$ we consider the following independent Poisson random variables:

$$J'_{i,j} \sim \mathrm{Poi}\Big(\frac{ct}{2N}\Big), \quad J''_{i,j} \sim \mathrm{Poi}\Big(\frac{c(1-t)}{2N_1}\Big), \quad J'''_{i,j} \sim \mathrm{Poi}\Big(\frac{c(1-t)}{2N_2}\Big). \tag{10.9}$$

We define the interpolating Hamiltonian

$$H_N(\sigma,t) = \sum_{i,j=1}^{N} J'_{i,j}\delta(\sigma_i,\sigma_j) + \sum_{i,j=1}^{N_1} J''_{i,j}\delta(\sigma_i,\sigma_j) + \sum_{i,j=N_1+1}^{N} J'''_{i,j}\delta(\sigma_i,\sigma_j), \tag{10.10}$$

which induces an interpolating random partition function

$$Z_N(\beta,t) = \sum_{\sigma\in[q]^N} e^{-\beta H_N(\sigma,t)}, \tag{10.11}$$

the expectation with respect to an interpolating random Boltzmann-Gibbs measure

$$\langle f(\sigma)\rangle_t = \frac{1}{Z_N(\beta,t)} \sum_{\sigma\in[q]^N} f(\sigma)e^{-\beta H_N(\sigma,t)}, \tag{10.12}$$

and an interpolating quenched pressure

$$p_N(\beta,t) = \frac{1}{N}\mathbb{E}\left[\log Z_N(\beta,t)\right]. \tag{10.13}$$

Since $p_N(\beta,1) = p_N(\beta)$ and $p_N(\beta,0) = \frac{N_1}{N}p_{N_1}(\beta) + \frac{N_2}{N}p_{N_2}(\beta)$ the proposition follows from the fundamental theorem of calculus if one can show that the interpolating pressure is monotonically non-decreasing in t.

Note that for a Poisson random variable X with parameter λ,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbb{P}[X=x] = e^{-\lambda}\frac{\lambda^{x-1}}{(x-1)!} - e^{-\lambda}\frac{\lambda^x}{x!}, \tag{10.14}$$

and hence, for a vector $X = (X_1,\dots,X_m)$ of m independent Poisson random variables X_i with parameter $\lambda_i(t)$ and a function $f : \mathbb{N}^m \to \mathbb{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[f(X)] = \mathbb{E}\left[\sum_{i=1}^{m}\frac{\mathrm{d}\lambda_i(t)}{\mathrm{d}t}\left(f(X_1,\dots,X_i+1,\dots,X_m) - f(X_1,\dots,X_i,\dots,X_m)\right)\right]. \tag{10.15}$$

We apply this to the function

$$f(t) = \frac{1}{N}\log Z_N(\beta,t), \tag{10.16}$$

which depends on the Poisson random variables $(J'_{i,j})_{i,j=1}^{N}$,$(J''_{i,j})_{i,j=1}^{N_1}$ and $(J'''_{i,j})_{i,j=N_1+1}^{N}$. Abusing notation, we write $f(t,X'_{i,j})$ for $f(t)$ with $J'_{i,j} = X'_{i,j}$. Then,

$$f(t,J'_{i,j}+1) - f(t,J'_{i,j}) = \frac{1}{N}\log\Bigg(\sum_{\sigma\in[q]^N} e^{-\beta H_N(\sigma,t)-\beta\delta(\sigma_i,\sigma_j)}\Bigg) - \frac{1}{N}\log Z_N(\beta,t) \tag{10.17}$$

$$= \frac{1}{N}\log\frac{\sum_{\sigma\in[q]^N} e^{-\beta\delta(\sigma_i,\sigma_j)}e^{-\beta H_N(\sigma,t)}}{Z_N(\beta,t)} = \frac{1}{N}\log\left\langle e^{-\beta\delta(\sigma_i,\sigma_j)}\right\rangle_t.$$

Similar computations hold for $J_{i,j}^{''}$ and $J_{i,j}^{'''}$. Using this, the derivative of the interpolating pressure equals

$$\frac{\mathrm{d}p_N(\beta,t)}{\mathrm{d}t} = \frac{c}{2}\mathbb{E}\left[\frac{1}{N^2}\sum_{i,j=1}^{N}\log\left\langle e^{-\beta\delta(\sigma_i,\sigma_j)}\right\rangle_t \right. \tag{10.18}$$
$$\left. -\frac{1}{NN_1}\sum_{i,j=1}^{N_1}\log\left\langle e^{-\beta\delta(\sigma_i,\sigma_j)}\right\rangle_t - \frac{1}{NN_2}\sum_{i,j=N_1+1}^{N}\log\left\langle e^{-\beta\delta(\sigma_i,\sigma_j)}\right\rangle_t\right].$$

Expression (10.18) can be further simplified using the identity

$$e^{-\beta\delta(\sigma_i,\sigma_j)} = 1-(1-e^{-\beta})\delta(\sigma_i,\sigma_j), \tag{10.19}$$

and the Taylor expansion

$$\log(1-x) = -\sum_{n=1}^{\infty}\frac{x^n}{n}, \qquad \forall |x|<1. \tag{10.20}$$

One obtains

$$\frac{\mathrm{d}p_N(\beta,t)}{\mathrm{d}t} = -\frac{c}{2}\mathbb{E}\left[\sum_{n=1}^{\infty}\frac{(1-e^{-\beta})^n}{n}\left(\frac{1}{N^2}\sum_{i,j=1}^{N}\left\langle\delta(\sigma_i,\sigma_j)\right\rangle_t^n\right.\right.$$
$$\left.\left. -\frac{1}{NN_1}\sum_{i,j=1}^{N_1}\left\langle\delta(\sigma_i,\sigma_j)\right\rangle_t^n - \frac{1}{NN_2}\sum_{i,j=N_1+1}^{N}\left\langle\delta(\sigma_i,\sigma_j)\right\rangle_t^n\right)\right]. \tag{10.21}$$

We want to rewrite this in terms of generalized overlaps. Note that

$$\frac{1}{N^2}\sum_{i,j=1}^{N}\delta(\sigma_i^{(1)},\sigma_i^{(1)})\cdots\delta(\sigma_i^{(n)},\sigma_i^{(n)}) = \sum_{r_1,\dots,r_n=1}^{q} q_N^2(r_1,\dots,r_n), \tag{10.22}$$

and hence

$$\frac{\mathrm{d}p_N(\beta,t)}{\mathrm{d}t} = -\frac{c}{2}\mathbb{E}\left[\sum_{n=1}^{\infty}\frac{(1-e^{-\beta})^n}{n}\right. \tag{10.23}$$
$$\left.\sum_{r_1,\dots,r_n=1}^{q}\left\langle q_N^2(r_1,\dots,r_n) - \frac{N_1}{N}q_{N_1}^2(r_1,\dots,r_n) - \frac{N_2}{N}q_{N_2}^2(r_1,\dots,r_n)\right\rangle_t\right].$$

Note that

$$q_N(r_1,\dots,r_n) = \frac{N_1}{N}q_{N_1}(r_1,\dots,r_n) + \frac{N_2}{N}q_{N_2}(r_1,\dots,r_n), \tag{10.24}$$

i.e., $q_N(r_1,\dots,r_n)$ is a convex linear combination of $q_{N_1}(r_1,\dots,r_n)$ and $q_{N_2}(r_1,\dots,r_n)$. Combining this with the fact that the function $f : x \mapsto x^2$ is a convex function shows that, a.s.,

$$q_N^2(r_1,\dots,r_n) - \frac{N_1}{N}q_{N_1}^2(r_1,\dots,r_n) - \frac{N_2}{N}q_{N_2}^2(r_1,\dots,r_n) \le 0, \tag{10.25}$$

and hence

$$\frac{\mathrm{d}p_N(\beta,t)}{\mathrm{d}t} \ge 0. \tag{10.26}$$

□

Remark 10.3. The same computation goes through also for the ferromagnetic model with the change $\beta \mapsto -\beta$. However $1 - e^{\beta} < 0$ for $\beta > 0$ and therefore the series expansion in (10.23), which is only allowed for $\beta < \log 2$, has alternating signs and monotonicity can not be derived anymore by inspection. We believe however that the interpolation is monotone also in the ferromagnetic case, though in the opposite direction. This belief is based on two facts. Firstly, pressure subadditivity for the ferromagnetic model on the Erdős-Rényi random graph has been checked numerically for small system sizes [6]. This is in agreement with a monotonic behavior of the interpolating pressure. Secondly, and more importantly, the numerical checks [6] for the Ising case ($q = 2$) show that for $0 \le t \le 1$ the series in (10.23) is dominated by the first term and therefore one would be left with the same interpolating pressure of the Curie-Weiss model which is known to be sub-additive. This is indeed rigorously shown at zero temperature in [96]. Although the ferromagnetic model has been fully solved in [34], it would be interesting to extend the monotonicity result to all temperatures.

10.4 Existence

We can now use superadditivity to prove the existence of the thermodynamic limit of the pressure and its realization as a supremum.

Proof of Theorem 10.1. The theorem is a direct consequence of Proposition 10.2 and Fekete's lemma [94] of which we recapitulate the proof in the present context. Suppose that $M \le N$. Then we can choose k and ℓ such that $N = kM + \ell$, with $0 \le \ell < M$. Because of superadditivity we know that

$$kMp_{kM} \ge kMp_M. \tag{10.27}$$

Hence, using superadditivity once more,

$$Np_N \ge kMp_M + \ell p_\ell \ge kMp_M + \min_{0\le i<M} ip_i. \tag{10.28}$$

Dividing both sides by N and taking $\liminf_{N\to\infty}$ gives

$$\liminf_{N\to\infty} p_N \ge \liminf_{N\to\infty} \left(\frac{N-\ell}{N} p_M + \frac{1}{N} \min_{0\le i<M} ip_i \right) = p_M. \tag{10.29}$$

Since this holds for all M we can take the supremum over M on the right-hand side, so that

$$\liminf_{N\to\infty} p_N \ge \sup_{M\in\mathbb{N}} p_M \ge \limsup_{M\to\infty} p_M, \tag{10.30}$$

where the last inequality trivially holds. Of course, it also holds that

$$\limsup_{N\to\infty} p_N \ge \liminf_{N\to\infty} p_N, \tag{10.31}$$

thus showing that $\lim_{N\to\infty} p_N$ exists and is equal to $\sup_{N\in\mathbb{N}} p_N$. To show that this supremum is finite, we observe that $e^{-\beta H(\sigma)} \le 1$, a.s., for all $\sigma \in [q]^N$ and hence

$$p_N = \frac{1}{N}\mathbb{E}[\log Z_N] = \frac{1}{N}\mathbb{E}\left[\log \sum_{\sigma\in[q]^N} e^{-\beta H(\sigma)} \right] \le \log q < \infty. \tag{10.32}$$

□

In Chapter 11 we improve this upper bound on $p_N(\beta)$.

11

Extended variational principle

In the previous chapter we have shown that the thermodynamic limit of the pressure exists and equals

$$p(\beta) = \sup_N p_N(\beta) < \infty. \tag{11.1}$$

In this chapter we give a different characterization for $p(\beta)$ by showing that

$$p(\beta) = \lim_{N\to\infty} \inf_{\mathscr{L}} G_N(\beta, \mathscr{L}), \tag{11.2}$$

for some explicit formula $G_N(\beta, \mathscr{L})$ and the infimum is over a suitable set of laws $\mathscr{L}$. This is called the *extended variational principle,* which was developed in [5] for the Sherrington-Kirkpatrick (SK) model. We prove in this chapter that it also holds for the antiferromagnet. Furthermore, we use it to derive upper bounds on the pressure by not taking the infimum over $\mathscr{L}$, but by choosing specific $\mathscr{L}$.

11.1 Results

To state the results in this chapter, we need the notion of exchangeable measures.

Definition 11.1 (Exchangeable measure). *Let $(\mu_\alpha)_{\alpha=1}^{\infty}$ be a random sequence such that each μ_α is positive and summable, almost surely. Also let $(\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}$ be a random array of elements of $[q]$. We write $\mathscr{L}$ for the measure which describes the joint distribution of*

$$((\mu_\alpha)_{\alpha\in\mathbb{N}},\ (\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}). \tag{11.3}$$

We say that $\mathscr{L}$ is exchangeable *if*

$$((\mu_\alpha)_{\alpha\in\mathbb{N}},\ (\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}) \stackrel{\mathrm{d}}{=} ((\mu_\alpha)_{\alpha\in\mathbb{N}},\ (\tau_{\alpha,\pi(k)})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}), \tag{11.4}$$

for every non-random permutation π of $\mathbb{N}$ which moves only finitely many k's.

Then, we have the following theorem:

Theorem 11.2 (Extended variational principle). *For all $\beta \geq 0$,*

$$p(\beta) = \lim_{N\to\infty} \inf_{\mathscr{L}} G_N(\beta, \mathscr{L}), \tag{11.5}$$

where the infimum is over all exchangeable laws $\mathscr{L}$, and where

$$G_N(\beta, \mathscr{L}) = G_N^{(1)}(\beta, \mathscr{L}) - G_N^{(2)}(\beta, \mathscr{L}), \tag{11.6}$$

with

$$G_N^{(1)}(\beta, \mathscr{L}) = \frac{1}{N}\mathbb{E}\Big[\log \sum_{\alpha=1}^{\infty} \mu_\alpha \sum_{\sigma\in[q]^N} \exp\Big(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{\alpha,k})\Big)\Big], \tag{11.7}$$

and

$$G_N^{(2)}(\beta, \mathscr{L}) = \frac{1}{N}\mathbb{E}\Big[\log \sum_{\alpha=1}^{\infty} \mu_\alpha \exp\Big(-\beta \sum_{k=1}^{L} \delta(\tau_{\alpha,2k-1}, \tau_{\alpha,2k})\Big)\Big], \tag{11.8}$$

where $\mathbb{E}$ is the expectation over $\mathscr{L}$ as well as the random variables $(I(k))_{k=1}^{\infty}$, K and L, where we assume that $(I(k))_{k=1}^{\infty}, K, L$ are all independent of one another and of $((\mu_\alpha), (\tau_{\alpha,k}))$, and $I(1), I(2), \ldots$ are i.i.d. uniform on $[N]$, K is Poisson with mean cN and L is Poisson with mean $cN/2$.

By not taking the infimum over all laws $\mathscr{L}$, but by choosing an explicit $\mathscr{L}$ we get an upper bound on the pressure. We give two examples in the next theorem:

Theorem 11.3 (Upper bounds on the pressure). *The following upper bounds on the pressure hold for all $\beta \geq 0$:*

(a) High-temperature solution. *For all $N \in \mathbb{N}$,*

$$p_N(\beta) \leq p^{\mathrm{HT}}(\beta) \equiv \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) + \log q. \tag{11.9}$$

(b) Replica-symmetric solution.

$$p(\beta) \leq \mathbb{E}\log\Big[\sum_{\sigma=1}^{q}\prod_{k=1}^{K_1}\Big(1-(1-e^{-\beta})P_k(\sigma)\Big)\Big] - \frac{c}{2}\mathbb{E}\log\Big[1-(1-e^{-\beta})\sum_{\sigma=1}^{q} P_1(\sigma)P_2(\sigma)\Big], \tag{11.10}$$

where K_1 is a Poisson random variable with mean c and the $P_k = (P_k(\sigma))_{\sigma\in[q]}$ are i.i.d. random vectors, satisfying, a.s., $P_k(\sigma) \geq 0$ for all $\sigma \in [q]$ and $\sum_{\sigma=1}^{q} P_k(\sigma) = 1$.

In the next chapter we show that the high-temperature solution is indeed correct for high temperatures.

11.2 Discussion

The extended variational principle. The extended variational principle was first introduced in [5] for the SK model. Before this, it was rigorously proved that the famous Parisi solution [92] was an upper bound for the pressure in [59]. The work [5], showed that this upper bound fits in the more general framework of the extended variational principle. The proof that Parisi's solution is correct was given in [100] by showing that the difference between the pressure and the Parisi solution vanishes in the thermodynamic limit.

Replica-symmetric solution. One can optimize over the distribution of P_k and it can be shown that the optimal choice is such that $(P_k(\sigma))_{\sigma\in[q]}$ satisfies the distributional recursion

$$(P_k(\sigma))_{\sigma\in[q]} \stackrel{\mathrm{d}}{=} \left(\frac{\prod_{k=1}^{K_1} \left(1-(1-e^{-\beta})P_k(\sigma)\right)}{\sum_{\sigma\in[q]} \prod_{k=1}^{K_1} \left(1-(1-e^{-\beta})P_k(\sigma)\right)} \right)_{\sigma\in[q]}, \tag{11.11}$$

which is comparable to (4.1). See [27] for details. The replica-symmetric solution is comparable to the solution of the ferromagnet in Chapter 5 and its Potts generalizations in [36, 37].

11.3 Extended variational principle

We begin by adding our system with N vertices to a second system with M vertices, where one should think of $M \gg N$. Or, equivalently, removing the system with N vertices from a system with $M+N$ vertices, thus leading to a so-called *cavity*.

Proposition 11.4. *Suppose that $N, M \in \mathbb{N}$ are chosen and let μ_M be any measure on $[q]^M$. Then,*

$$p_N(\beta) \le G_{N,M}(\beta, \mu_M), \tag{11.12}$$

where

$$G_{N,M}(\beta,\mu_M) = G^{(1)}_{N,M}(\beta,\mu_M) - G^{(2)}_{N,M}(\beta,\mu_M), \tag{11.13}$$

with

$$G^{(1)}_{N,M}(\beta,\mu_M) = \frac{1}{N}\mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu_M(\tau) \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{i=1}^{N}\sum_{j=1}^{M} K_{i,j}\delta(\sigma_i,\tau_j)\right)\right], \tag{11.14}$$

where the $K_{i,j}$'s are i.i.d. Poisson random variables with parameter c/M, and

$$G^{(2)}_{N,M}(\beta,\mu_M) = \frac{1}{N}\mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu_M(\tau) \exp\left(-\beta \sum_{i,j=1}^{M} L_{i,j}\delta(\tau_i,\tau_j)\right)\right], \tag{11.15}$$

where the $L_{i,j}$'s are i.i.d. Poisson random variables with parameters $cN/(2M^2)$.

Proof. For a $t \in [0,1]$ we consider the following independent Poisson random variables:

$$(\tilde{J}_{i,j})_{i,j\in[N]}, \qquad (\tilde{K}_{i,j})_{i\in[N];\ j=\in[M]}, \qquad (\tilde{L}_{i,j})_{i,j\in[M]}, \tag{11.16}$$

such that

$$\mathbb{E}[\tilde{J}_{i,j}] = \frac{(1-t)c}{2N}, \quad \mathbb{E}[\tilde{K}_{i,j}] = \frac{ct}{M}, \quad \mathbb{E}[\tilde{L}_{i,j}] = \frac{(1-t)cN}{2M^2}, \tag{11.17}$$

for all appropriate indices i, j. We define

$$H_{N,M}(\sigma,\tau,t)=\sum_{i,j=1}^{N}\tilde{J}_{i,j}\delta(\sigma_i,\sigma_j)+\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{K}_{i,j}\delta(\sigma_i,\tau_j)+\sum_{i,j=1}^{M}\tilde{L}_{i,j}\delta(\tau_i,\tau_j), \tag{11.18}$$

and we also define

$$Z_{N,M}(\beta,\mu_M,t)=\sum_{\tau\in[q]^M}\mu_M(\tau)\sum_{\sigma\in[q]^N}e^{-\beta H_{N,M}(\sigma,\tau,t)}, \tag{11.19}$$

and

$$\langle f(\sigma,\tau)\rangle_t=\frac{1}{Z_{N,M}(\beta,\mu_M,t)}\sum_{\tau\in[q]^M}\mu_M(\tau)\sum_{\sigma\in[q]^N}f(\sigma,\tau)e^{-\beta H_{N,M}(\sigma,\tau,t)}. \tag{11.20}$$

Then

$$\frac{1}{N}\mathbb{E}[\log Z_{N,M}(\beta,\mu_M,0)]=p_N(\beta)+G^{(2)}_{N,M}(\beta,\mu_M), \tag{11.21}$$

since if $t=0$ then $H_{N,M}(\sigma,\tau,0)$ splits into a summand only depending on σ and one only depending on τ. Furthermore,

$$\frac{1}{N}\mathbb{E}[\log Z_{N,M}(\beta,\mu_M,1)]=G^{(1)}_{N,M}(\beta,\mu_M). \tag{11.22}$$

Moreover, as in the proof of Proposition 10.2, one can show that

$$\begin{aligned}
&\frac{d}{dt}\Big(\frac{1}{N}\mathbb{E}[\log Z_{N,M}(\beta,\mu_M,t)]\Big)\\
&\quad=-\frac{c}{2}\mathbb{E}\Bigg[\frac{1}{N^2}\sum_{i,j=1}^{N}\log\left\langle e^{-\beta\delta(\sigma_i,\sigma_j)}\right\rangle_t\\
&\qquad-\frac{2}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\log\left\langle e^{-\beta\delta(\sigma_i,\tau_j)}\right\rangle_t+\frac{1}{M^2}\sum_{i,j=1}^{M}\log\left\langle e^{-\beta\delta(\tau_i,\tau_j)}\right\rangle_t\Bigg]\\
&\quad=\frac{c}{2}\mathbb{E}\Bigg[\sum_{n=1}^{\infty}\frac{(1-e^{-\beta})^n}{n}\Bigg(\frac{1}{N^2}\sum_{i,j=1}^{N}\left\langle\delta(\sigma_i,\sigma_j)\right\rangle_t^n\\
&\qquad-\frac{2}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\langle\delta(\sigma_i,\tau_j)\right\rangle_t^n+\frac{1}{M^2}\sum_{i,j=1}^{M}\left\langle\delta(\tau_i,\tau_j)\right\rangle_t^n\Bigg)\Bigg].
\end{aligned} \tag{11.23}$$

This can again be rewritten in terms of generalized multi-overlaps as

$$\begin{aligned}
&\frac{c}{2}\mathbb{E}\Bigg[\sum_{n=1}^{\infty}\frac{(1-e^{-\beta})^n}{n}\sum_{r_1,\dots,r_n=1}^{q}\Big\langle q_N^2(r_1,\dots,r_n)\\
&\qquad-2q_N(r_1,\dots,r_n)q_M(r_1,\dots,r_n)+q_M^2(r_1,\dots,r_n)\Big\rangle_t\Bigg] \qquad (11.24)\\
&\quad=\frac{c}{2}\mathbb{E}\Bigg[\sum_{n=1}^{\infty}\frac{(1-e^{-\beta})^n}{n}\sum_{r_1,\dots,r_n=1}^{q}\Big\langle\big(q_N(r_1,\dots,r_n)-q_M(r_1,\dots,r_n)\big)^2\Big\rangle_t\Bigg],
\end{aligned}$$

which is obviously nonnegative for all t. This proves that

$$\begin{aligned} p_N(\beta) + G^{(2)}_{N,M}(\beta,\mu_M) &= \frac{1}{N}\mathbb{E}[\log Z_{N,M}(\beta,\mu_M,0)] \\ &\leq \frac{1}{N}\mathbb{E}[\log Z_{N,M}(\beta,\mu_M,1)] = G^{(1)}_{N,M}(\beta,\mu_M). \end{aligned} \tag{11.25}$$

In other words,

$$p_N(\beta) \leq G^{(1)}_{N,M}(\beta,\mu_M) - G^{(2)}_{N,M}(\beta,\mu_M). \tag{11.26}$$

□

In order to accommodate a limit where M approaches ∞, we mention an equivalent version of the function $G_{N,M}(\beta,\mu_M)$:

Lemma 11.5. *Let $I(1), I(2), \ldots$ be i.i.d. random variables uniformly chosen from $[N]$, and let $J(1), J(2), \ldots$ be i.i.d. random variables uniformly chosen from $[M]$. Independently of this, let K be a Poisson random variable with parameter cN and let L be a Poisson random variable with parameter $cN/2$. Then,*

$$G^{(1)}_{N,M}(\beta,\mu_M) = \frac{1}{N}\mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu_M(\tau) \sum_{\sigma\in[q]^N} \exp\left(-\beta\sum_{k=1}^{K}\delta(\sigma_{I(k)},\tau_{J(k)})\right)\right], \tag{11.27}$$

and

$$G^{(2)}_{N,M}(\beta,\mu_M) = \frac{1}{N}\mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu_M(\tau)\exp\left(-\beta\sum_{k=1}^{L}\delta(\tau_{J(2k-1)},\tau_{J(2k)})\right)\right]. \tag{11.28}$$

Proof. This follows from a property of Poisson random variables known as Poisson thinning. Previously we had i.i.d. Poisson random variables $K_{i,j}$ and $L_{i,j}$. Now we have just two Poisson random variables K and L, but we have i.i.d. uniform random variables $I(k)$, $J(k)$. The Poisson thinning property refers to the fact that the families

$$\hat{K}_{i,j} = \#\{k \leq K : (I(k),J(k)) = (i,j)\} \text{ and } \hat{L}_{i,j} = \#\{k \leq L : (J(2k-1),J(2k)) = (i,j)\}, \tag{11.29}$$

are distributed identically to $K_{i,j}$ and $L_{i,j}$, respectively. □

Corollary 11.6. *Suppose that μ_M is a random measure on $[q]^M$. Then,*

$$p_N(\beta) \leq \mathbb{E}\left[G_{N,M}(\beta,\mu_M)\right], \tag{11.30}$$

where the symbol $\mathbb{E}$ denotes the expectation with respect to μ_M.

Proof. For each random realization of μ_M,

$$p_N(\beta) \leq G_{N,M}(\beta,\mu_M), \tag{11.31}$$

almost surely, according to Proposition 11.4. Hence, the corollary follows by elementary properties of the expectation. □

The following proposition may be viewed as the $M\to\infty$ limit of Proposition 11.4:

Proposition 11.7. *For all $\beta \ge 0$ and exchangeable $\mathscr{L}$,*

$$p_N(\beta) \le G_N(\beta, \mathscr{L}), \tag{11.32}$$

where $G_N(\beta, \mathscr{L})$ is defined in Theorem 11.2.

Proof. For any fixed M, consider a random realization of the sequence $(\mu_\alpha)_{\alpha\in\mathbb{N}}$ and $(\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}$. Define the random measure μ_M on $[q]^M$ as

$$\mu_M(\tau) = \sum_{\alpha=1}^{\infty} \mu_\alpha \mathbb{1}_{\{(\tau_{\alpha,1},\ldots,\tau_{\alpha,M})=\tau\}}, \tag{11.33}$$

for each $\tau \in [q]^M$. This is the empirical measure, but where we merely truncate the full sequence $(\tau_{\alpha,1}, \tau_{\alpha,2}, \ldots)$ to the first M components of the spin. Another useful way to state the same thing is to notice that for any non-random function $f : [q]^M \to \mathbb{R}$,

$$\sum_{\tau\in[q]^M} f(\tau)\mu_M(\tau) = \sum_{\alpha=1}^{\infty} \mu_\alpha f(\tau_{\alpha,1}, \ldots, \tau_{\alpha,M}). \tag{11.34}$$

In fact, it is not necessary that f is non-random, merely that it is independent of $(\mu_\alpha)_{\alpha\in\mathbb{N}}$ and $(\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}$. Note that μ_M is a random measure, but according to Corollary 11.6 this still gives an upper bound. Specifically,

$$p_N(\beta) \le \mathbb{E}\left[G_{N,M}(\beta, \mu_M)\right], \tag{11.35}$$

where the expectation is over the law $\mathscr{L}$ for $(\mu_\alpha)_{\alpha\in\mathbb{N}}$ and $(\tau_{\alpha,k})_{\alpha\in\mathbb{N},\, k\in\mathbb{N}}$, from which μ_M is derived as a measurable function.

According to Lemma 11.5, we may write

$$G^{(1)}_{N,M}(\beta, \mu_M) = \frac{1}{N}\mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu_M(\tau) \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{J(k)})\right)\right], \tag{11.36}$$

where $I(1), I(2), \ldots$ are i.i.d. uniform on $[N]$, and $J(1), J(2), \ldots$ are i.i.d. uniform on $[M]$, and independently of this, K is a Poisson random variable with parameter cN. Note that, according to (11.34), we may write

$$\begin{aligned} &\sum_{\tau\in[q]^M} \mu_M(\tau) \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{J(k)})\right) \\ &\qquad = \sum_{\alpha=1}^{\infty} \mu_\alpha \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{\alpha,J(k)})\right). \end{aligned} \tag{11.37}$$

Conditionally on the event that $J(1), \ldots, J(K)$ are all distinct elements of $[M]$,

$$\begin{aligned} &\frac{1}{N}\log \sum_{\alpha=1}^{\infty} \mu_\alpha \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{\alpha,J(k)})\right) \\ &\qquad \stackrel{\mathrm{d}}{=} \frac{1}{N}\log \sum_{\alpha=1}^{\infty} \mu_\alpha \sum_{\sigma\in[q]^N} \exp\left(-\beta \sum_{k=1}^{K} \delta(\sigma_{I(k)}, \tau_{\alpha,k})\right) \end{aligned} \tag{11.38}$$

where we replaced the random indices $J(1),\dots,J(K)$ by the non-random indices $1,\dots,K$, because we assumed that $(\tau_{\alpha,1},\dots,\tau_{\alpha,M})$ are exchangeable, meaning equal in distribution under finite permutations. Note that here we use the fact that K and $J(1),J(2),\dots$ are independent of $(\mu_\alpha)_{\alpha\in\mathbb{N}}$ and $(\tau_{\alpha,k})_{\alpha\in\mathbb{N},\ k\in\mathbb{N}}$. Moreover, conditional on the value of K, the probability that $J(1),\dots,J(K)$ are all distinct is

$$\frac{M(M-1)\cdots(M-K+1)}{M^K}. \tag{11.39}$$

If we take a single realization of K for all M's, then we see that this conditional probability converges to 1, pointwise, almost surely. So we are justified in making the rearrangement in (11.38), with high probability. Moreover, conditioning on K, we see that the function on the left hand side of (11.38) is bounded in the interval

$$\left[\frac{1}{N}\log\sum_{\alpha=1}^{\infty}\mu_\alpha+\log(q)-\frac{\beta K}{N},\frac{1}{N}\log\sum_{\alpha=1}^{\infty}\mu_\alpha+\log(q)\right]. \tag{11.40}$$

This is summable against the distribution of K. Therefore, by the dominated convergence theorem,

$$\lim_{M\to\infty}G_{N,M}^{(1)}(\beta,\mu_M)=G_N^{(1)}(\beta,\mathscr{L}). \tag{11.41}$$

A similar argument holds for the second term $G_N^{(2)}(\beta,\mathscr{L})$. □

We can now prove the extended variational principle:

Proof of Theorem 11.2. By taking the infimum over $\mathscr{L}$ in Proposition 11.7 and then taking $\liminf_{N\to\infty}$ it follows that

$$p(\beta)\le\liminf_{N\to\infty}\inf_{\mathscr{L}}G_N(\beta,\mathscr{L}). \tag{11.42}$$

It remains to prove that

$$p(\beta)\ge\limsup_{N\to\infty}\inf_{\mathscr{L}}G_N(\beta,\mathscr{L}), \tag{11.43}$$

because then, combined with the fact that the limit superior is always greater or equal to the limit inferior,

$$p(\beta)\le\liminf_{N\to\infty}\inf_{\mathscr{L}}G_N(\beta,\mathscr{L})\le\limsup_{N\to\infty}\inf_{\mathscr{L}}G_N(\beta,\mathscr{L})\le p(\beta), \tag{11.44}$$

and hence all must be equal. This also shows that $\lim_{N\to\infty}\inf_{\mathscr{L}}G_N(\beta,\mathscr{L})$ indeed exists.

To prove (11.43), we first use superadditivity to show that

$$p(\beta)=\limsup_{N\to\infty}\liminf_{M\to\infty}\frac{1}{N}\left((M+N)p_{M+N}(\beta)-Mp_M(\beta)\right). \tag{11.45}$$

To show (11.45), we write (suppressing the dependence on β)

$$\begin{aligned}\frac{M+NK}{NK}p_{NK}-\frac{M}{NK}p_M&=\frac{1}{K}\sum_{k=0}^{K-1}\left(\frac{M+Nk+N}{N}p_{M+Nk+N}-\frac{M+Nk}{N}p_{M+Nk}\right)\\&\ge\inf_{M'\ge M}\left(\frac{M'+N}{N}p_{M'+N}-\frac{M'}{N}p_{M'}\right).\end{aligned} \tag{11.46}$$

By first taking the limit $K \to \infty$ and then the limit $M \to \infty$ we get

$$p(\beta) \geq \liminf_{M\to\infty} \left(\frac{M+N}{N} p_{M+N} - \frac{M}{N} p_M \right). \tag{11.47}$$

It remains to take the limit superior of $N \to \infty$ to prove the lower bound in (11.45). To get the upper bound in (11.45), note that

$$\frac{M+N}{N} p_{M+N} - \frac{M}{N} p_M \geq p_N, \tag{11.48}$$

by superadditivity, so that

$$\limsup_{N\to\infty} \liminf_{M\to\infty} \left(\frac{M+N}{N} p_{M+N} - \frac{M}{N} p_M \right) \geq \limsup_{N\to\infty} \liminf_{M\to\infty} p_N = p(\beta), \tag{11.49}$$

proving (11.45).

By definition,

$$(M+N)p_{M+N}(\beta) = \mathbb{E}\left[\log Z_{M+N}(\beta)\right] = \mathbb{E}\left[\log \sum_{\tau\in[q]^M} \sum_{\sigma\in[q]^N} e^{-\beta H_{M,N}(\tau,\sigma)} \right], \tag{11.50}$$

where, with random variables $J_{i,j}, J^{'}_{i,j} \sim \text{Poi}\left(\frac{c}{2(M+N)}\right)$ and $K_{i,j} \sim \text{Poi}\left(\frac{c}{M+N}\right)$ that are all independent of each other,

$$H_{M,N}(\tau,\sigma) = \sum_{i,j=1}^{M} J_{i,j}\delta(\tau_i,\tau_j) + \sum_{j=1}^{M}\sum_{i=1}^{N} K_{i,j}\delta(\tau_j,\sigma_i) + \sum_{i,j=1}^{N} J_{i,j}\delta(\sigma_i,\sigma_j). \tag{11.51}$$

We let

$$\mu^*_M(\tau) = \exp\Big(-\beta \sum_{i,j=1}^{M} J^{'}_{i,j}\delta(\tau_i,\tau_j)\Big), \tag{11.52}$$

to rewrite (11.50) as

$$\begin{aligned}
&(M+N)p_{M+N}(\beta) \\
&= \mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu^*_M(\tau) \sum_{\sigma\in[q]^N} \exp\Big(-\beta \sum_{j=1}^{M}\sum_{i=1}^{N} K_{i,j}\delta(\tau_j,\sigma_i) - \beta \sum_{i,j=1}^{N} J^{'}_{i,j}\delta(\sigma_i,\sigma_j)\Big)\right] \\
&\geq \mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu^*_M(\tau) \sum_{\sigma\in[q]^N} \exp\Big(-\beta \sum_{j=1}^{M}\sum_{i=1}^{N} K_{i,j}\delta(\tau_j,\sigma_i) - \beta \sum_{i,j=1}^{N} J^{'}_{i,j}\Big)\right] \qquad (11.53) \\
&= \mathbb{E}\left[\log \sum_{\tau\in[q]^M} \mu^*_M(\tau) \sum_{\sigma\in[q]^N} \exp\Big(-\beta \sum_{j=1}^{M}\sum_{i=1}^{N} K_{i,j}\delta(\tau_j,\sigma_i)\Big)\right] - \beta N^2 \mathbb{E}[J^{'}_{i,j}].
\end{aligned}$$

Using Poisson thinning we can rewrite this, with $K_M \sim \text{Poi}\left(\frac{cMN}{M+N}\right)$ independent of everything else, $J(k)$ i.i.d. uniform on $[M]$ and $I(k)$ i.i.d. uniform on $[N]$ both also independent of everything else, as

$$\mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \mu_M^*(\tau) \sum_{\sigma\in[q]^N} \exp\Bigg(-\beta \sum_{k=1}^{K_M} \delta(\tau_{J(k)}, \sigma_{I(k)})\Bigg)\Bigg] - \beta N^2 \frac{c}{2(M+N)}$$
$$\geq \mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \mu_M^*(\tau) \sum_{\sigma\in[q]^N} \exp\Bigg(-\beta \sum_{k=1}^{K_M+K_M'} \delta(\tau_{J(k)}, \sigma_{I(k)})\Bigg)\Bigg] - o_M(1), \quad (11.54)$$

where $K_M' \sim \text{Poi}\left(\frac{cN^2}{M+N}\right)$ independent of everything else and where $o_M(1)$ is a function going to zero for $M \to \infty$. Note that $K_M + K_M'$ is again Poisson distributed with expectation $\frac{cMN}{M+N} + \frac{cN^2}{M+N} = cN$. Hence, with $K \sim \text{Poi}(cN)$ independent of everything else, the above equals

$$\mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \mu_M^*(\tau) \sum_{\sigma\in[q]^N} \exp\Bigg(-\beta \sum_{k=1}^{K} \delta(\tau_{J(k)}, \sigma_{I(k)})\Bigg)\Bigg] - o_M(1)$$
$$= N G_{N,M}^{(1)}(\beta, \mu_M^*) - o_M(1), \quad (11.55)$$

with $G_{N,M}^{(1)}(\beta, \mu_M^*)$ as defined in (11.14).

Similarly, with $\tilde{J}_{i,j} \sim \text{Poi}\left(\frac{c}{2M}\right)$ i.i.d. independent of everything else,

$$M p_M(\beta) = \mathbb{E}[\log Z_M(\beta)] = \mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \exp\Bigg(-\beta \sum_{i,j=1}^{M} \tilde{J}_{i,j} \delta\tau_i, \tau_j\Bigg)\Bigg]$$
$$\leq \mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \exp\Bigg(-\beta \sum_{i,j=1}^{M} \tilde{J}_{i,j} \delta\tau_i, \tau_j - \beta \sum_{i,j=1}^{M} \tilde{J}_{i,j}'(\delta\tau_i, \tau_j - 1)\Bigg)\Bigg]$$
$$= \mathbb{E}\Bigg[\log \sum_{\tau\in[q]^M} \exp\Bigg(-\beta \sum_{i,j=1}^{M} (\tilde{J}_{i,j} + \tilde{J}_{i,j}') \delta\tau_i, \tau_j + \beta \sum_{i,j=1}^{M} \tilde{J}_{i,j}'\Bigg)\Bigg], \quad (11.56)$$

where $\tilde{J}_{i,j}' \sim \text{Poi}\left(\frac{cN^2}{2M^2(M+N)}\right)$ i.i.d. independent of everything else. Note that

$$\tilde{J}_{i,j} + \tilde{J}_{i,j}' \stackrel{d}{=} J_{i,j} + L_{i,j}, \quad (11.57)$$

if we take $J_{i,j} \sim \text{Poi}\left(\frac{c}{2(M+N)}\right)$ and $L_{i,j} \sim \text{Poi}\left(\frac{cN}{2M^2}\right)$ all independent of each other and everything else, because

$$\frac{c}{2M} + \frac{cN^2}{2M^2(M+N)} = \frac{c}{2(M+N)} + \frac{cN}{2M^2}. \quad (11.58)$$

Hence,

$$
\begin{aligned}
Mp_M(\beta) &\le \mathbb{E}\Big[\log \sum_{\tau\in[q]^M} \exp\Big(-\beta \sum_{i,j=1}^{M}(J_{i,j}+L_{i,j})\delta(\tau_i,\tau_j)\Big)\Big] + \beta M^2 \mathbb{E}[\tilde{J}'_{i,j}] \\
&= \mathbb{E}\Big[\log \sum_{\tau\in[q]^M} \mu_M^*(\tau)\exp\Big(-\beta \sum_{i,j=1}^{M} L_{i,j}\delta(\tau_i,\tau_j)\Big)\Big] + \beta M^2 \frac{cN^2}{2M^2(M+N)} \\
&= NG^{(2)}_{N,M}(\mu_M^*) + o_M(1), \qquad (11.59)
\end{aligned}
$$

with $G^{(2)}_{N,M}(\mu_M^*)$ as defined in (11.14). Combining the above gives

$$
\begin{aligned}
\frac{M+N}{N} p_{M+N} - \frac{M}{N} p_M &\ge G^{(1)}_{N,M}(\mu_M^*) - G^{(2)}_{N,M}(\mu_M^*) + o_M(1) \\
&= G_{N,M}(\mu_M^*) + o_M(1) \ge \inf_{\mathscr{L}} G_N(\beta,\mathscr{L}) + o_M(1). \qquad (11.60)
\end{aligned}
$$

Now, using (11.45),

$$
\begin{aligned}
p(\beta) &= \limsup_{N\to\infty} \liminf_{M\to\infty} \frac{1}{N}\big((M+N)p_{M+N}(\beta) - Mp_M(\beta)\big) \\
&\ge \limsup_{N\to\infty} \liminf_{M\to\infty} \big(\inf_{\mathscr{L}} G_N(\beta,\mathscr{L}) + o_M(1)\big) \\
&= \limsup_{N\to\infty} \inf_{\mathscr{L}} G_N(\beta,\mathscr{L}), \qquad (11.61)
\end{aligned}
$$

proving the theorem. □

11.4 Upper bounds on the pressure

In Proposition 11.7 we show that for every exchangeable law $\mathscr{L}$ we get an upper bound on the pressure. We now give two examples proving Theorem 11.3.

Proof of Theorem 11.3(a). Fix $m \in (0,1)$ and let Λ_m be the measure on $(0,\infty)$ with $\mathrm{d}\Lambda_m = mx^{-m-1}\mathrm{d}x$. Suppose that $(\xi_i)_{i\in\mathbb{N}}$ is a Poisson point process with intensity measure Λ_m. This has the property that, for positive i.i.d. random variables $\lambda_1, \lambda_2, \dots$ also independent of $(\xi_i)_{i\in\mathbb{N}}$, [32]

$$
\{\lambda_\alpha \xi_\alpha\}_{\alpha=1}^{\infty} \overset{\mathrm{d}}{=} \{\mathbb{E}[\lambda_\alpha^m]^{1/m}\xi_\alpha\}_{\alpha=1}^{\infty}. \qquad (11.62)
$$

Let

$$
\hat{\xi}_\alpha = \frac{\xi_\alpha}{\sum_{\alpha=1}^{\infty}\xi_\alpha}. \qquad (11.63)
$$

Then,

$$
\mathbb{E}\Big[\log \sum_{\alpha=1}^{\infty}\lambda_\alpha\hat{\xi}_\alpha\Big] = \mathbb{E}\Big[\log \sum_{\alpha=1}^{\infty}\mathbb{E}[\lambda_\alpha^m]^{1/m}\xi_\alpha\Big] - \mathbb{E}\Big[\log\sum_{\alpha=1}^{\infty}\xi_\alpha\Big] = \frac{1}{m}\log \mathbb{E}[\lambda_1^m]. \quad (11.64)
$$

We apply Proposition 11.7 with $(\mu_\alpha)_{\alpha\in\mathbb{N}} = (\hat{\xi}_\alpha)_{\alpha\in\mathbb{N}}$ and $\tau_{\alpha,k}$ i.i.d. uniformly on $[q]$. We start with $G_N^{(2)}$ which is easier. We can write this as

$$G_N^{(2)} = \frac{1}{N}\mathbb{E}\Big[\log\sum_{\alpha=1}^{\infty}\hat{\xi}_\alpha\lambda_\alpha\Big], \tag{11.65}$$

with

$$\lambda_\alpha = \exp\Big(-\beta\sum_{k=1}^{L}\delta(\tau_{\alpha,2k-1},\tau_{\alpha,2k})\Big). \tag{11.66}$$

Note that these λ_α are not independent of each other because they all depend on the same L. Therefore it is important to first condition on L, because then the λ_α are conditionally independent. Then,

$$\mathbb{E}[\lambda_1^m|L] = \prod_{k=1}^{L}\mathbb{E}\left[e^{-m\beta\delta(\tau_{1,2k-1},\tau_{1,2k})}\right] = \Big(1-\frac{1}{q}+\frac{1}{q}e^{-m\beta}\Big)^L. \tag{11.67}$$

Using this and (11.64),

$$\begin{aligned} G_N^{(2)} &= \frac{1}{N}\mathbb{E}\Big[\mathbb{E}\Big[\log\sum_{\alpha=1}^{\infty}\hat{\xi}_\alpha\lambda_\alpha\,\Big|\,L\Big]\Big] = \frac{1}{N}\mathbb{E}\left[\frac{L}{m}\log\Big(1-\frac{1}{q}+\frac{1}{q}e^{-m\beta}\Big)\right] \\ &= \frac{c}{2m}\log\Big(1-\frac{1-e^{-m\beta}}{q}\Big). \end{aligned} \tag{11.68}$$

We can rewrite $G_N^{(1)}$ as

$$G_N^{(1)} = \frac{1}{N}\mathbb{E}\Big[\log\sum_{\alpha=1}^{\infty}\mu_\alpha\sum_{\sigma\in[q]^N}\exp\Big(-\beta\sum_{i=1}^{N}\sum_{k=1}^{K_i}\delta(\sigma_i,\tau_{\alpha,i,k})\Big)\Big], \tag{11.69}$$

where each $K_i = |\{k\le K : I(k)=i\}|$ is an independent Poisson random variable with parameter c, because the sum of independent Poisson random variables is again Poisson. Once again the random variables $K_1,\dots,K_N$ are the same for all α's, not independent. Therefore, we condition on $K_1,\dots,K_N$. In this case

$$\lambda_\alpha = \sum_{\sigma\in[q]^N}\exp\Big(-\beta\sum_{i=1}^{N}\sum_{k=1}^{K_i}\delta(\sigma_i,\tau_{\alpha,i,k})\Big) = \prod_{i=1}^{N}\sum_{\sigma_i=1}^{q}\exp\Big(-\beta\sum_{k=1}^{K_i}\delta(\sigma_i,\tau_{\alpha,i,k})\Big). \tag{11.70}$$

The $\tau_{\alpha,i,k}$'s are all independent. In particular they are independent for different values of i. Therefore,

$$\mathbb{E}[\lambda_\alpha^m\mid K_1,\dots,K_N] = \prod_{i=1}^{N}\mathbb{E}\Big[\Big(\sum_{\sigma_i=1}^{q}e^{-\beta\sum_{k=1}^{K_i}\delta(\sigma_i,\tau_k)}\Big)^m\,\Big|\,K_i\Big]. \tag{11.71}$$

Using this,

$$G_N^{(1)} = \frac{1}{m}\mathbb{E}\Big[\log\mathbb{E}\Big[\Big(\sum_{\sigma=1}^{q}e^{-\beta\sum_{k=1}^{K_1}\delta(\sigma,\tau_k)}\Big)^m\,\Big|\,K_1\Big]\Big]. \tag{11.72}$$

We now take the limit $m \nearrow 1$. Then,

$$G_N^{(2)} = \frac{c}{2} \log\left(1 - \frac{1-e^{-\beta}}{q}\right). \tag{11.73}$$

For $G_N^{(1)}$ note that in the limit $m \nearrow 1$

$$\mathbb{E}\left[\sum_{\sigma=1}^{q} e^{-\beta \sum_{k=1}^{K_1} \delta(\sigma,\tau_k)} \,\middle|\, K_1\right] = \sum_{\sigma=1}^{q} \prod_{k=1}^{K_1} \mathbb{E}\left[e^{-\beta\delta(\sigma,\tau_k)}\right] = q\left(1 - \frac{1-e^{-\beta}}{q}\right)^{K_1}. \tag{11.74}$$

Hence,

$$G_N^{(1)} = \mathbb{E}\left[\log\left\{q\left(1 - \frac{1-e^{-\beta}}{q}\right)^{K_1}\right\}\right] = \log q + c\log\left(1 - \frac{1-e^{-\beta}}{q}\right). \tag{11.75}$$

Combining gives

$$p_N(\beta) \le G_N^{(1)} - G_N^{(2)} = \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) + \log q. \tag{11.76}$$

□

In the theorem above we used the Poisson point process $\{\xi_\alpha\}_{\alpha=1}^{\infty}$ to get the upper bound. To apply Proposition 11.7, however, it is sufficient that the index set $\{\alpha\}$ is countable. Therefore we may consider instead of a single countable index, a pair $\alpha = (\alpha_1, \alpha_2)$. We then construct the points ξ_α in a hierarchical way, i.e., for each α_1 we construct new point process $\{\xi_{\alpha_2}^{(2)}(\alpha_1)\}_{\alpha_2=1}^{\infty}$. We show that for appropriate choice of these point processes and a random probability measure on τ_α, we get the replica-symmetric solution.

Proof of Theorem 11.3(b). Let $\{\xi_{\alpha_1}^{(1)}\}_{\alpha_1=1}^{\infty}$ be a Poisson point process with intensity measure Λ_{m_1}, with $m_1 \in (0,1)$. For $\alpha_1 = 1, 2, \ldots$, we let $\{\xi_{\alpha_2}^{(2)}(\alpha_1)\}_{\alpha_2=1}^{\infty}$ be i.i.d. Poisson point processes with intensity measure Λ_{m_2}, also independent from the process $\{\xi_{\alpha_1}^{(1)}\}_{\alpha_1=1}^{\infty}$. Then we define

$$\xi_\alpha = \xi_{\alpha_1}^{(1)} \cdot \xi_{\alpha_2}^{(2)}(\alpha_1), \tag{11.77}$$

and its normalized version

$$\hat{\xi}_\alpha = \frac{\xi_\alpha}{\sum_{\alpha\in\mathbb{N}^2} \xi_\alpha}. \tag{11.78}$$

Note that

$$\left\{\sum_{\alpha_2} \xi_\alpha\right\}_{\alpha_1\in\mathbb{N}} \stackrel{d}{=} \left\{\mathbb{E}\left[\left(\sum_{\alpha_2} \xi_{\alpha_2}^{(2)}(\alpha_1)\right)^{m_1}\right]^{1/m_1} \xi_{\alpha_1}^{(1)}\right\}_{\alpha_1\in\mathbb{N}}, \tag{11.79}$$

because of (11.62). These expectations are only finite for $m_1 < m_2$ and hence we assume that this holds.

We now apply Proposition 11.7 with $(\mu_\alpha)_{\alpha\in\mathbb{N}^2} = (\hat{\xi}_\alpha)_{\alpha\in\mathbb{N}^2}$. Furthermore, we choose the distributions P_k such that they satisfy the assumptions stated in the theorem. Then, for all k, the $\tau_{\alpha,k}, \alpha \in \mathbb{N}$, are, given P_k, i.i.d. with distribution P_k. Then, using (11.62) twice, we can show that

$$G_N^{(1)} = \frac{1}{m_2} \mathbb{E} \log \mathbb{E} \left[\left(\sum_{\sigma=1}^{q} \exp\left(-\beta \sum_{k=1}^{K_1} \delta(\sigma, \tau_k) \right) \right)^{m_2} \Big| K_1, (P_k)_{k\geq 1} \right], \tag{11.80}$$

where we write τ_k for a generic $\tau_{\alpha,k}$. Similarly,

$$G_N^{(2)} = \frac{c}{2m_2} \mathbb{E} \log \mathbb{E} \left[e^{-\beta m_2 \delta(\tau_1, \tau_2)} \Big| P_1, P_2 \right]. \tag{11.81}$$

Taking the limit $m_2 \nearrow 1$ and explicitly taking the expectation over the τ's then proves the theorem. $\square$

12

Phase transition

We now prove that the system undergoes a phase transition if the connectivity constant c is large enough. Recall that

$$\beta_c = \inf\left\{\beta : p(\beta) \neq p^{\mathrm{HT}}(\beta)\right\}. \tag{12.1}$$

We show that the system undergoes a phase transition by giving bounds on β_c showing that $\beta_c \in (0,\infty)$. We use a constrained second-moment argument to show that there exists a $\beta^{2\mathrm{nd}} > 0$, such that $p(\beta) = p^{\mathrm{HT}}(\beta)$ for $\beta < \beta^{2\mathrm{nd}}$. Furthermore, we prove that for c large enough, there exists a $\beta^{\mathrm{en}} < \infty$ such that the entropy becomes negative for $\beta > \beta^{\mathrm{en}}$ if the high-temperature solution were to be true. This contradiction shows that $p(\beta) \neq p^{\mathrm{HT}}(\beta)$ for $\beta > \beta^{\mathrm{en}}$.

12.1 Results

We have the following bounds on the critical temperature:

Theorem 12.1 (Bounds on β_c)**.** *It holds that*

$$\beta^{2\mathrm{nd}} \leq \beta_c \leq \beta^{\mathrm{en}}, \tag{12.2}$$

with

$$\beta^{2\mathrm{nd}} = \begin{cases} -\log\left(1 - \frac{2}{1+\sqrt{c}}\right) & \text{when } q = 2, c > 1, \\ -\log\left(1 - \sqrt{\frac{2(q-1)\log(q-1)}{c}}\right) & \text{when } q > 2, c > 2(q-1)\log(q-1), \\ \infty & \text{otherwise.} \end{cases} \tag{12.3}$$

and

$$\beta^{\mathrm{en}} = \inf\left\{\beta : \log q + \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) < -\frac{\beta c}{2}\frac{e^{-\beta}}{q-1+e^{-\beta}}\right\}, \tag{12.4}$$

which is finite if

$$c > \frac{-2\log q}{\log(1-1/q)}. \tag{12.5}$$

We prove lower bound in Section 12.3 using a constrained second-moment method similar to that in [1]. The upper bound on β_c is proved in Section 12.4 by showing that the entropy becomes negative for $\beta > \beta^{\text{en}}$, which is not possible.

If $\beta^{\text{en}} = \infty$, then the condition in the infimum must not be true in the limit $\beta \to \infty$, i.e.,

$$\log q + \frac{c}{2} \log(1 - 1/q) \geq 0, \tag{12.6}$$

and hence,

$$c \leq \frac{-2 \log q}{\log(1 - 1/q)}. \tag{12.7}$$

We can thus conclude that $\beta^{\text{en}} < \infty$ if

$$c > \frac{-2 \log q}{\log(1 - 1/q)}. \tag{12.8}$$

12.2 Discussion

Critical temperatures in the physics literature. In [72], it is suggested that the high-temperature solution becomes unstable at $\beta^{\text{HT}} = -\log\left(1 - \frac{q}{1+\sqrt{c}}\right)$ and hence the high-temperature solution is not correct for $\beta > \beta^{\text{HT}}$. This can be made rigorous, as was shown in [28]. In combination with Theorem 12.1, this proves that, for $q = 2$, $\beta_c = \beta^{2\text{nd}} = \beta^{\text{HT}}$.

Note that β^{HT} is finite for $c > (q-1)^2$. For $q = 2, 3, 4$ it holds that $(q-1)^2 < \frac{-2\log q}{\log(1-1/q)}$ so that the instability of the high-temperature solution occurs before the entropy becomes negative, thus suggesting a second-order phase transition. When $q \geq 5$, however, the opposite is true and hence the system undergoes a first-order phase transition. In [72] it is suggested that the system has replica symmetry breaking when $q \geq 4$.

Graph coloring. The graph coloring problem has been studied using the antiferromagnetic Potts model in the physics literature, for example in [84, 85, 102]. There it is suggested that the system undergoes several phase transitions at $\beta = \infty$ for fixed q when the connectivity constant c increases. For very small c, almost all solutions to the graph coloring problem are in a giant cluster, where *"clusters are groups of nearby solutions that are in some sense disconnected from each other"* [102]. When c increases, at a certain point the giant cluster splits into an exponential number of small clusters. After this a condensation and a rigidity transition occur. Eventually, the graph becomes uncolorable.

This last phase transition from colorable to uncolorable is rigourously analyzed in [1], where it is shown that if q is the smallest integer satisfying $c < 2q \log q$, then the *chromatic number* of the graph, i.e., the least amount of colors necessary to color the graph properly, is equal to q or $q + 1$ with probability 1 in the thermodynamic limit.

12.3 Constrained second-moment method

In this section we prove the lower bound on β_c by showing that for $\beta < \beta^{2\text{nd}}$ it holds that $p(\beta) = p^{\text{HT}}(\beta)$. Note that $p(\beta) \leq p^{\text{HT}}(\beta)$ immediately follows from Theorem 11.3(a) and it thus remains to show that $p(\beta) \geq p^{\text{HT}}(\beta)$. For this we follow [1] where a *constrained* *second-moment method* was used for the graph coloring problem, which coincides with the $\beta = \infty$ case of our model.

We first restrict the partition function to *balanced* configurations, i.e., to configurations $\sigma \in [q]^{qN}$ such that the number of σ_i being equal to r is N for all $r \in [q]$. We denote this balanced partition function by $\widetilde{Z}_N(\beta)$, i.e.,

$$\widetilde{Z}_N(\beta) = \sum_{\substack{\sigma \in [q]^{qN} \\ \sigma \text{ balanced}}} e^{-\beta H(\sigma)}. \tag{12.9}$$

Furthermore, we condition on the event that $|J| \equiv \sum_{i,j=1}^{N} J_{i,j} \approx cN/2$. This conditioning on the typical event is important, because else the fluctuations of $|J|$ will cause the second-moment method to fail. The above constraining is allowed because of the following lemma:

Lemma 12.2 (Constraints on the pressure)**.** *For all $\beta, c > 0$,*

$$p(\beta) \geq \lim_{N \to \infty} \frac{1}{qN} \mathbb{E}\left[\log \widetilde{Z}_{qN}(\beta) \,\middle|\, |J| \in B_{qN}\right], \tag{12.10}$$

where

$$B_N = \left(\frac{cN}{2} - N^{2/3}, \frac{cN}{2} + N^{2/3}\right). \tag{12.11}$$

Proof. First of all, note that because of Theorem 10.1 we know that $p(\beta)$ exists and hence we can take any subsequence of $N \to \infty$. Hence,

$$p(\beta) = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}[\log Z_N(\beta)] = \lim_{N\to\infty} \frac{1}{qN}\mathbb{E}[\log Z_{qN}(\beta)] \geq \lim_{N\to\infty} \frac{1}{qN}\mathbb{E}[\log \widetilde{Z}_{qN}(\beta)], \tag{12.12}$$

because the sum in $\widetilde{Z}_{qN}(\beta)$ is over less configurations than the sum in $Z_{qN}(\beta)$. We can rewrite,

$$\begin{aligned}\frac{1}{qN}\mathbb{E}[\log \widetilde{Z}_{qN}] = \frac{1}{qN}&\mathbb{E}\left[\log \widetilde{Z}_{qN}(\beta) \,\middle|\, |J| \in B_{qN}\right] \mathbb{P}[|J| \in B_{qN}] \\ &+ \frac{1}{qN}\mathbb{E}\left[\log \widetilde{Z}_{qN}(\beta) \,\middle|\, |J| \notin B_{qN}\right] \mathbb{P}[|J| \notin B_{qN}].\end{aligned} \tag{12.13}$$

As in (10.32), we can show that a.s.,

$$\frac{1}{qN} \log \widetilde{Z}_{qN}(\beta) \leq \log q. \tag{12.14}$$

Note that $|J| \sim \text{Poi}\left(N^2 \frac{c}{2N}\right) = \text{Poi}\left(\frac{cN}{2}\right)$, so that by the Chebychev inequality,

$$\mathbb{P}[|J| \notin B_{qN}] = \mathbb{P}\left[\big||J| - cN/2\big| \geq N^{2/3}\right] \leq \frac{cN}{2N^{4/3}} = o(1). \tag{12.15}$$

Hence,

$$\frac{1}{qN}\mathbb{E}[\log \widetilde{Z}_{qN}] = \frac{1}{qN}\mathbb{E}\left[\log \widetilde{Z}_{qN}(\beta) \,\middle|\, |J| \in B_{qN}\right] + o(1), \tag{12.16}$$

and the lemma follows. □

We want to use a second-moment method to show that

$$\lim_{N\to\infty}\frac{1}{qN}\mathbb{E}\left[\log\widetilde{Z}_{qN}(\beta)\,\middle|\,|J|\in B_{qN}\right]=p^{\mathrm{HT}}(\beta), \tag{12.17}$$

and thus need to control the first and second moment of $\widetilde{Z}_{qN}(\beta)$ under the condition that $|J|\in B_{qN}$. For this, we start with the following technical lemma:

Lemma 12.3. *For $\sigma,\sigma^{(1)},\sigma^{(2)}\in[q]^N$,*

$$\mathbb{E}\left[e^{-\beta H(\sigma,J)}\,\middle|\,|J|\right]=\left(1-(1-e^{-\beta})\sum_{r_1=1}^{q}q_N^2(r_1)\right)^{|J|}, \tag{12.18}$$

and

$$\mathbb{E}\left[e^{-\beta(H(\sigma^{(1)},J)+H(\sigma^{(2)},J))}\,\middle|\,|J|\right]=\left(1-(1-e^{-\beta})\sum_{r_1=1}^{q}\left((q_N^{(1)}(r_1))^2+(q_N^{(2)}(r_1))^2\right)\right.$$
$$\left.+(1-e^{-\beta})^2\sum_{r_1,r_2=1}^{q}q_N^2(r_1,r_2)\right)^{|J|}, \tag{12.19}$$

where $q_N(r_1)$ and $q_N(r_1,r_2)$ are the overlaps introduced in (10.7).

Proof. We first compute the conditional probability

$$\mathbb{P}\left[\bigcap_{i,j=1}^{N}\{J_{i,j}=k_{i,j}\}\,\middle|\,|J|=k\right]=\frac{\prod_{i,j=1}^{N}e^{-c/(2N)}\frac{(c/(2N))^{k_{i,j}}}{k_{i,j}!}\mathbb{1}_{\{\sum_{i,j=1}^{N}k_{i,j}=k\}}}{e^{-cN/2}\frac{(cN/2)^k}{k!}}=\frac{1}{N^{2k}}\frac{k!}{\prod_{i,j=1}^{N}k_{i,j}!}. \tag{12.20}$$

Hence,

$$\begin{aligned}\mathbb{E}\left[e^{-\beta H(\sigma,J)}\,\middle|\,|J|=k\right]&=\frac{1}{N^{2k}}\sum_{|J|=k}\prod_{i,j=1}^{N}\left(e^{-\beta\delta(\sigma_i,\sigma_j)}\right)^{J_{i,j}}\frac{k!}{\prod_{i,j=1}^{N}k_{i,j}!}\\&=\left(\frac{1}{N^2}\sum_{i,j=1}^{N}e^{-\beta\delta(\sigma_i,\sigma_j)}\right)^{k},\end{aligned} \tag{12.21}$$

which follows from the multinomial formula. The first statement of the lemma now follows by again observing that $e^{-\beta\delta(\sigma_i,\sigma_j)}=1-(1-e^{-\beta})\delta(\sigma_i,\sigma_j)$. The proof of the second statement is similar.

□

For all balanced configurations $q_N(r)=1/q$. Using this and the previous lemma we can now control the constrained first moment of $\widetilde{Z}_{qN}(\beta)$:

Lemma 12.4 (Constrained first moment)**.** *For all $\beta,c>0$,*

$$\lim_{N\to\infty}\frac{1}{qN}\log\mathbb{E}\left[\widetilde{Z}_{qN}(\beta)\,\middle|\,|J|\in B_{qN}\right]=p^{\mathrm{HT}}(\beta). \tag{12.22}$$

Proof. For a balanced configuration $q_N(r_1) = 1/q$ for all $r_1 \in [q]$. Hence, it follows from Lemma 12.3 that

$$\mathbb{E}\left[\widetilde{Z}_{qN}(\beta) \,\Big|\, |J| \in B_{qN}\right] = Q_{qN}\mathbb{E}\left[\left(1 - \frac{1-e^{-\beta}}{q}\right)^{|J|} \,\Big|\, |J| \in B_{qN}\right], \tag{12.23}$$

where Q_{qN} denotes the number of balanced configurations in $[q]^{qN}$. Using $|J| \in B_{qN}$ and $1 - \frac{1-e^{-\beta}}{q} \leq 1$,

$$\left(1 - \frac{1-e^{-\beta}}{q}\right)^{\frac{cqN}{2}+(qN)^{2/3}} \leq \mathbb{E}\left[\left(1 - \frac{1-e^{-\beta}}{q}\right)^{|J|} \,\Big|\, |J| \in B_{qN}\right] \leq \left(1 - \frac{1-e^{-\beta}}{q}\right)^{\frac{cqN}{2}-(qN)^{2/3}}. \tag{12.24}$$

Hence,

$$\lim_{N\to\infty} \frac{1}{qN} \log \mathbb{E}\left[\left(1 - \frac{1-e^{-\beta}}{q}\right)^{|J|} \,\Big|\, |J| \in B_{qN}\right] = \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right). \tag{12.25}$$

For the number of balanced configurations

$$\lim_{N\to\infty} \frac{1}{qN} \log Q_{qN} = \lim_{N\to\infty} \frac{1}{qN} \log \frac{(qN)!}{(N!)^q} = \log q, \tag{12.26}$$

by Stirling's formula. Combining the above proves the lemma. □

Computing the second moment is more difficult. We first give a variational expression in the next lemma:

Lemma 12.5 (Constrained second moment). *For all $\beta, c > 0$,*

$$\lim_{N\to\infty} \frac{1}{qN} \log \mathbb{E}\left[\widetilde{Z}_{qN}(\beta)^2 \,\Big|\, |J| \in B_{qN}\right] = \sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)), \tag{12.27}$$

where

$$\phi^{(2)}(q(r_1,r_2)) = \frac{c}{2}\log\left(1 - \frac{2(1-e^{-\beta})}{q} + (1-e^{-\beta})^2 \sum_{r_1,r_2=1}^{q} q^2(r_1,r_2)\right) - \sum_{r_1,r_2=1}^{q} q(r_1,r_2)\log q(r_1,r_2), \tag{12.28}$$

and $\mathscr{B}([q]^2)$ denotes those overlaps for which the marginals are balanced, i.e.,

$$\sum_{r_1=1}^{q} q^2(r_1,r_2) = 1/q, \qquad \sum_{r_2=1}^{q} q^2(r_1,r_2) = 1/q, \qquad \text{and} \qquad q^2(r_1,r_2) \in [0,1]. \tag{12.29}$$

Proof. We can write, with $\mathbb{E}_s$ denoting the expectation over a configuration $s = (s^{(1)}, s^{(2)}) \in [q]^{2N}$ uniformly chosen from all configurations for which both $s^{(1)}$ and $s^{(2)}$ are balanced,

$$\widetilde{Z}_{qN}(\beta)^2 = \sum_{\substack{\sigma^{(1)},\sigma^{(2)}\in[q]^{qN} \\ \sigma^{(1)},\sigma^{(2)} \text{ balanced}}} e^{-\beta(H(\sigma^{(1)})+H(\sigma^{(2)}))} = Q_{qN}^2 \mathbb{E}_s\left[e^{-\beta(H(s^{(1)})+H(s^{(2)}))}\right]$$

$$= Q_{qN}^2 \mathbb{E}_s\left[\left(1 - \frac{2(1-e^{-\beta})}{q} + (1-e^{-\beta})^2 \sum_{r_1,r_2=1}^{q} q_N^2(r_1,r_2)\right)^{|J|}\right]. \tag{12.30}$$

Since $|J|/N \to c/2$ for $N \to \infty$ if $|J| \in B_N$, we can use Sanov's theorem and Varadhan's lemma to obtain [65]:

$$\lim_{N\to\infty} \frac{1}{qN} \log \mathbb{E}\left[\widetilde{Z}_{qN}(\beta)^2 \,\middle|\, |J| \in B_{qN} \right]$$
$$= 2\log q + \sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \left\{ \frac{c}{2} \log\left(1 - \frac{2(1-e^{-\beta})}{q} + (1-e^{-\beta})^2 \sum_{r_1,r_2=1}^{q} q^2(r_1,r_2)\right) \right.$$
$$\left. - \sum_{r_1,r_2=1}^{q} q(r_1,r_2)\log\left(q^2 q(r_1,r_2)\right) \right\} = \sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)), \tag{12.31}$$

where we also used (12.26). □

For $q = 2$ the above optimization problem is one-dimensional and can thus be analyzed explicitly:

Proposition 12.6 (The case $q = 2$)**.** *For $q = 2$ and $\beta < \beta^{2\mathrm{nd}}$,*

$$\sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)) = 2p^{\mathrm{HT}}(\beta). \tag{12.32}$$

Proof. By choosing $q(r_1,r_2) = 1/q^2$, we immediately get

$$\sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)) \geq 2p^{\mathrm{HT}}(\beta). \tag{12.33}$$

Note that this lower bound holds for all q.

For the upper bound, there exists a $\theta \in [0,1/2]$ such that $\theta = q(1,1) = q(2,2)$ and $1/2 - \theta = q(1,2) = q(2,1)$. We want to show that

$$\phi^{(2)}(\theta) - 2p^{\mathrm{HT}}(\beta) \leq 0. \tag{12.34}$$

We bound, using $\log(1+x) \leq x$,

$$\frac{c}{2}\log\left(\frac{1 - \frac{2(1-e^{-\beta})}{q} + (1-e^{-\beta})^2 \sum_{r_1,r_2=1}^{q} q^2(r_1,r_2)}{\left(1 - \frac{(1-e^{-\beta})}{q}\right)^2} \right)$$
$$= \frac{c}{2}\log\left(1 + \frac{4(1-e^{-\beta})^2}{(1+e^{-\beta})^2}(4\theta^2 - 2\theta + \frac{1}{4})\right) \leq \frac{c}{2}\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2 (4\theta-1)^2. \tag{12.35}$$

Furthermore,

$$-\sum_{r_1,r_2=1}^{q} q(r_1,r_2)\log\left(q^2 q(r_1,r_2)\right) - 2\log q = -2\theta\log(2\theta) - (1-2\theta)\log(1-2\theta). \tag{12.36}$$

Hence,

$$\phi^{(2)}(\theta) - 2p^{\mathrm{HT}}(\beta) \leq \frac{c}{2}\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2 (4\theta-1)^2 - 2\theta\log(2\theta) - (1-2\theta)\log(1-2\theta) \equiv \psi(\theta). \tag{12.37}$$

Computing the derivative gives

$$\psi'(\theta) = 4c\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2(4\theta-1) - 2\log(2\theta) + 2\log(1-2\theta), \tag{12.38}$$

so that $\psi'(1/4) = 0$. The second derivative equals

$$\psi''(\theta) = 16c\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2 - \frac{2}{\theta(1-2\theta)}. \tag{12.39}$$

This is maximal in $\theta = 1/4$, for which $\frac{2}{\theta(1-2\theta)} = 16$. Hence, $\psi(\theta)$ is a concave function for

$$c\left(\frac{1-e^{-\beta}}{1+e^{-\beta}}\right)^2 \leq 1, \tag{12.40}$$

so that in this case $\theta = 1/4$ is the unique maximizer of $\psi(\theta)$ and $\psi(1/4) = 0$. Hence, (12.40) gives the condition for which

$$\sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)) \leq 2p^{\mathrm{HT}}(\beta). \tag{12.41}$$

The condition (12.40) can be rewritten as

$$\beta \leq -\log\left(1 - \frac{2}{1+\sqrt{c}}\right). \tag{12.42}$$

□

For $q \geq 3$ the optimization problem is more difficult. We use the results in [1] to get bounds on the values β for which the above is also true:

Proposition 12.7 (The case $q \geq 3$). *For $q \geq 3$ and $\beta < \beta^{\mathrm{2nd}}$,*

$$\sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)) = 2p^{\mathrm{HT}}(\beta). \tag{12.43}$$

Proof. Again, it follows that

$$\sup_{q(r_1,r_2)\in\mathscr{B}([q]^2)} \phi^{(2)}(q(r_1,r_2)) \geq 2p^{\mathrm{HT}}(\beta), \tag{12.44}$$

by choosing $q(r_1,r_2) = 1/q^2$. For the upper bound, let $a_{r_1,r_2} = q \cdot q(r_1,r_2)$ so that

$$\phi^{(2)}(q(r_1,r_2)) = \frac{c}{2}\log\left(1 - \frac{2(1-e^{-\beta})}{q} + \frac{(1-e^{-\beta})^2}{q^2}\sum_{r_1,r_2=1}^{q} a_{r_1,r_2}^2\right) - \frac{1}{q}\sum_{r_1,r_2=1}^{q} a_{r_1,r_2}\log a_{r_1,r_2} + \log q. \tag{12.45}$$

Note that the matrix $A = (a_{r_1,r_2})_{r_1,r_2\in[q]}$ is a doubly stochastic matrix, i.e., all rows and columns add up to one. Let ρ be the 2-norm of A, i.e., $\rho = \sum_{r_1,r_2=1}^{q} a_{r_1,r_2}^2$. Then, [1, Theorem 9] tells us that

$$-\sum_{r_1,r_2=1}^{q} a_{r_1,r_2}\log a_{r_1,r_2} \le \max_m\left\{m\log q + (q-m)f(q,m,\rho)\ ;\ 0\le m\le \frac{q(q-\rho)}{q-1}\right\}, \tag{12.46}$$

for some explicit function f. It thus suffices to show that, for all $1\le\rho\le q$ and $0\le m\le \frac{q(q-\rho)}{q-1}$,

$$\frac{c}{2}\log\left(1-\frac{2(1-e^{-\beta})}{q}+\frac{(1-e^{-\beta})^2}{q^2}\rho\right)+\frac{m\log q}{q}+\frac{q-m}{q}f(q,m,\rho)+\log q \le 2\log q + c\log\left(1-\frac{1-e^{-\beta}}{q}\right), \tag{12.47}$$

which can be rewritten as

$$\frac{c}{2}\log\left(\frac{1-\frac{2(1-e^{-\beta})}{q}+\frac{(1-e^{-\beta})^2}{q^2}\rho}{\left(1-\frac{1-e^{-\beta}}{q}\right)^2}\right)\le\left(1-\frac{m}{q}\right)\left(\log q - f(q,m,\rho)\right). \tag{12.48}$$

Note that

$$\log\left(\frac{1-\frac{2(1-e^{-\beta})}{q}+\frac{(1-e^{-\beta})^2}{q^2}\rho}{\left(1-\frac{1-e^{-\beta}}{q}\right)^2}\right)=\log\left(1+\frac{(\rho-1)(1-e^{-\beta})^2}{(q-(1-e^{-\beta}))^2}\right)\le\frac{(\rho-1)(1-e^{-\beta})^2}{(q-1)^2}. \tag{12.49}$$

Hence, (12.47) holds if

$$\frac{c}{2}(1-e^{-\beta})^2\frac{\rho-1}{(q-1)^2}\le\left(1-\frac{m}{q}\right)\left(\log q - f(q,m,\rho)\right). \tag{12.50}$$

In [1, Theorem 7] it is proved that this holds for

$$\frac{c}{2}(1-e^{-\beta})^2\le(q-1)\log(q-1), \tag{12.51}$$

and hence we need that

$$\beta\le-\log\left(1-\sqrt{\frac{2(q-1)\log(q-1)}{c}}\right). \tag{12.52}$$

□

The fact that $\beta_c \ge \beta^{2\text{nd}}$ now follows from a standard second-moment method argument.

12.4 Bounds from entropy positivity

The entropy for the Potts model is the following generalization of (3.11):

Definition 12.8. *The random* entropy *density equals*

$$S_N = -\frac{1}{N}\sum_{\sigma\in[q]^N}\mu(\sigma)\log\mu(\sigma). \tag{12.53}$$

Then we have the following lemma.

Lemma 12.9 (Entropy positivity)**.** *For $\beta \geq 0, \delta > 0$ and $N \in \mathbb{N}$, a.s.,*

$$0 \leq S_N(\beta) \leq -\beta\frac{\psi_N(\beta)-\psi_N(\beta-\delta)}{\delta} + \psi_N(\beta). \tag{12.54}$$

Proof. Note that $\mu(\sigma) \leq 1$, a.s., and hence $\log(\mu(\sigma)) \leq 0$. From this the lower bound immediately follows.

For the upper bound, note that

$$\frac{\partial}{\partial\beta}\psi_N(\beta) = \frac{\partial}{\partial\beta}\frac{1}{N}\log Z_N(\beta) = -\frac{1}{N}\frac{\sum_{\sigma\in[q]^N}H(\sigma)e^{-\beta H(\sigma)}}{Z_N(\beta)} = -\frac{1}{N}\langle H(\sigma)\rangle. \tag{12.55}$$

Combined with

$$\log(\mu(\sigma)) = \log\left(\frac{e^{-\beta H(\sigma)}}{Z_N(\beta)}\right) = -\beta H(\sigma) - \log Z_N(\beta), \tag{12.56}$$

this gives

$$\begin{aligned} S_N(\beta) &= \frac{1}{N}\sum_{\sigma\in[q]^N}\mu(\sigma)\left(\beta H(\sigma)+\log Z_N\right) = \frac{\beta}{N}\langle H(\sigma)\rangle + \psi_N(\beta) \\ &= -\beta\frac{\partial}{\partial\beta}\psi_N(\beta) + \psi_N(\beta). \end{aligned} \tag{12.57}$$

Furthermore, $\psi_N(\beta)$ is a convex function of β, because

$$\begin{aligned} \frac{\partial^2}{\partial\beta^2}\psi_N(\beta) &= \frac{1}{N}\frac{\sum_{\sigma\in[q]^N}(H(\sigma))^2e^{-\beta H(\sigma)}}{Z_N(\beta)} + \frac{1}{N}\left(\frac{\sum_{\sigma\in[q]^N}H(\sigma)e^{-\beta H(\sigma)}}{Z_N(\beta)}\right)^2 \\ &= \left\langle (H(\sigma))^2\right\rangle - \langle H(\sigma)\rangle^2 \geq 0. \end{aligned} \tag{12.58}$$

Hence,

$$\frac{\partial}{\partial\beta}\psi_N(\beta) \geq \frac{\psi_N(\beta)-\psi_N(\beta-\delta)}{\delta}. \tag{12.59}$$

Combining this with (12.57) proves the lemma. □

We can use this lemma to give an upper bound on β_c.

Proposition 12.10. *For $q > 1$,*

$$\beta_c \leq \beta^{\text{en}} \equiv \inf\left\{\beta : \log q + \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) < -\frac{\beta c}{2}\frac{e^{-\beta}}{q-1+e^{-\beta}}\right\}. \tag{12.60}$$

Proof. Suppose that $\beta_c > \beta^{\text{en}}$. Then, there is a β such that $p(\beta) = p^{\text{HT}}(\beta)$ and $p(\beta-\delta) = p^{\text{HT}}(\beta-\delta)$, and

$$\log q + \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) < -\frac{\beta c}{2}\frac{e^{-\beta}}{q-1+e^{-\beta}}. \tag{12.61}$$

But we know from Lemma 12.9 that, a.s.,

$$-\beta\frac{\psi_N(\beta) - \psi_N(\beta-\delta)}{\delta} + \psi_N(\beta) \geq 0, \tag{12.62}$$

and hence this also holds in expectation:

$$-\beta\frac{p_N(\beta) - p_N(\beta-\delta)}{\delta} + p_N(\beta) \geq 0. \tag{12.63}$$

Taking the limit of $N \to \infty$ then gives

$$-\beta\frac{p^{\text{HT}}(\beta) - p^{\text{HT}}(\beta-\delta)}{\delta} + p^{\text{HT}}(\beta) \geq 0, \tag{12.64}$$

because of our assumption on β. Since $p^{\text{HT}}(\beta)$ is analytic, we can take the limit $\delta \searrow 0$, so that

$$-\beta\frac{\partial}{\partial\beta}p^{\text{HT}}(\beta) + p^{\text{HT}}(\beta) \geq 0. \tag{12.65}$$

Using $p^{\text{HT}}(\beta) = \log q + \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right)$ then gives,

$$\log q + \frac{c}{2}\log\left(1 - \frac{1-e^{-\beta}}{q}\right) \geq -\frac{\beta c}{2}\frac{e^{-\beta}}{q-1+e^{-\beta}}, \tag{12.66}$$

which is in contradiction with (12.61). □

Bibliography

[1] D. Achlioptas and A. Naor. The two possible values of the chromatic number of a random graph. *Annals of Mathematics*, **162**:1335–1351, (2005).

[2] M. Aizenman. Geometric analysis of ϕ^4 fields and Ising models: parts I and II. *Communications in Mathematical Physics*, **86**:1–48, (1982).

[3] M. Aizenman and R. Fernández. On the critical behavior of the magnetization in high-dimensional Ising models. *Journal of Statistical Physics*, **44**:393–454, (1986).

[4] M. Aizenman and R. Graham. On the renormalized coupling constant and the susceptibility in ϕ_4^4 field theory and the Ising model in four dimensions. *Nuclear Physics B*, **225**(2):261–288, (1983).

[5] M. Aizenman, R. Sims and S.L. Starr. Extended variational principle for the Sherrington-Kirkpatrick spin-glass model. *Physical Review B*, **68**:214403, (2003).

[6] D. Alberici. Thermodynamical limit in diluted spin systems. *Laurea Thesis*, Università di Bologna, (2010).

[7] R. Albert, I. Albert and G.L. Nakarado. Structural vulnerability of the North American power grid. *Physical Review E*, **69**:025103, (2004).

[8] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, **74**:47–97, (2002).

[9] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, **25**(2):812-854, (1997).

[10] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, **12**:1454–1508, (2007).

[11] D. Aldous and J.M. Steele. The objective method: probabilistic combinatorial optimization. In: *Probability on Discrete Structures*, pages 1–72, Springer, Berlin, (2004).

[12] L. Backstrom, P. Boldi, M. Rosa, J. Ugander and S. Vigna. Four degrees of separation. Preprint, arXiv:1111.4570, (2011).

[13] F. Ball, D. Sirl and P. Trapman. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Advances in Applied Probability*, **41**(3):765–796, (2009).

[14] F. Ball, D. Sirl and P. Trapman. Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Mathematical Biosciences*, **224**(2):53–73, (2010).

[15] A. Basak and A. Dembo. Ferromagnetic Ising measures on large locally tree-like graphs. Preprint, arXiv:1205.4749, (2012).

[16] R.J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press, London, (1982).

[17] I. Benjamini, R. Lyons and O. Schramm. Unimodular random trees. Preprint, arXiv:1207.1752, (2012).

[18] N. Berger, C. Borgs, J.T. Chayes, A. Saberi. A weak local limit for preferential attachment graphs. Preprint, (2009).

[19] S. Bhamidi, R. van der Hofstad and G. Hooghiemstra. Extreme value theory, Poisson-Dirichlet distributions and first passage percolation on random networks. *Advances in Applied Probability*, **42**:706–738, (2010).

[20] S. Bhamidi, R. van der Hofstad and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *The Annals of Applied Probability*, **20**(5):1907–1965, (2010).

[21] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York, (1999).

[22] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, **1**:311–316, (1980).

[23] B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge, (2001).

[24] B. Bollobás, S. Janson and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, **31**:3–122, (2007).

[25] S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *Annals of Probability*, **37**(6):2332–2356, (2009).

[26] J. Cohen and M. Enserink. As swine flu circles globe, scientists grapple with basic questions. *Science*, **324**:572–573, (2009).

[27] P. Contucci, S. Dommers, C. Giardinà and S. Starr. Antiferromagnetic Potts model on the Erdős-Rényi random graph. Preprint, arXiv:1106.4714v3, (2011).

[28] P. Contucci, S. Dommers, C. Giardinà and S. Starr. Antiferromagnetic Potts model on the Erdős-Rényi random graph. Preprint, arXiv:1106.4714v4, (2012). To appear in *Communications in Mathematical Physics*.

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

[30] P. Contucci and S. Ghirlanda. Modeling society with statistical mechanics: an application to cultural contact and immigration. *Quality & Quantity*, **41**(4):569–578, (2007).

[31] P. Contucci and C. Giardinà. Mathematics and social sciences: a statistical mechanics approach to immigration. *ERCIM News*, **73**:34–35, (2008).

[32] P. Contucci and C. Giardinà. *Perspectives on Spin Glasses.* Cambridge University Press, Cambridge, (2012).

[33] M. Deijfen, R. van der Hofstad and G. Hooghiemstra. Scale-free percolation. Preprint, arXiv:1103.0208, (2011).

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

[35] 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).

[36] A. Dembo, A. Montanari, A. Sly and N. Sun. The replica symmetric solution for Potts models on d-regular graphs. Preprint, arXiv:1207.5500, (2012).

[37] A. Dembo, A. Montanari and N. Sun. Factor models on locally tree-like graphs. Preprint, arXiv:1110.4821, (2011).

[38] 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).

[39] S. Dommers, C. Giardinà and R. van der Hofstad. Ising critical exponents on random trees and graphs. Preprint, arXiv:1211.3005, (2012).

[40] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes. Ising models on networks with an arbitrary distribution of connections. *Physical Review E*, **66**:016104, (2002).

[41] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes. Potts model on complex networks. *The European Physics Journal B*, **38**:177–182, (2004).

[42] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, **80**(4):1275–1335, (2008).

[43] J.H. Dshalalow. *Real Analysis: An Introduction to the Theory of Real Functions and Integration.* Chapman & Hall/CRC, London, (2001).

[44] A.C.D. van Enter, V.N. Ermolaev, G. Iacobelli and C. Külske. Gibbs-non-Gibbs properties for evolving Ising models on trees. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **48**(3):774–791, (2012).

[45] A.C.D. van Enter, R. Fernández, F. den Hollander and F. Redig. Possible loss and recovery of Gibbsianness during stochastic evolution of Gibbs measures. *Communications in Mathematical Physics*, **226**(1):101–130, (2002).

[46] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Science*, **5**:17–61, (1960).

[47] H. van den Esker, R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with infinite mean degrees. *Extremes*, **8**:111–140, (2006).

[48] W. Evans, C. Kenyon, Y. Peres, L.J. Schulman. Broadcasting on trees and the Ising model. *The Annals of Applied Probability*, **10**(2):410–433, (2000).

[49] Facebook. *One Billion Fact Sheet*. Available from: http://newsroom.fb.com/download-media/4227, (2012).

[50] C. Faloutsos, M. Faloutsos and P. Faloutsos. On power-law relationships of the Internet topology. *Computer Communication Review*, **29**:251–262, (1999).

[51] D.A. Fell and A. Wagner. The small world of metabolism. Nature Biotechnology, **18**:1121–1122, (2000).

[52] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II, 2nd Edition*. John Wiley and Sons, New York, (1971).

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

[54] S. Franz and M. Leone. Replica bounds for optimization problems and diluted spin systems. *Journal of Statistical Physics*, **111**:535–564, (2003).

[55] K. Frenken. A complexity approach to innovation networks. The case of the aircraft industry (1909–1997). *Research Policy*, **29**:257–272, (2000).

[56] E.N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, **30**(4):1141–1144, (1959).

[57] R.B. Griffiths. Correlations in Ising ferromagnets. II. External magnetic fields. *Journal of Mathematical Physics*, **8**(3):484–489, (1967).

[58] R.B. Griffiths, C.A. Hurst and S. Sherman. Concavity of magnetization of an Ising ferromagnet in a positive external field. *Journal of Mathematical Physics*, **11**(3):790–795, (1970).

[59] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in Mathematical Physics*, **233**:1–12, (2003).

[60] F. Guerra and F.L. Toninelli. The thermodynamic limit in mean field spin glass models. *Communications in Mathematical Physics*, **230**:71–79, (2002).

[61] F. Guerra and F.L. Toninelli. The high-temperature region of the Viana-Bray diltued spin glass model. *Journal of Statistical Physics*, **115**(1–2), 531–555, (2004).

[62] R. Guimerà, S. Mossa, A. Turtschi and L.A.N. Amaral. The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences*, **102**(22):7794–7799, (2005).

[63] R. van der Hofstad. Random graphs and complex networks. *Lecture notes*, Preprint, (2012).

[64] R. van der Hofstad and G. Hooghiemstra. Universality for distances in power-law random graphs. *Journal of Mathematical Physics*, **49**:125209, (2008).

[65] F. den Hollander. *Large Deviations*. American Mathematical Society, Providence, (2000).

[66] IMS Research. Internet Connected Devices Approaching 10 Billion, to exceed 28 Billion by 2020. *Press Release*, available from: http://imsresearch.com/, (2012).

[67] K. Itô (ed.). *Encyclopedic Dictionary of Mathematics, Second Edition.* The MIT Press, Cambridge, (1993).

[68] S. Janson and M.J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, **34**(2):197–216, (2008).

[69] S. Janson, T. Łuczak and A. Ruciński. *Random Graphs*. John Wiley & Sons, New York, (2000).

[70] H. Jeong, S.P. Mason, A.-L. Barabási and Z.N. Oltvai. Lethality and centrality in protein networks. *Nature*, **411**:41–42, (2001).

[71] D.G. Kelly and S. Sherman. General Griffiths' inequalities on correlations in Ising ferromagnets. *Journal of Mathematical Physics*, **9**(3):466–484, (1968).

[72] F. Krząkała and L. Zdeborová. Potts glass on random graphs. *Europhysics Letters*, **81**:57005, (2008).

[73] T. Kuran. The East European revolution of 1989: is it surprising that we were surprised? *The American Economic Review*, **81**(2):121–125, (1991).

[74] M. Leone, A. Vázquez, A. Vespignani and R. Zecchina. Ferromagnetic ordering in graphs with arbitrary degree distribution. *The European Physical Journal B*, **28**:191–197, (2002).

[75] D. Liben-Nowell and J. Kleinberg. Tracing information flow on a global scale using Internet chain-letter data. *Proceedings of the National Academy of Sciences*,**105**(12):4633–4638, (2008).

[76] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Communications in Mathematical Physics*, **125**:337–353, (1989).

[77] R. Lyons. Random walks and percolation on trees. *The Annals of Probability*, **18**(3):931–958, (1990).

[78] R. Lyons, R. Pemantle and Y. Peres. Ergodic theory on Galton-Watson trees: speed of the random walk and dimension of harmonic measure. *Ergodic Theory and Dynamical Systems*, **15**(3):593–619, (1995).

[79] M. Mézard and A. Montanari. Reconstruction on trees and spin glass transition. *Journal of Statistical Physics*, **124**(6):1317–1350, (2006).

[80] D.S. Mitrinović. *Analytic Inequalities*. Springer-Verlag, Berlin, (1970).

[81] A. Montanari, E. Mossel and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probability Theory and Related Fields*, **152**:31–51, (2012).

[82] A. Montanari and A. Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, **107**(47):20196–20201, (2010).

[83] T. Mountford, J.-C. Mourrat, D. Valesin and Q. Yao. Exponential extinction time of the contact process on finite graphs. Preprint, arXiv:1203.2972, (2012).

[84] J. van Mourik and D. Saad. Random graph coloring: statistical physics approach. *Physical Review E*, **66**:056120, (2002).

[85] R. Mulet, A. Pagnani, M. Weigt and R. Zecchina. Coloring random graphs. *Physical Review Letters*, **89**(26):268701, (2002).

[86] E.J. Neves and R.H. Schonmann. Critical droplets and metastability for a Glauber dynamics at very low temperatures. *Communications in Mathematical Physics*, **137**:209–230, (1991).

[87] M.E.J. Newman. The structure and function of complex networks. *SIAM Review*, **45**(2):167–256, (2003).

[88] M.E.J. Newman. *Networks: An Introduction*. Oxford University Press, Oxford, (2010).

[89] M. Niss. History of the Lenz–Ising model 1920–1950: from ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, **59**(3):267–318, (2005).

[90] M. Niss. History of the Lenz–Ising Model 1950–1965: from irrelevance to relevance. *Archive for History of Exact Sciences*, **63**(3):243–287, (2009).

[91] M. Niss. History of the Lenz–Ising Model 1965–1971: the role of a simple model in understanding critical phenomena. *Archive for History of Exact Sciences*, **65**(6):625–658, (2011).

[92] G. Parisi. A sequence of approximated solutions to the S-K model for spin glasses. *Journal of Physics A*, **13**:L115–L121, (1980).

[93] B. Pittel. On tree census and the giant component in sparse random graphs. *Random Structures & Algorithms*, **1**(3):311–342, (1990).

[94] G. Pólya and G. Szegő. *Problems and Theorems in Analysis: Series, Integral Calculus, Theory of Functions*. Springer, Berlin, (1978).

[95] S. Ruan, W. Wang and S.A. Levin. The effect of global travel on the spread of SARS. *Mathematical Biosciences and Engineering*, **3**(1):205–218, (2006).

[96] L. De Sanctis and F. Guerra. Mean field dilute ferromagnet: high temperature and zero temperature behavior. *Journal of Statistical Physics*, **132**:759–785, (2008).

[97] C.M. Schneider, L. de Arcangelis and H.J. Herrmann. Scale free networks by preferential depletion. Preprint, arXiv:1103.1396, (2011).

[98] S.H. Strogatz. Exploring complex networks. *Nature*, **410**:268–276, (2001).

[99] M. Talagrand. *Spin Glasses: A Challenge for Mathematicians*. Springer, Berlin, (2003).

[100] M. Talagrand. The Parisi formula. *Annals of Mathematics*, **163**:221–263, (2006).

[101] J.G. White, E. Southgate, J.N. Thomson and S. Brenner. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philosophical Transactions of the Royal Society B*, **314**:1–340, (1986).

[102] L. Zdeborová and F. Krząkała. Phase transitions in the coloring of random graphs. *Physical Review E*, **76**(3):031131, (2007).

INDEX

Curriculum Vitae

Sander Dommers was born on August 22, 1984 in Dordrecht, the Netherlands. In 2002, he finished the gymnasium at the Stedelijk Dalton Lyceum in Dordrecht and started a joint program in Computer Science and Engineering and in Industrial and Applied Mathematics at Eindhoven University of Technology. After finishing the first year (propedeuse) of these two programs, he continued his Bachelor and Master studies in Industrial and Applied Mathematics.

Sander finished his Bachelor with a thesis on 'The machine repair model' under supervision of Onno Boxma. His Master specialization was Statistics, Probability and Operations Research which he finished with a thesis on 'Distances in power-law random graphs' under supervision of Remco van der Hofstad.

During his studies, Sander was a member of the Eindhoven Student Dance Association Footloose, of which he was also president in the board in the academic year 2005–2006. He also played the clarinet in the Eindhoven Student Symphony Orchestra Ensuite, part of ESMG Quadrivium.

After graduating in 2009, Sander started as a PhD student under supervision of Remco van der Hofstad and Cristian Giardinà in the Stochastics section of the Department of Mathematics and Computer Science at Eindhoven University of Technology. He also spent a three and a four week period at the University of Modena and Reggio Emilia. The research carried out as a PhD student resulted in this thesis.

Besides doing research and teaching, Sander was a member of the Departmental Council from 2009 to 2012. He also continued to play the clarinet in Ensuite and was the treasurer in the board of SCOrE, which facilitates a meeting place for members of cultural student associations in Eindhoven. To show high school students how fun mathematics can be, he helped organizing and supervising mathematics summer camps of Vierkant voor Wiskunde and was in the problem committee of the Dutch Mathematical Kangaroo competition.

In 2013, Sander started as a postdoctoral research fellow at the University of Bologna.

LIST OF PUBLICATIONS

Journal articles

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).

S. Dommers, R. van der Hofstad and G. Hooghiemstra. Diameters in preferential attachment models. *Journal of Statistical Physics*, **139**(1):72–107, (2010).

Preprints

S. Dommers, C. Giardinà and R. van der Hofstad. Ising critical exponents on power-law random graphs. Preprint, arXiv:1211.3005, (2012).

P. Contucci, S. Dommers, C. Giardinà and S. Starr. Antiferromagnetic Potts model on the Erdős-Rényi random graph. Preprint, arXiv:1106.4714, (2012). To appear in *Communications in Mathematical Physics*.

Report

S. Cardon, S. Dommers, C. Eksin, R. Sitters, A. Stougie and L. Stougie. A PTAS for the Multiple Depot Vehicle Routing Problem. *SPOR Report No. 2008-03, Eindhoven University of Technology*, (2008).

Acknowledgements

This thesis is the result of four years of research. Finishing this thesis would not have been possible without the help and support of many others. I therefore would like to thank several people.

First, I would like to express my sincere gratitude to my promotor and supervisor Remco van der Hofstad. I was very lucky to have such a briljant supervisor. I would like to thank him for infecting me with his enthusiasm for random graphs and for teaching me how to be a better mathematician. I really enjoyed working together.

Second, I would like to thank my copromotor and second supervisor Cristian Giardinà. We had many interesting discussions and I admire his patience while explaining things to me. I am also grateful for his hospitality during my visits to Reggio Emilia and Modena, making my stays there very pleasant.

I am also grateful that I had the opportunity to work with Pierluigi Contucci and Shannon Starr and would like to thank them for a pleasant collaboration.

I am also indebted to Sem Borst, Aernout van Enter, Roberto Fernández, Koen Frenken and Francesco Guerra for willing to be a member of my PhD defense committee, reading my thesis and giving useful comments.

I would also like to thank all my colleagues in the Stochastics section and at Eurandom for the great atmosphere. I really enjoyed the many coffee breaks, playing foosball and pool, going to the pubquiz and the numerous other social get-togethers we had. I also had a great time in the offices I worked in with Robert and Tim, and for shorter periods also with Maren, Richard and Bartek.

I also thank my friends and family for their support. Finally, I would like to thank Relinde for always being there for me and cheering me up on many occasions. It was also very useful that she finished her PhD before me, so that I didn't have to find out many things related to the defense myself.