The Annals of Applied Probability
2023, Vol. 33, No. 3, 1961–2003
https://doi.org/10.1214/22-AAP1857

FLUCTUATIONS IN MEAN-FIELD ISING MODELS

BY NABARUN DEB[a] AND SUMIT MUKHERJEE[b]

Department of Statistics, Columbia University, [a]*nd2560@columbia.edu,* [b]*sm3949@columbia.edu*

In this paper, we study the fluctuations of the average magnetization in an Ising model on an approximately d_N regular graph G_N on N vertices. In particular, if G_N satisfies a "spectral gap" condition, we show that whenever $d_N \gg \sqrt{N}$, the fluctuations are universal and the same as that of the Curie–Weiss model in the entire ferromagnetic parameter regime. We give a counterexample to demonstrate that the condition $d_N \gg \sqrt{N}$ is tight, in the sense that the limiting distribution changes if $d_N \sim \sqrt{N}$ except in the high temperature regime. By refining our argument, we extend universality in the high temperature regime up to $d_N \gg N^{1/3}$. Our results include universal fluctuations of the average magnetization in Ising models on regular graphs, Erdős–Rényi graphs (directed and undirected), stochastic block models, and sparse regular graphons. In fact, our results apply to general matrices with nonnegative entries, including Ising models on a Wigner matrix, and the block spin Ising model. As a by-product of our proof technique, we obtain Berry–Esseen bounds for these fluctuations, exponential concentration for the average of spins, tight error bounds for the mean-field approximation of the partition function, and tail bounds for various statistics of interest.

1. Introduction. The Ising model is a discrete Markov random field which was initially introduced as a mathematical model of ferromagnetism in Statistical Physics, and has received extensive attention in Probability (cf. [1, 10, 14, 17, 23, 24, 29, 32, 39] and references therein) and Statistics (cf. [5, 16, 21, 25, 34, 35, 38] and references therein). The model can be described by the following probability mass function in $\boldsymbol{\sigma} := (\sigma_1, \dots, \sigma_N) \in \{-1, 1\}^N$:

$$\mathbb{P}(\boldsymbol{\sigma}) := \frac{1}{Z_N(\beta, B)} \exp\left(\frac{\beta}{2}\boldsymbol{\sigma}^\top A_N \boldsymbol{\sigma} + B\sum_{i=1}^N \sigma_i\right). \tag{1.1}$$

Here A_N is a symmetric $N \times N$ matrix with nonnegative entries, and has zeroes on its diagonal, and $\beta > 0$ and $B \in \mathbb{R}$ are scalar parameters often referred to in the Statistical Physics literature as *inverse temperature* and *external magnetic field* respectively. The factor $Z_N(\beta, B)$ is the normalizing constant/partition function of the model. The most common choice of the coupling matrix A_N is the adjacency matrix of a graph G_N on N vertices, scaled by the average degree $\overline{d}_N := \frac{1}{N}\sum_{i,j=1}^N G_N(i, j)$. Here and throughout the rest of the paper, we use the notation G_N to denote both a graph and its adjacency matrix. A pivotal quantity of interest which has attracted extensive attention in the literature is the average sum of spins/magnetization density, defined by

$$\overline{\boldsymbol{\sigma}} := \frac{\sum_{i=1}^N \sigma_i}{N}.$$

The fluctuations for $\overline{\boldsymbol{\sigma}}$ are mostly known for very few choices of the graph G_N, including the complete graph (see, e.g., [14, 19, 21]), the directed Erdős–Rényi graph (see [26]), and sparse Erdős–Rényi graphs (see [24]). In this paper, we focus on studying fluctuations of $\overline{\boldsymbol{\sigma}}$, when

Received January 2021; revised July 2021.
MSC2020 subject classifications. Primary 82B20; secondary 82B26.
Key words and phrases. Berry–Esseen bound, Ising model, regular graphs, mean-field, partition function.

A_N is the scaled adjacency matrix of an approximately regular graph G_N. The motivation for this work is the recent paper [4], where the authors show universal asymptotics of the partition function $Z_N(\beta, B)$ on *any* sequence of approximately regular graphs with diverging average degree, which is governed by the mean-field prediction formula. In particular, it follows from [4], Theorem 2.1, that the mean-field prediction formula is asymptotically universal in the sense that

$$\frac{1}{N}\log Z_N(\beta, B) \stackrel{N\to\infty}{\longrightarrow} \sup_{\mathbf{x}\in[-1,1]}\left\{\frac{\beta x^2}{2} + Bx - \frac{1+x}{2}\log\frac{1+x}{2} - \frac{1-x}{2}\log\frac{1-x}{2}\right\}$$

for any sequence of approximately d_N regular graphs G_N with $d_N \to \infty$. A natural follow up question is to what extent this universality extends to other properties of such "mean-field" Ising models. In this paper we try to address this question by studying the universal behavior of the statistic $\overline{\boldsymbol{\sigma}}$.

Our main results (see Theorems 1.1–1.4) show that $\overline{\boldsymbol{\sigma}}$ exhibits universal fluctuations for a large class of "approximately regular" graphs with $\overline{d}_N$ diverging "fast enough", across all parameter regimes for (β, B). Our proof techniques yield tight error bounds for the Mean-Field approximation of the partition function (see Theorem 1.5), exponential concentration for the average of spins (see Theorem 1.6 and Corollary 1.1), and tail bounds for various statistics of interest (see Lemmas 2.1–2.3). One of our main contributions is that our results hold even if the minimum and maximum eigenvalue of A_N have the same magnitude asymptotically (see Remark 2.1). Our assumptions on A_N are thus significantly weaker than the expander type assumptions prevalent in the literature. For ease of exposition, in Section 1.2, we outline our proof techniques in the special case where G_N is regular.

1.1. *Main results.* We begin with a definition which partitions the parameter set $\{(\beta, B) : \beta > 0, B \in \mathbb{R}\}$ into different domains.

DEFINITION 1.1. Let

$$\Theta_{11} := \{(\beta, 0) : 0 < \beta < 1\}, \qquad \Theta_{12} := \{(\beta, B) : \beta > 0, B \neq 0\},$$
$$\Theta_2 := \{(\beta, 0) : \beta > 1\}, \qquad \Theta_3 := (1, 0).$$

Finally, let $\Theta_1 := \Theta_{11} \cup \Theta_{12}$. We will refer to Θ_1 as the uniqueness regime, Θ_2 as the nonuniqueness regime, and Θ_3 as the critical point. The names of the different regimes are motivated by the next lemma, the proof of which follows from simple calculus (see, e.g., [17], page 144, Section 1.1.3).

LEMMA 1.1. *Consider the fixed point equation*

$$\phi(x) = 0 \quad \text{where } \phi(x) := x - \tanh(\beta x + B). \tag{1.2}$$

(a) *If* $(\beta, B) \in \Theta_{11}$, *then* (1.2) *has a unique solution at* $t = 0$, *and* $\phi'(0) > 0$.

(b) *If* $(\beta, B) \in \Theta_{12}$, *then* (1.2) *has a unique root* t *with the same sign as that of* B, *and* $\phi'(t) > 0$.

(c) *If* $(\beta, B) \in \Theta_2$, *then* (1.2) *has two nonzero roots* $\pm t$ *of this equation, where* $t > 0$, *and* $\phi'(\pm t) > 0$.

(d) *If* $(\beta, B) \in \Theta_3$, *then* (1.2) *has a unique solution at* $t = 0$, *and* $\phi'(0) = 0$.

We will use t as defined in the above lemma throughout the paper, noting that t does depend on (β, B). The following result summarizes the fluctuations of $\overline{\boldsymbol{\sigma}}$ in the Curie–Weiss model (see [21]), which is the Ising model on the complete graph.

LEMMA 1.2. *Suppose $\boldsymbol{\sigma}$ is a random vector from the Curie–Weiss model $\mathbb{P}^{\mathrm{CW}}$ with p.m.f.*

$$\mathbb{P}^{\mathrm{CW}}(\boldsymbol{\sigma}) = \frac{1}{Z_N^{\mathrm{CW}}(\beta, B)} \exp\left(\frac{N\beta}{2}\overline{\boldsymbol{\sigma}}^2 + B\sum_{i=1}^N \sigma_i\right). \tag{1.3}$$

Let $Z_\tau \sim N(0, \tau)$ with $\tau := \frac{1-t^2}{1-\beta(1-t^2)}$ for $(\beta, B) \notin \Theta_3$, and let W be a continuous random variable with density proportional to $e^{-x^4/12}$. Then the following hold:

$$\sqrt{N}(\overline{\boldsymbol{\sigma}} - t) \stackrel{d}{\to} Z_\tau \quad \textit{if } (\beta, B) \in \Theta_1,$$

$$\sqrt{N}\big(\overline{\boldsymbol{\sigma}} - M(\boldsymbol{\sigma})\big) \stackrel{d}{\to} Z_\tau \quad \textit{if } (\beta, B) \in \Theta_2,$$

$$N^{1/4}\overline{\boldsymbol{\sigma}} \stackrel{d}{\to} W \quad \textit{if } (\beta, B) \in \Theta_3.$$

Here $M(\boldsymbol{\sigma})$ is a random variable which equals t if $\overline{\boldsymbol{\sigma}} \geq 0$, and $-t$ otherwise, whenever $(\beta, B) \in \Theta_2$.

We will now explore to what extent the fluctuations of $\overline{\boldsymbol{\sigma}}$ are universal. We need the following notation to state our main results.

DEFINITION 1.2.

(i) Given two positive sequences x_N, y_N, we use the notation $x_N \lesssim y_N$ to denote the existence of a finite constant C free of N, such that $x_N \leq C y_N$.

(ii) Given a symmetric matrix A_N, let $R_i := \sum_{j=1}^N A_N(i, j)$ denote the row sums of A_N, and let $(\lambda_1(A_N), \ldots, \lambda_N(A_N))$ denote its eigenvalues arranged in decreasing order. Let $\|A_N\|_F$ and $\|A_N\|_{\mathrm{op}}$ denote the Frobenius norm and the operator norm of A_N respectively.

(iii) Given two real valued random variables X, Y, define the Kolmogorov–Smirnov distance between X and Y by

$$d_{\mathrm{KS}}(X, Y) := \sup_{x \in \mathbb{R}} \big|\mathbb{P}(X \leq x) - P(Y \leq x)\big|.$$

THEOREM 1.1. *Suppose that $(\beta, B) \in \Theta_1$. Assume further that the sequence of matrices A_N satisfies the following two conditions*:

$$\max_{1 \leq i \leq N} R_i \lesssim 1, \tag{1.4}$$

$$\lim_{N \to \infty} \lambda_1(A_N) = 1. \tag{1.5}$$

If $\boldsymbol{\sigma}$ is a random vector from the Ising model (1.1), then we have

$$d_{\mathrm{KS}}\big(\sqrt{N}(\overline{\boldsymbol{\sigma}} - t), Z_\tau\big) \lesssim \frac{1}{\sqrt{N}}\left(\|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2 + t\left|\sum_{i=1}^N (R_i - 1)\right|\right), \tag{1.6}$$

where Z_τ is defined as in Lemma 1.2.

Note that Theorem 1.1 leaves out the parameter regime $\Theta_2 \cup \Theta_3$. The following example shows that such a universal behavior is not expected in this parameter regime, unless we assume some notion of connectivity for A_N.

EXAMPLE 1.1. With N even, let A_N be the adjacency matrix of two disjoint complete graphs $K_{N/2}$, scaled by $N/2$. Then the following hold:

(a) If $(\beta, B) \in \Theta_2$, then $\overline{\boldsymbol{\sigma}} \xrightarrow{d} \frac{1}{2}\delta_0 + \frac{1}{4}(\delta_t + \delta_{-t})$.

(b) If $(\beta, B) \in \Theta_3$, then $N^{1/4}\overline{\boldsymbol{\sigma}} \xrightarrow{d} (W_1 + W_2)/2^{3/4}$, where W_1, W_2 are i.i.d. with the same distribution as that of W, with W defined as in Lemma 1.2.

The above example shows that if we want universal fluctuations in the regimes $\Theta_2 \cup \Theta_3$, the matrix A_N needs to be "connected" in some asymptotic sense. If A_N is exactly the adjacency matrix of a d_N regular graph G_N scaled by d_N, then $\lambda_1(A_N) = 1$, and it is easy to check that the graph G_N is connected iff there is a spectral gap, that is, $\lambda_2(A_N) < 1$. Motivated by this, we propose the following asymptotic notion of a spectral gap.

DEFINITION 1.3. We say a sequence of symmetric matrices $\{A_N\}_{N\geq 1}$ with nonnegative entries satisfies the spectral gap condition, if

$$\limsup_{N\to\infty} \frac{\lambda_2(A_N)}{\lambda_1(A_N)} < 1. \tag{1.7}$$

We note that assumption (1.7) is somewhat weak in the sense that it does not imply connectivity in general. In particular this allows the existence of small disconnected subgraphs in G_N, as shown in the following example.

EXAMPLE 1.2. Let G_N denote a graph which is the disjoint union of a d_N regular graph G_{1,N_1} on N_1 vertices, and an arbitrary graph G_{2,N_2} on N_2 vertices, with $N_1 + N_2 = N$ and $N_2 = o(d_N)$. Then the average degree of the whole graph G_N is $\widetilde{d}_N = d_N(1 + o(1))$. It is easy to check that if G_{1,N_1} satisfies (1.7), then G_N satisfies (1.7), even though G_N is disconnected.

Under the assumption of a spectral gap, our next result shows universal fluctuations in the nonuniqueness regime.

THEOREM 1.2. *Suppose that $(\beta, B) \in \Theta_2$. Assume further that the sequence of matrices A_N satisfies* (1.4), (1.5), *and* (1.7). *If $\boldsymbol{\sigma}$ is a random vector from the Ising model* (1.1), *then we have*

$$d_{\mathrm{KS}}\big(\sqrt{N}(\overline{\boldsymbol{\sigma}} - M(\boldsymbol{\sigma})), Z_\tau\big) \lesssim \frac{1}{\sqrt{N}}\left(\|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2 + \left|\sum_{i=1}^N (R_i - 1)\right|\right), \tag{1.8}$$

where $M(\boldsymbol{\sigma})$ and Z_τ are defined as in Lemma 1.2.

To prove universal fluctuations in the critical regime, we need a stronger notion of regularity on A_N, that is,

$$\limsup_{N\to\infty} N^{1/4} \max_{1\leq i\leq N} |R_i - 1| \lesssim 1. \tag{1.9}$$

THEOREM 1.3. *Suppose that $(\beta, B) \in \Theta_3$. If $\boldsymbol{\sigma}$ is a random vector from the Ising model* (1.1) *where A_N satisfies* (1.7) *and* (1.9). *Then we have*

$$\begin{aligned} d_{\mathrm{KS}}(N^{1/4}\overline{\boldsymbol{\sigma}}, W) &\lesssim \frac{\varepsilon_N}{\sqrt{N}} + \frac{\varepsilon_N r_N}{N^{1/4}} \\ &\quad + \frac{(\log N)^2}{N^{1/4}}\sqrt{\sum_{i=1}^N (R_i - 1)^2 + N^{-1/2}\left[\sum_{i=1}^N (R_i - 1)\right]^2}, \end{aligned} \tag{1.10}$$

where

$$r_N := \sqrt{(\log N)^3 \max_{1\le i\le N} \sum_{j=1}^N A_N(i,j)^2 + \log N \max_{1\le i\le N} |R_i - 1|},$$

$$\varepsilon_N := \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2 + \frac{1}{N}\sum_{i=1}^N (R_i-1)^2 + \log N,$$

and W is as in Lemma 1.2.

REMARK 1.1. Using these results, in Section 1.3 we will show that for any sequence of d_N regular graphs satisfying the spectral gap condition (see (1.7)), the fluctuation of $\overline{\boldsymbol{\sigma}}$ is universal in $\Theta_1 \cup \Theta_2$ if $d_N \gg \sqrt{N}$, and in Θ_3 if $d_N \gg \sqrt{N}\log N$. We now give an example to show that the above conditions are actually tight (up to log factor in the critical regime). The proof of this example will appear in an upcoming draft [36].

EXAMPLE 1.3. Let G_N denote the line graph of the complete graph K_n, so that $N = \binom{n}{2} = \frac{n^2}{2}(1+o(1))$. This is a regular graph with degree $d_N = 2(n-2) = 2\sqrt{2N}(1+o(1))$, and its top two eigenvalues are $\lambda_1(G_N) = 2(n-2)$ and $\lambda_2(G_N) = n-2$ (see [15], Lemma 2). It follows that $A_N = \frac{1}{d_N} G_N$ does satisfy (1.7), and

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}}\|A_N\|_F^2 = \sqrt{N} \max_{1\le i\le N} \sum_{j=1}^N A_N(i,j)^2 = \frac{1}{2\sqrt{2}} \ne 0.$$

In this case we have the following limiting distributions across different regimes:

$$\sqrt{N}(\overline{\boldsymbol{\sigma}}_N - t) + \mu \xrightarrow{w} Z_\tau \quad \text{if } (\beta, B) \in \Theta_1,$$

$$\sqrt{N}\big(\overline{\boldsymbol{\sigma}}_N - M(\boldsymbol{\sigma})\big) + \operatorname{sgn}\big(M(\boldsymbol{\sigma})\big)\mu \xrightarrow{w} Z_\tau \quad \text{if } (\beta, B) \in \Theta_2,$$

$$N^{1/4}\overline{\boldsymbol{\sigma}}_N \xrightarrow{w} \widetilde{W} \quad \text{if } (\beta, B) \in \Theta_3,$$

where $\mu := \frac{\beta t}{\sqrt{2}(1-\beta(1-t^2))\cdot(2-\beta(1-t^2))}$ is strictly larger than 0 if $(\beta, B) \in \Theta_{12} \cup \Theta_2 \cup \Theta_3$, and $\widetilde{W}$ has density proportional to $\exp(-\frac{w^4}{12} - \frac{w^2}{\sqrt{2}})$. Therefore, the fluctuations do not match that of the Curie–Weiss model unless $(\beta, B) \in \Theta_{11}$.

Note that in the above example, $\overline{\boldsymbol{\sigma}}$ has a different limit compared to the Curie–Weiss model in $\Theta_{12} \cup \Theta_2 \cup \Theta_3$, but continues to have universal fluctuations in the high parameter regime Θ_{11}. We now state a modified theorem for the regime Θ_{11}, which shows that in this regime we can do better.

THEOREM 1.4. *Suppose that $(\beta, B) \in \Theta_{11}$, and A_N satisfies*

$$\lim_{N\to\infty} \max_{1\le i\le N} R_i = 1. \tag{1.11}$$

If $\boldsymbol{\sigma}$ is a random vector from the Ising model (1.1), then setting $\alpha_N := \max_{1\le i\le N} \sum_{j=1}^N A_N(i, j)^2$ we have

$$\begin{aligned} d_{\mathrm{KS}}(\sqrt{N}\overline{\boldsymbol{\sigma}}, Z_\tau) \lesssim{}& \frac{1}{\sqrt{N}} + \frac{\|A_N\|_F^2 \sqrt{\alpha_N \log N}}{\sqrt{N}} \\ &+ \big[1 + \|A_N\|_F \alpha_N \log N\big]\sqrt{\frac{\sum_{i=1}^N (R_i-1)^2}{N}}, \end{aligned} \tag{1.12}$$

where Z_τ is defined as in Lemma 1.2.

REMARK 1.2. It follows from the above result that in the regime Θ_{11}, $\overline{\boldsymbol{\sigma}}$ has universal fluctuations on regular graphs of degree $d_N \gg (N \log N)^{1/3}$. We believe this is not tight, and universal fluctuations should hold on any sequence of regular graphs with $d_N \to \infty$. In [26] the authors prove such a result when G_N is a nonsymmetric Erdős–Rényi graph in the regime Θ_{11} (details in example section below).

Note that we only expect a similar behavior as in the Curie–Weiss model, if the underlying graphs are approximately regular and have large degree. Quantifying this philosophy, the bounds in each of the theorems have two terms, the first term controls the sparsity of the underlying graph/matrix, and the second term controls the extent of regularity of the graph/matrix. Recall example 1.3, which suggests that the term controlling the sparsity is optimal. In a similar spirit, the following example suggests that the term controlling the extent of regularity is also optimal.

EXAMPLE 1.4.

(a) Assume that $\sqrt{N}$ is an integer, and let G_N be the disjoint union of two complete graphs of size $N - \sqrt{N}$ and $\sqrt{N}$ respectively. Let $\overline{d}_N$ denote the average degree of G_N and $A_N = (\overline{d}_N)^{-1} G_N$. In this case

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (R_i - 1)^2 > 0,$$

but every other term in the RHS of (1.6) converges to 0. If $\boldsymbol{\sigma}$ is a random vector from the Ising model (1.1) with $B \neq 0$, then $\sqrt{N}(\overline{\boldsymbol{\sigma}} - t) \xrightarrow{w} \mu + Z_\tau$, where $\mu := \frac{\beta t(1-t^2)}{1-\beta(1-t^2)} + \tanh(B) - t \neq 0$.

(b) With $G_N = K_N$, let $A_N = \frac{1}{N-\sqrt{N}} G_N$. In this case

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (R_i - 1) > 0,$$

but every other term in the RHS of (1.6) converges to 0. If $\boldsymbol{\sigma}$ is a random vector from the Ising model (1.1) with $B \neq 0$, then $\sqrt{N}(\overline{\boldsymbol{\sigma}} - t) \xrightarrow{w} \mu + Z_\tau$, where $\mu := \frac{\beta t(1-t^2)}{1-\beta(1-t^2)} \neq 0$.

The main ingredient of our proof technique is comparing the Ising model on an approximately regular graph to that of an i.i.d. model/Curie–Weiss model. As a byproduct of this approach, we also obtain quantitative bounds for the following asymptotics of the log partition function via the mean-field prediction formula, defined via the following lower bound (cf. [4]):

$$\log Z_N(\beta, B) \geq \sup_{\boldsymbol{\sigma}\in[-1,1]^N} \left\{ \frac{\beta}{2} \boldsymbol{\sigma}^\top A_N \boldsymbol{\sigma} + B \sum_{i=1}^{N} \sigma_i - \sum_{i=1}^{N} I(\sigma_i) \right\},$$

where $I(x) := \frac{1+x}{2} \log \frac{1+x}{2} + \frac{1-x}{2} \log \frac{1-x}{2}$ is the binary entropy function. By choosing $\boldsymbol{\sigma} = t\mathbf{1}$ with t as defined in Lemma 1.1, we get the further lower bound

$$\log Z_N(\beta, B) \geq N \left\{ \frac{\beta t^2}{2} + Bt - I(t) \right\} + \frac{\beta t^2}{2} \sum_{i=1}^{N} (R_i - 1) =: \mathcal{M}_N(\beta, B). \tag{1.13}$$

It follows from [4], Theorem 2.1, that $\log Z_N(\beta, B) - \mathcal{M}_N(\beta, B) = o(N)$, as soon as $\|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2 = o(N)$. Our next result gives a bound to the approximation error of the partition function $Z_N(\beta, B)$ by $\mathcal{M}_N(\beta, B)$, which we henceforth refer to as the mean-field prediction in this paper.

THEOREM 1.5. *Let A_N satisfy* (1.4) *and* (1.5).

(a) *If $(\beta, B) \in \Theta_1$ then we have*

$$\log Z_N(\beta, B) - \mathcal{M}_N(\beta, B) \lesssim \|A_N\|_F^2 + t^2 \sum_{i=1}^N (R_i - 1)^2.$$

(b) *If $(\beta, B) \in \Theta_2$, then the same conclusion as in part* (*a*) *holds under the extra assumption that A_N satisfies* (1.7).

(c) *If $(\beta, B) \in \Theta_3$, then under the extra assumption that A_N satisfies* (1.7) *we have*

$$\log Z_N(\beta, B) - \mathcal{M}_N(\beta, B) \lesssim \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)^2\right]^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2 + \log N.$$

REMARK 1.3. To see how the error bounds of the above theorem compare to existing error bounds for the mean-field prediction formula in the literature, let us take the example where A_N is the (scaled) adjacency matrix of a d_N-regular graph G_N. In this case, the above theorem gives the error bound $O(N/d_N)$ for the mean-field prediction formula. This immediately improves the bounds from [4], Theorem 1.1—$o(N)$, [25], Theorem 1.1—$O(N/d_N^{1/3})$, [20], Example 3—$O(N/d_N^{1/2-o(1)})$ under strong expander type conditions not needed here) and [2], Corollary 2.9 and Example 2.10—$O(N/\sqrt{d_N})$.

For our next result, define an i.i.d. probability measure $\mathbb{Q}$ on $\{-1, 1\}^N$ by setting

$$\mathbb{Q}(\sigma_1, \dots, \sigma_N) := \big(\exp(-\beta t - B) + \exp(\beta t + B)\big)^{-N} \exp\left((\beta t + B) \sum_{i=1}^N \sigma_i\right). \tag{1.14}$$

Our next theorem shows that if an event is unlikely under the above i.i.d. measure/ the Curie–Weiss model (depending on (β, B)), then it is also unlikely under an Ising model on an approximately regular graph with large degree.

THEOREM 1.6. *Let A_N satisfy* (1.4) *and* (1.5). *Also, let $\mathcal{E}_N \subset \{-1, 1\}^N$ be arbitrary.*

(a) *If $(\beta, B) \in \Theta_1$, then we have*

$$\log \mathbb{P}(\mathcal{E}_N) \lesssim \log \mathbb{Q}(\mathcal{E}_N) + \|A_N\|_F^2 + t^2 \sum_{i=1}^N (R_i - 1)^2.$$

(b) *If $(\beta, B) \in \Theta_2$, then under the further assumption* (1.7) *we have*

$$\log \mathbb{P}(\mathcal{E}_N) \lesssim \log \mathbb{P}^{\mathrm{CW}}(\mathcal{E}_N) + \|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2.$$

(c) *If $(\beta, B) \in \Theta_3$, then under the further assumption* (1.7) *we have*

$$\log \mathbb{P}(\mathcal{E}_N) \lesssim \log \mathbb{P}^{\mathrm{CW}}(\mathcal{E}_N) + \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)^2\right]^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2 + \log N.$$

As an application of the above theorem, we immediately get the following exponential concentration for $\overline{\boldsymbol{\sigma}}$.

COROLLARY 1.1. *Suppose A_N satisfies* (1.4), (1.5), *and*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N (R_i-1)^2=0, \qquad \lim_{N\to\infty}\frac{1}{N}\|A_N\|_F^2=0.$$

- *If $(\beta, B)\in\Theta_1$, then for every $\delta>0$ we have*

$$\limsup_{n\to\infty}\frac{1}{N}\log\mathbb{P}\big(|\overline{\boldsymbol{\sigma}}-t|>\delta\big)<0.$$

The same conclusion holds for $(\beta, B)\in\Theta_3$, under the extra assumption that A_N satisfies (1.7).

- *If $(\beta, B)\in\Theta_2$, then under the extra assumption that A_N satisfies* (1.7), *for every $\delta>0$ we have*

$$\limsup_{n\to\infty}\frac{1}{N}\log\mathbb{P}\big(\big|\overline{\boldsymbol{\sigma}}-M(\boldsymbol{\sigma})\big|>\delta\big)<0,$$

where $M(\boldsymbol{\sigma})$ is defined as in Lemma 1.2.

Similar concentration results can be obtained for other higher order polynomials of $\boldsymbol{\sigma}$, as studied in [1, 12, 23] and the references therein. However, these papers focus exclusively on the high temperature regime Θ_{11} whereas our result applies to all temperatures. The references cited above can deal with nonferromagnetic interactions and general external fields as well. We note in passing that it should be possible to extend our proof technique to general nonconstant magnetic fields.

1.2. *Proof overview.* For the sake of simplicity, we focus on the case where A_N is the adjacency matrix of a d_N regular graph scaled by d_N. For verifying Theorem 1.1, following [14], Theorem 2.1, form an exchangeable pair $(\boldsymbol{\sigma},\boldsymbol{\sigma}')$ as follows:

Let I denote a randomly sampled index from $\{1,2,\dots,N\}$. Given $I=i$, replace σ_i with an independent ±1 valued random variable σ_i' with mean $\mathbb{E}[\sigma_i|(\sigma_j, j\neq i)]=\tanh(\beta m_i(\boldsymbol{\sigma})+B)$, where $m_i(\boldsymbol{\sigma}):=\sum_{j=1}^N A_N(i,j)\sigma_j$. Then, setting $\boldsymbol{\sigma}':=(\sigma_1,\dots,\sigma_{i-1},\sigma_i',\sigma_{i+1},\dots,\sigma_N)$, we have that $(\boldsymbol{\sigma},\boldsymbol{\sigma}')$ is an exchangeable pair. With $T_N:=\sqrt{N}(\overline{\boldsymbol{\sigma}}-t)$ and $T_N':=\sqrt{N}(\overline{\boldsymbol{\sigma}'}-t)$, a simple computation using a Taylor series expansion of $\tanh(\beta x+B)$ around $x=t$ gives

$$\begin{aligned}\mathbb{E}[T_N-T_N'|\boldsymbol{\sigma}]&=N^{-3/2}\sum_{i=1}^N\big(\sigma_i-\tanh\big(\beta m_i(\boldsymbol{\sigma})+B\big)\big)\\&=N^{-3/2}\Bigg[\sum_{i=1}^N\big(\sigma_i-\tanh(\beta t+B)\big)+\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-t\big)\operatorname{sech}^2(\beta t+B)\\&\quad+O_P\Bigg(\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-t\big)^2\Bigg)\Bigg].\end{aligned}$$

Since G_N is regular, we have $\sum_{i=1}^N\sigma_i=\sum_{i=1}^N m_i(\boldsymbol{\sigma})$. Also t satisfies $t=\tanh(\beta t+B)$, and so the above display gives

$$\mathbb{E}[T_N-T_N'|\boldsymbol{\sigma}]=\frac{T_N}{N}\big(1-\beta(1-t^2)\big)+N^{-3/2}O_P\Bigg(\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-t\big)^2\Bigg). \tag{1.15}$$

By [14], Theorem 1.2, T_N approximately satisfies Stein's equation if we show that the second term in the RHS above is negligible, that is, $S_N:=\sum_{i=1}^N(m_i(\boldsymbol{\sigma})-t)^2=o(\sqrt{N})$. This is the content of Lemma 2.1, which bounds the exponential moment of S_N to show that $S_N = O_P(\frac{N}{d_N})$. Thus we require $d_N \gg \sqrt{N}$ to ensure the linear term in (1.15) dominates the error term. The main ingredient for Lemma 2.1 is a version of the Hanson–Wright inequality for $\{-1, +1\}$ valued random variables (cf. Lemma 4.1). Justifying the above steps gives a proof of Theorem 1.1. The proof of Theorem 1.2 follows on similar lines, after replacing t above by $M(\boldsymbol{\sigma})$, where $M(\boldsymbol{\sigma}) = t$ if $\overline{\boldsymbol{\sigma}} \geq 0$, and $M(\boldsymbol{\sigma}) = -t$ otherwise, as defined in Lemma 1.2.

The above program does not work for Theorem 1.3, which deals with the critical regime Θ_3. This is because $\beta(1-t^2) = 1$, and so the linear term in (1.15) vanishes. With $T_N = N^{1/4}\overline{\boldsymbol{\sigma}}$, using a Taylor series expansion of $\tanh(x)$ around $x = \overline{\boldsymbol{m}}(\boldsymbol{\sigma}) := N^{-1}\sum_{i=1}^N m_i(\boldsymbol{\sigma})$ and following similar steps as the derivation of (1.15) we have

$$\begin{aligned} \mathbb{E}[T_N - T_N' | \boldsymbol{\sigma}] = N^{-7/4}\Bigg[&\sum_{i=1}^N \big(\sigma_i - \tanh(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))\big) + \frac{\tanh''(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))}{2}\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2 \\ &+ O_P\Bigg(\sum_{i=1}^N \big|m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big|^3\Bigg)\Bigg]. \end{aligned} \tag{1.16}$$

Noting that $\overline{\boldsymbol{m}}(\boldsymbol{\sigma}) = \overline{\boldsymbol{\sigma}}$, the leading term in the RHS of (1.16) equals $N^{-3/4}(\overline{\boldsymbol{\sigma}} - \tanh(\overline{\boldsymbol{\sigma}})) \approx \frac{1}{3N^{3/4}}\overline{\boldsymbol{\sigma}}^3$. From here, provided one can ignore the two error terms in (1.16), we can use [14], Theorem 1.2, to show that T_N converges in distribution to the nonnormal limit W, as desired. The main obstacle is the nontrivial step of bounding the error terms in (1.16). To this effect, note that the error terms in the RHS of (1.16) can be bounded as follows:

$$\frac{\tanh''(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))}{2}\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2 = O_P(\overline{\boldsymbol{\sigma}}\widetilde{S}_N),$$

$$\sum_{i=1}^N \big|m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big|^3 = O_P\Big(\max_{1\le i\le N}\big|m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big|\widetilde{S}_N\Big),$$

where $\widetilde{S}_N := \sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2$. Thus in contrast to what happened before, it no longer suffices to bound only the quadratic term $\widetilde{S}_N$, but instead we also need to bound $\max_{i\in[N]} |m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|$. The estimate for $\widetilde{S}_N$ is done by introducing an auxiliary variable to express $\boldsymbol{\sigma}$ as a mixture of i.i.d. distributions, and then using Lemma 4.1. The more challenging task is to bound $\max_{i\in[N]} |m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|$, which we achieve by using a novel recursive argument, as follows:

Using the method of concentration via exchangeable pairs, we first show the approximate fixed point equation

$$m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}) \overset{P}{\approx} \sum_{j=1}^N A_N(i,j)\big(m_j(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big).$$

Writing $\widetilde{\mathbf{m}}(\boldsymbol{\sigma}) := (m_1(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}), \dots, m_N(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))$, and recursing the above fixed point equation we get $\widetilde{\mathbf{m}} \approx A_N^k \widetilde{\mathbf{m}}$, which gives

$$\big|\widetilde{m}_i(\boldsymbol{\sigma})\big| \approx \Bigg|\sum_{j=1}^N (A_N^k)(i,j)\widetilde{m}_j(\boldsymbol{\sigma})\Bigg| \lesssim \sqrt{\sum_{j=1}^N \widetilde{m}_j(\boldsymbol{\sigma})^2}\sqrt{(A_N^{2k})(i,i)}. \tag{1.17}$$

Since A_N is the scaled adjacency matrix of a connected regular graph, there can be at most two eigenvalues with absolute value 1, and so for k large enough the contribution of all other eigenvalues to A_N^{2k} should be negligible. Also the corresponding normalized eigenvectors for these two eigenvalues must have all entries equal to $\frac{1}{\sqrt{N}}$ in absolute value. This suggests the approximate inequality

$$(A_N^{2k})(i,i) \le \frac{2}{N} \tag{1.18}$$

for k large enough. In Lemma 6.2, we show the above bound for general regular matrices with nonnegative entries satisfying the spectral gap condition (1.7), but no condition on the minimum eigenvalue (i.e., no expander type condition). Of course such a result is not correct if (1.7) does not hold, as then G_N can be disconnected. Plugging the bound (1.18) in (1.17) along with the estimate $\widetilde{S}_N = \sum_{i=1}^N \widetilde{m}_i(\boldsymbol{\sigma})^2 = O_P(\frac{N}{d_N})$ gives

$$\max_{1\le i\le N} \widetilde{m}_i(\boldsymbol{\sigma}) = O_P\left(\sqrt{\frac{N}{d_N} \times \frac{2}{N}}\right) \lesssim \frac{1}{\sqrt{d_N}}.$$

Because of standard union bounds, we incur a log factor and deduce the estimate

$$\max_{1\le i\le N} |m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})| = \max_{1\le i\le N} |\widetilde{m}_i(\boldsymbol{\sigma})| = O_P\left(\sqrt{\frac{\log N}{d_N}}\right).$$

Plugging this bound back into (1.16) shows that both the error terms are negligible, and hence gives an approximate Stein's equation for W thereby completing the proof of Theorem 1.3.

For verifying Theorem 1.4 in the regime Θ_{11}, we use a modified version of (1.15) with $T_N = \sqrt{N}\overline{\boldsymbol{\sigma}}$, where we expand $\tanh(\beta x)$ around $x = 0$:

$$\begin{aligned}\mathbb{E}[T_N - T_N' | \boldsymbol{\sigma}] &= \frac{1}{N^{3/2}} \sum_{i=1}^N (\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma}))) \\ &= \frac{1}{N^{3/2}} \left[\sum_{i=1}^N (\sigma_i - \beta m_i(\boldsymbol{\sigma})) + O_P\left(\sum_{i=1}^N |m_i(\boldsymbol{\sigma})|^3\right)\right] \\ &= \frac{(1-\beta)T_N}{\sqrt{N}} + O_P\left(\max_{1\le i\le N} |m_i(\boldsymbol{\sigma})| \sum_{i=1}^N m_i(\boldsymbol{\sigma})^2\right).\end{aligned}$$

As before, to complete the proof one needs to show that the error term above is negligible. The quadratic term $S_N = \sum_{i=1}^N m_i(\boldsymbol{\sigma})^2$ is controlled using Lemma 2.1, and the max term is controlled by setting up another fixed equation (see Lemma 2.2 part (a)).

The above sketch works for exactly regular graphs. To handle approximately regular graphs/matrices, we need to bound the moments of $\mathbf{c}^{\mathrm{T}}\boldsymbol{\sigma}$ where $c_i = R_i - 1$, in the regimes Θ_{11} and Θ_3. This requires another recursive argument, and is carried out in Lemma 2.2 part (b) and Lemma 2.3 part (b) for regimes Θ_{11} and Θ_3 respectively. In fact, the proof of Lemma 2.2 applies to general vectors $\mathbf{c}$, and the proof of Lemma 2.3 can be modified to handle this case.

1.3. *Examples*. As mentioned before, the most common example of a coupling matrix A_N in model (1.1) is the scaled adjacency matrix $\frac{1}{\overline{d}_N} G_N$, where G_N is the adjacency matrix of a simple labeled graph on N vertices with degree vector $(d_1, \dots, d_N)$, and $\overline{d}_N := \frac{1}{N}\sum_{i=1}^N d_i$ is the average degree of G_N. The scaling discussed in the above definition ensures that the resulting Ising model has nontrivial phase transition properties (see, e.g., [4, 34]). Below we consider some specific examples of graphs to illustrate our theorems.

(a) *Regular graphs*: Let G_N be a d_N regular graph. Then $\|A_N\|_F^2 = \frac{N}{d_N}$ and $R_i = 1$, and so applying Theorems 1.1, 1.2, 1.3, and 1.4 give

$$d_{\mathrm{KS}}(\sqrt{N}(\overline{\boldsymbol{\sigma}} - t), Z_\tau) \lesssim \sqrt{\frac{N \log N}{d_N^3}} + \frac{1}{\sqrt{N}} \quad \text{if } (\beta, B) \in \Theta_{11},$$

$$
\begin{aligned}
d_{\mathrm{KS}}\big(\sqrt{N}(\overline{\boldsymbol{\sigma}}-t), Z_\tau\big) &\lesssim \frac{\sqrt{N}}{d_N} \quad \text{if } (\beta, B) \in \Theta_{12},\\
d_{\mathrm{KS}}\big(\sqrt{N}\big(\overline{\boldsymbol{\sigma}}-M(\boldsymbol{\sigma})\big), Z_\tau\big) &\lesssim \frac{\sqrt{N}}{d_N} \quad \text{if } (\beta, B) \in \Theta_{2} \text{ and } G_N \text{ satisfies (1.7)},\\
d_{\mathrm{KS}}\big(N^{1/4}\overline{\boldsymbol{\sigma}}, W\big) &\lesssim \left(\frac{\sqrt{N}\log N}{d_N}\right)^{3/2} + \frac{\sqrt{N}}{d_N}\\
&\quad + \frac{\log N}{\sqrt{N}} \quad \text{if } (\beta, B) \in \Theta_{3} \text{ and } G_N \text{ satisfies (1.7)},
\end{aligned}
$$

where Z_τ and W are defined as in Lemma 1.2. In particular this means that $\overline{\boldsymbol{\sigma}}$ has the same fluctuations as that of the Curie–Weiss model as soon as

$$
\begin{aligned}
&d_N \gg (N\log N)^{1/3} \quad \text{if } (\beta, B) \in \Theta_{11},\\
&d_N \gg \sqrt{N} \quad \text{if } (\beta, B) \in \Theta_{12},\\
&d_N \gg \sqrt{N} \quad \text{if } (\beta, B) \in \Theta_{2} \text{ and (1.7) holds},\\
&d_N \gg \sqrt{N}\log N \quad \text{if } (\beta, B) \in \Theta_{3} \text{ and (1.7) holds}.
\end{aligned}
\tag{1.19}
$$

Further, as already shown in Example 1.3, the requirement $d_N \gg \sqrt{N}$ is sharp in the regimes $\Theta_{12} \cup \Theta_2 \cup \Theta_3$. Note that for the particular case of the Curie–Weiss model at criticality we get the convergence rate of $N^{-1/2}\log N$, which matches the rate obtained in [14] up to the log factor. In fact, it is easy to modify our argument in the special case of the Curie–Weiss model to get rid of the log factor. We observe that for the case of random d_N regular graphs, condition (1.7) holds with high probability, as $\lambda_2(G_N) = O_P(\sqrt{d_N}) \ll d_N$ (see [11]), and so our results apply directly to random regular graphs if d_N satisfies (1.19). We stress that our results apply to regular bipartite graphs as well, and does not need the graph to be an expander as in [10].

(b) *Erdős–Rényi graphs*: Suppose $G_N \sim \mathcal{G}(N, p_N)$ is the symmetric Erdős–Rényi random graph with $0 < p_N \le 1$. Define $A_N(i,j) := \frac{1}{(N-1)p_N} G_N(i,j)$, and note that

$$
\begin{aligned}
\max_{1\le i\le N} |R_i - 1| &= O_P\left(\sqrt{\frac{\log N}{N p_N}}\right),\\
\left|\sum_{i=1}^N (R_i - 1)\right| &= O_P\left(\frac{1}{\sqrt{p_N}}\right),\\
\sum_{i=1}^N (R_i - 1)^2 &= O_P\left(\frac{1}{p_N}\right).
\end{aligned}
\tag{1.20}
$$

Since $\lambda_2(G_N) = O_P(\sqrt{N p_N}) \ll N p_N$ ([22], Theorem 1.1), (1.7) holds as well. Then our theorems conclude universal fluctuations for $\overline{\boldsymbol{\sigma}}$ as soon as

$$
\begin{aligned}
&p_N \gg (\log N)^{1/3} N^{-2/3} \quad \text{if } (\beta, B) \in \Theta_{11},\\
&p_N \gg N^{-1/2} \quad \text{if } (\beta, B) \in \Theta_{12} \cup \Theta_2,\\
&p_N \gg (\log N)^4 N^{-1/2} \quad \text{if } (\beta, B) \in \Theta_{3},
\end{aligned}
\tag{1.21}
$$

both in the quenched and annealed setting. We note that our results also apply to the asymmetric Erdős–Rényi random graph $\widetilde{\mathcal{G}}(N, p_N)$, under the same regime of p_N as in the symmetric case. This is because an Ising model on the asymmetric Erdős–Rényi graph is equivalent to an Ising model with the symmetric coupling matrix $A_N(i,j) = \frac{\widetilde{\mathcal{G}}_N(i,j)+\widetilde{\mathcal{G}}_N(j,i)}{2(N-1)p_N}$, which is approximately regular, as

$$R_i = \sum_{j=1}^N A_N(i,j) = \frac{1}{2(N-1)p_N}\sum_{j=1}^N \bigl(\widetilde{\mathcal{G}}_N(i,j)+\widetilde{\mathcal{G}}_N(j,i)\bigr) \sim \frac{\mathrm{Bin}(2(N-1),p_N)}{2(N-1)p_N} \overset{P}{\approx} 1,$$

where the last approximation (in the sense of (1.20)) follows by a standard application of Chernoff's inequality. The asymmetric case was studied recently in [26], where the authors derive fluctuations as soon as $Np_N \to \infty$, but only in the subparameter regime $\Theta_{11} \cup \Theta_3$. The authors conjecture similar results for the symmetric case, which we are able to verify partially in this paper. Moreover, our theorems apply simultaneously to both the symmetric and the asymmetric cases with explicit convergence rates. A few months after our paper was submitted, [27] was uploaded where the authors obtain fluctuations for the magnetization in the asymmetric case for the parameter regime $\Theta_{12} \cup \Theta_2$ when $N^{1/3}p_N \to \infty$, in [27], Theorems 1.1 and 1.3. In contrast, our results show universal fluctuations in the larger regime $N^{1/2}p_N \to \infty$, and apply to both the symmetric and asymmetric cases, from which the fluctuation results for the magnetization in [27] follow as corollaries. On the other hand, in [27], Theorem 1.4, the authors derive a central limit theorem for the log partition function when $N^{1/3}p_N \to \infty$, a direction which is not explored in our paper.

(c) *Balanced stochastic block model*: Suppose G_N is a stochastic block model with two communities of size $N/2$ (assume N is even). Let the probability of an edge within the community be a_N, and across communities be b_N. This is the well-known stochastic block model, which has received considerable attention in Probability, Statistics, and Machine Learning (see [18, 30, 33] and references within). If we take $A_N = \frac{2}{N(a_N+b_N)}G_N$, universal asymptotics hold for $\overline{\boldsymbol{\sigma}}$ as soon as $p_N := \frac{a_N+b_N}{2}$ satisfies (1.21), and $\liminf_{N\to\infty}\frac{b_N}{a_N} > 0$ (needed to ensure (1.7)). Similar results hold when the number of communities is larger than two.

(d) *Sparse regular graphons*: Suppose that W be a symmetric measurable function from $[0,1]^2$ to $[0,1]$, such that $\int_{[0,1]} W(x,y)\,dy = a > 0$ for all $x \in [0,1]$, and $\lambda_2(W) < a$, where $\{\lambda_i(W)\}_{i\ge 1}$ are the countable set of ordered eigenvalues. Also let $(U_1,\dots,U_N) \overset{\text{i.i.d.}}{\sim} U(0,1)$. For $\gamma \in (0,1]$, let

$$\{G_N(i,j)\}_{1\le i<j\le N} \overset{\text{i.i.d.}}{\sim} \mathrm{Bern}\biggl(\frac{W(U_i,U_j)}{N^\gamma}\biggr).$$

Such random graph models have been studied in the literature under the name W random graphons (cf. [6–9, 31]). In this case for the choice $A_N = \frac{1}{Np_N}G_N$ with $p_N = aN^{-\gamma}$, universal fluctuation holds as soon as $\gamma < 1/2$. Indeed, note that $\mathbb{E}[R_i|U_1,\dots,U_N] = (aN)^{-1}\sum_{j=1}^N W(U_i,U_j)$ and write

$$R_i - 1 = \biggl[R_i - \frac{\sum_{j=1}^N W(U_i,U_j)}{aN}\biggr] + \biggl[\frac{\sum_{j=1}^N W(U_i,U_j)}{aN} - 1\biggr].$$

By using Bernstein's inequality conditional on $(U_1,\dots,U_N)$, the first term is $O_P((Np_N)^{-1/2})$. Similarly, by applying Bernstein's inequality conditional on U_i, the second term is $O_P(N^{-1/2})$. An application of the union bound then implies that (1.20) holds. Also with W_N denoting the $N\times N$ matrix with $W_N(i,j) = W(U_i,U_j)$, using [3], Corollary 3.3, we have $\|A_N - (aN)^{-1}W_N\|_{\mathrm{op}} = O_P(\frac{\sqrt{N}}{Np_N})$. Since W_N converges in cut norm to W, it follows using [31], Section 11.6, that

$$\lim_{N\to\infty}\lambda_2(A_N) = a^{-1}\lim_{N\to\infty}\frac{\lambda_2(W_N)}{N} = a^{-1}\lambda_2(W) < 1$$

and so A_N satisfies (1.7). By our results, universal fluctuations hold for $\overline{\boldsymbol{\sigma}}$ as soon as (1.19) holds.

(e) *Block spin Ising model*: Suppose that N is even, and

$$\begin{aligned} A_N(i,j) &= a_N \quad \text{if } i,j \le N/2 \text{ or } i,j > N/2 \\ &= b_N \quad \text{if } i \le N/2,\, j > N/2, \text{ or } i > N/2,\, j \le N/2. \end{aligned}$$

A_N can be thought of as the expectation of a stochastic block model with two communities. In the particular case $a_N = \frac{\beta}{N}$, $b_N = \frac{\alpha}{N}$, this model has been studied in [5, 32] under the name block spin Ising model. Again in this case universal asymptotics holds for $\overline{\boldsymbol{\sigma}}$ as soon as $d_N := \frac{N(a_N+b_N)}{2}$ satisfies (1.19), and $\liminf_{N\to\infty} \frac{b_N}{a_N} > 0$. This in particular matches the results obtained from [32], Theorems 1.2, 1.4, which studies the subparameter regime $\Theta_{11} \cup \Theta_3$. Our results apply to the whole parameter regime of (β, B) and a wide regime of scalings of (a_N, b_N), providing explicit convergence rates. Similar extension holds when the matrix A_N has more than two groups as well.

(f) *Wigner matrices*: To demonstrate that our techniques apply to examples well beyond scaled adjacency matrices, let A_N be a Wigner matrix with its entries $\{A_N(i,j), 1 \le i < j \le N\}$ i.i.d. from a distribution F scaled by $N\mu$, where F is a distribution on nonnegative reals with finite exponential moment and mean $\mu > 0$. In this case we have

$$\max_{1\le i\le N} |R_i - 1| = O_P\left(\sqrt{\frac{\log N}{N}}\right), \qquad \left|\sum_{i=1}^N (R_i - 1)\right| = O_P(1), \qquad \sum_{i=1}^N (R_i-1)^2 = O_P(1).$$

Also [3], Corollary 3.5, shows that $\|A_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\|_{\text{op}} = N^{-1/2}$, and so (1.7) holds. Thus our theorems apply giving universal fluctuations for $\overline{\boldsymbol{\sigma}}$.

2. Main technical lemmas. In this section, we state our main technical lemmas which could be of independent interest. Our first result in this section is an exponential moment control lemma in all parameter regimes, which is one of the main estimates of this paper, and is itself new. The proof of this is deferred to Section 4.

LEMMA 2.1. *Suppose $\boldsymbol{\sigma}$ is an observation from* (1.1), *with A_N satisfying* (1.4) *and* (1.5).

(a) *If $(\beta, B) \in \Theta_1$, then there exists a fixed positive number $\delta > 0$ such that*

$$\log \mathbb{E}\left[\exp\left(\frac{\delta}{2}\sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - t)^2\right)\right] \lesssim \|A_N\|_F^2 + t^2 \sum_{i=1}^N (R_i - 1)^2. \tag{2.1}$$

(b) *If $(\beta, B) \in \Theta_2$, then the conclusion of part* (*a*) *holds under the additional assumption that A_N satisfies* (1.7).

(c) *If $(\beta, B) \in \Theta_3$, then under the additional assumption that A_N satisfies* (1.7) *there exists a fixed positive number $\delta > 0$ such that*

$$\begin{aligned} &\log \mathbb{E}\left[\exp\left(\frac{\delta}{2}\sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2\right)\right] \\ &\quad \lesssim \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i-1)^2\right]^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2 + \log N, \end{aligned} \tag{2.2}$$

where $\overline{\boldsymbol{m}}(\boldsymbol{\sigma}) := N^{-1}\sum_{i=1}^N m_i(\boldsymbol{\sigma})$.

Our next lemma establishes a uniform control on the $m_i(\boldsymbol{\sigma})$'s and a second moment bound on a linear statistic of interest, when $(\beta, B) \in \Theta_{11}$. The proof of this lemma is deferred to Section 5.

LEMMA 2.2. *Assume that $\boldsymbol{\sigma}$ is an observation from* (1.1) *with* $(\beta, B) \in \Theta_{11}$, *and* A_N *satisfies* (1.11). *Setting* $\alpha_N = \max_{1\le i\le N}\sum_{j=1}^N A_N(i,j)^2$ *as in Theorem* 1.4, *the following conclusions hold*:

(a) $\log\mathbb{P}(\max_{1\le i\le N}|m_i(\boldsymbol{\sigma})| \ge \lambda\sqrt{\alpha_N\log N}) \lesssim -\lambda^2$, *for any* $\lambda > 0$.
(b) $\mathbb{E}[\sum_{i=1}^N (R_i-1)\sigma_i]^2 \lesssim (\sum_{i=1}^N (R_i-1)^2)[1+\|A_N\|_F^2\alpha_N^2(\log N)^2]$.

Our final lemma yields uniform control on $\max_{1\le i\le N}|m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})|$ and moment bounds on linear statistics of interest, when $(\beta, B)\in\Theta_3$. Its proof has been deferred to Section 5.

LEMMA 2.3. *Suppose $\boldsymbol{\sigma}$ is an observation from* (1.1) *with* $(\beta, B)\in\Theta_3$, *such that* A_N *satisfies* (1.9) *and* (1.7). *Suppose further that the RHS of* (1.10) *is bounded. Then the following conclusions hold*:

(a) $\log\mathbb{P}(\max_{1\le i\le N}|m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})| \ge \lambda\sqrt{\alpha_N(\log N)^3} + \log N\max_{1\le i\le N}|R_i-1|) \lesssim -\lambda^2$, *for any* $\lambda > 0$.
(b) $\mathbb{E}[\sum_{i=1}^N (R_i-1)\sigma_i]^2 \lesssim (\sum_{i=1}^N (R_i-1)^2 + N^{-1/2}[\sum_{i=1}^N (R_i-1)]^2)(\log N)^4$.
(c) $N^{3/2}\mathbb{E}(\overline{\boldsymbol{\sigma}}^6)\lesssim 1$.

REMARK 2.1 (On the minimum eigenvalue of A_N). Note that our results work even when $\lambda_N(A_N)\to -1$, as opposed to stronger spectral gap assumptions such as $\max_{2\le i\le N}|\lambda_i(A_N)|\to 0$. This has been achieved by a new matrix theoretic estimate (see Lemma 6.2) which shows that under (1.7), A_N can have at most one eigenvalue "close" to -1 (see Remark 6.1 for connections to graph theory).

3. Proof of main results. We first state a lemma which will be needed in all parameter regimes.

LEMMA 3.1. *Suppose $\boldsymbol{\sigma}$ is an observation from* (1.1) *for some* A_N *satisfying* (1.4), *and* $\beta > 0$, $B\in\mathbb{R}$.

(a) *Recalling that* $m_i(\boldsymbol{\sigma}) = \sum_{j=1}^N A_N(i,j)\sigma_j$, *we have*

$$\mathbb{E}\left[\sum_{i=1}^N \big(\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma})+B)\big)\tanh(\beta m_i(\boldsymbol{\sigma})+B)\right]^2 \lesssim N.$$

(b) *For any* $\mathbf{c} = (c_1,\ldots,c_n)\in\mathbb{R}^n$ *we have*

$$\log\mathbb{P}\left(\left|\sum_{i=1}^N c_i\big(\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma})+B)\big)\right| > t\right) \lesssim -\frac{t^2}{\|\mathbf{c}\|_2^2}.$$

Here, part (a) follows by invoking [13], Lemma 3.2, and (b) can be obtained by making minor adjustments in the proof of [35], Lemma 1.

3.1. *Proofs of Theorems* 1.1 *and* 1.2. In this section, we will prove Theorems 1.1 and 1.2 using Theorem 1.6, Lemmas 2.1, 3.1, and Proposition 6.1. The statement of Proposition 6.1 is deferred to Section 6.3 as its scope is limited to the Curie–Weiss model introduced in (1.3).

Without loss of generality we may assume that the RHS of (1.6) and (1.8) are bounded by 1, because otherwise the bound is trivial. Recall the definition of $M(\boldsymbol{\sigma})$ for $(\beta, B)\in\Theta_2$ from Lemma 1.2 and set $M(\boldsymbol{\sigma}) = t$ for $(\beta, B)\in\Theta_1$. We have not made the dependence of $M(\boldsymbol{\sigma})$ on (β, B) explicit for notational simplicity. From Section 1.2, recall that $T_N = \sqrt{N}(\overline{\boldsymbol{\sigma}} - M(\boldsymbol{\sigma}))$

and $T_N' = \sqrt{N}(\overline{\boldsymbol{\sigma'}} - M(\boldsymbol{\sigma}'))$ where $\boldsymbol{\sigma}$ is an observation from the Ising model (1.1), and $\boldsymbol{\sigma}'$ is generated as follows: Let I denote a randomly sampled index from $\{1, 2, \dots, N\}$. Given $I = i$, replace σ_i with an independent ± 1 valued random variable σ_i' with mean $\tanh(\beta m_i(\boldsymbol{\sigma}) + B) = \mathbb{E}[\sigma_i | (\sigma_j, j \neq i)]$, and let $\boldsymbol{\sigma}' := (\sigma_1, \dots, \sigma_{i-1}, \sigma_i', \sigma_{i+1}, \dots, \sigma_N)$.

With this setup, a direct computation gives

$$\begin{aligned}\mathbb{E}[T_N - T_N' | T_N] = {} & \frac{1}{N^{3/2}} \sum_{i=1}^N \mathbb{E}[\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma}) + B) | T_N] \\ & - \sqrt{N}\mathbb{E}[M(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma'}) | T_N],\end{aligned} \tag{3.1}$$

where the second term in the RHS above can be expanded as

$$\begin{aligned}&\sum_{i=1}^N \tanh(\beta m_i(\boldsymbol{\sigma}) + B) \\ &\quad = N \tanh(\beta M(\boldsymbol{\sigma}) + B) + \beta(1-t^2)\sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma})) + \sum_{i=1}^N \xi_i (m_i - M(\boldsymbol{\sigma}))^2 \\ &\quad = N M(\boldsymbol{\sigma}) + \beta(1-t^2)\sum_{i=1}^N (\sigma_i - M(\boldsymbol{\sigma})) + \beta(1-t^2) M(\boldsymbol{\sigma}) \sum_{i=1}^N (R_i - 1) \\ &\qquad + \beta(1-t^2)\sum_{i=1}^N (R_i - 1)(\sigma_i - M(\boldsymbol{\sigma})) + \sum_{i=1}^N \xi_i (m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma}))^2\end{aligned} \tag{3.2}$$

for random variables $(\xi_i)_{1 \le i \le N}$ satisfying $\max_{1 \le i \le N} |\xi_i| \lesssim 1$, where the second line uses the identity $M(\boldsymbol{\sigma}) = \tanh(\beta M(\boldsymbol{\sigma}) + B)$. Setting $h_i = \beta(1-t^2)(R_i - 1)$ and plugging (3.2) into (3.1) we get

$$\begin{aligned}\mathbb{E}[T_N - T_N' | T_N] = {} & \underbrace{\frac{T_N}{N}(1 - \beta(1-t^2))}_{g(T_N)} - \underbrace{\frac{1}{N\sqrt{N}} \sum_{i=1}^N \mathbb{E}[\xi_i (m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma}))^2 | T_N]}_{H_1(T_N)} \\ & \underbrace{- \frac{1}{N\sqrt{N}} \mathbb{E}\left[\sum_{i=1}^N h_i (\sigma_i - M(\boldsymbol{\sigma})) \Big| T_N\right]}_{H_2(T_N)} \\ & \underbrace{- \sqrt{N}\mathbb{E}[M(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma}') | T_N] - N^{-3/2}\beta(1-t^2) M(\boldsymbol{\sigma}) \sum_{i=1}^N (R_i - 1).}_{H_3(T_N)}\end{aligned} \tag{3.3}$$

Next, we observe that

$$\begin{aligned}&\limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}\left(|T_N - T_N'| \ge \frac{3}{\sqrt{N}}\right) \\ &\quad \le \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}\left(\sqrt{N}|M(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma}')| \ge \frac{1}{\sqrt{N}}\right) \\ &\quad \le \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}(M(\boldsymbol{\sigma}) \neq M(\boldsymbol{\sigma}')) < 0,\end{aligned} \tag{3.4}$$

where the last inequality for $(\beta, B) \in \Theta_2$ follows on using part (b) of Theorem 1.6 with $\mathcal{E}_N := \{\sum_{i=1}^N \sigma_i \in \{-2, -1, 0, 1, 2\}\}$ along with part (c) of Proposition 6.1 to note that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big(M(\boldsymbol{\sigma}) \neq M(\boldsymbol{\sigma}')\big) \le \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}^{\mathrm{CW}}(\mathcal{E}_N) < 0. \tag{3.5}$$

From (3.3), we choose $g(x) = x(1-\beta(1-t^2))/N$. With this choice, observe that $G(x) := \int_0^x g(y)\,dy = (1-\beta(1-t^2))x^2/2N$. We now set $c_0 := N/(1-t^2)$ and $c_1 := (2\pi\tau)^{-1/2}$, and note the existence of positive constants c_2 and c_3 free of N such that assumptions (H1) and (H3) from [14], page 465, are all satisfied. By a slight variant of [14], Theorem 1.2 (see Section 6.4) and (3.4), we then have

$$\begin{aligned} d_{\mathrm{KS}}(T_N, Z_\tau) \lesssim{}& \mathbb{E}\left|1 - \frac{N}{2(1-t^2)}\mathbb{E}\big[(T_N - T_N')^2 | T_N\big]\right| + \frac{c_1 \max(c_3, 1)}{\sqrt{N}} + \frac{\mathbb{E}|T_N| + 1}{\sqrt{N}} \\ &+ \frac{N}{c_1}\mathbb{E}\left[\sum_{a=1}^{3}\big|H_a(T_N)\big|\right] + \exp(-c_2 N). \end{aligned} \tag{3.6}$$

As we will see later in the proof, the $\exp(-c_2N)$ term is of a smaller order than the other terms in the RHS above. However, we choose to present it in this form so as to emphasize that (3.6) does not follow from a direct application of [14], Theorem 1.2, for $(\beta, B) \in \Theta_2$.

We will now estimate each term in the RHS of (3.6). Proceeding to control $\mathbb{E}|H_1(T_N)|$ we have

$$N\sqrt{N}\big|H_1(T_N)\big| = \left|\sum_{i=1}^N \mathbb{E}\big(\xi_i\big(m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma})\big)^2\big)\Big|T_N\right| \lesssim \mathbb{E}\left(\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma})\big)^2\Big|T_N\right), \tag{3.7}$$

and so

$$N\sqrt{N}\mathbb{E}\big|H_1(T_N)\big| \lesssim \mathbb{E}\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma})\big)^2 \le \eta_N \tag{3.8}$$

using Lemma 2.1, with $\eta_N := \|A_N\|_F^2 + t^2\sum_{i=1}^N (R_i - 1)^2$. Next, we have

$$\begin{aligned} N\sqrt{N}\big|H_2(T_N)\big| \le{}& \left|\mathbb{E}\left(\sum_{i=1}^N h_i\big(\sigma_i - \tan\big(\beta m_i(\boldsymbol{\sigma}) + B\big)|T_N\big)\right)\right| \\ &+ \left|\mathbb{E}\left(\sum_{i=1}^N h_i\big(\tanh\big(\beta m_i(\boldsymbol{\sigma}) + B\big) - \tan\big(\beta M(\boldsymbol{\sigma}) + B\big)\big)\Big|T_N\right)\right| \end{aligned} \tag{3.9}$$

and so

$$\begin{aligned} N\sqrt{N}\mathbb{E}\big|H_2(T_N)\big| &\lesssim \sqrt{\sum_{i=1}^n h_i^2} + \sqrt{\sum_{i=1}^N h_i^2}\sqrt{\mathbb{E}\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma})\big)^2} \\ &\lesssim \sqrt{\sum_{i=1}^N (R_i - 1)^2(1 + \sqrt{\eta_N})} \lesssim \eta_N + \sum_{i=1}^N (R_i - 1)^2, \end{aligned} \tag{3.10}$$

where the penultimate line uses part (b) of Lemma 3.1, and the last line again uses (3.8). Also observe that

$$N\sqrt{N}\big|H_3(T_N)\big| \lesssim N^2\mathbb{E}\big(\big|M(\boldsymbol{\sigma}) - M(\boldsymbol{\sigma}')\big|T_N\big) + t\left|\sum_{i=1}^N (R_i - 1)\right|, \tag{3.11}$$

where the first term has an expectation which is exponentially small in N using (3.5). Finally we have

$$\left|\mathbb{E}\left[1-\frac{N}{2(1-t^2)}(T_N-T_N')^2\Big|T_N\right]\right|\lesssim \mathbb{E}\left|\mathbb{E}\left[1-\frac{(\sigma_I-\sigma_I')^2}{2(1-t^2)}\Big|T_N\right]\right|+N^2\mathbb{E}\left[\left|M(\boldsymbol{\sigma})-M(\boldsymbol{\sigma}')\right|\right].$$

The second term on the RHS above is exponentially small, by (3.5). For the first term on the RHS, note that

$$\begin{aligned}\mathbb{E}[1-(\sigma_I-\sigma_I')^2/2(1-t^2)|\boldsymbol{\sigma}] &= \frac{1}{N(1-t^2)}\sum_{i=1}^N(\mathbb{E}[\sigma_i\sigma_i'|\boldsymbol{\sigma}]-t^2)\\ &\lesssim N^{-1}\left|\sum_{i=1}^N(\sigma_i\tanh(\beta m_i(\boldsymbol{\sigma})+B)-t^2)\right|.\end{aligned}$$

As a result we have

$$\begin{aligned}&\mathbb{E}\left|\mathbb{E}[1-(\sigma_I-\sigma_I')^2/2(1-t^2)|T_N]\right|\\ &\quad\lesssim \mathbb{E}\left|\sum_{i=1}^N(\sigma_i-\tanh(\beta m_i(\boldsymbol{\sigma})+B))\tanh(\beta m_i(\boldsymbol{\sigma})+B)\right|\\ &\qquad+\frac{1}{\sqrt{N}}\sqrt{\sum_{i=1}^N\mathbb{E}(m_i(\boldsymbol{\sigma})-M(\boldsymbol{\sigma}))^2}\\ &\quad\lesssim \frac{1}{\sqrt{N}}+\frac{1}{\sqrt{N}}\sqrt{\sum_{i=1}^N\mathbb{E}(m_i(\boldsymbol{\sigma})-M(\boldsymbol{\sigma}))^2}\le\frac{1+\sqrt{\eta_N}}{\sqrt{N}},\end{aligned}\tag{3.12}$$

where we have used (3.8), and part (a) of Lemma 3.1. We now claim that

$$\mathbb{E}T_N^2\lesssim 1.\tag{3.13}$$

Given this claim, combining the estimates from (3.6), (3.4), (3.8), (3.10), (3.11), and (3.12) we get

$$d_{\mathrm{KS}}(T_N,Z_\tau)\lesssim\frac{1}{\sqrt{N}}+\frac{\eta_N}{\sqrt{N}}+\frac{1}{\sqrt{N}}\sum_{i=1}^N(R_i-1)^2+\frac{t}{\sqrt{N}}\left|\sum_{i=1}^N(R_i-1)\right|,$$

where Z_τ is defined as in Lemma 1.2. The desired bound follows on noting that $\eta_N\gtrsim\|A_N\|_F^2\gtrsim 1$.

It thus suffices to prove (3.13). To this effect, using (3.3) we get

$$\left|\mathbb{E}[T_N-T_N'|T_N]-\frac{T_N}{N}(1-\beta(1-t^2))\right|\lesssim\sum_{a=1}^3|H_a(T_N)|.$$

By multiplying both sides of the above display by NT_N and taking expectation gives

$$\mathbb{E}[T_N^2]\lesssim|\mathbb{E}[N(T_N-T_N')T_N]|+\mathbb{E}\left[N|T_N|\left(\sum_{a=1}^3|H_a(T_N)|\right)\right],$$

where we have used the fact that $\beta(1-t^2)<1$. This follows from Lemma 1.1, parts (b) and (c), on noting that $\phi'(t)=1-\beta(1-t^2)$ where $\phi(\cdot)$ is defined as in Lemma 1.1. By the exchangeability of T_N and T_N' we have

$$\mathbb{E}[N(T_N-T_N')T_N]=\mathbb{E}[N(T_N'-T_N)T_N']=\frac{1}{2}\mathbb{E}[N(T_N-T_N')^2]\lesssim 1.$$

Also, from (3.7), (3.9), and (3.11) we have

$$N\sum_{a=1}^{3}\mathbb{E}\big[H_a(T_N)\big]^2 \lesssim \frac{\eta_N + \sum_{i=1}^{N}(R_i-1)^2}{\sqrt{N}} \lesssim 1,$$

where the last bound uses the fact that the RHS of (1.6) and (1.8) are bounded. Using Chebyshev's inequality then gives

$$\mathbb{E}(T_N^2) \lesssim 1 + \sqrt{\mathbb{E}(T_N^2)}\sqrt{\sum_{a=1}\mathbb{E}\big(NH_a(T_N)\big)^2} \lesssim 1 + \sqrt{\mathbb{E}(T_N^2)}$$

which implies $\mathbb{E}(T_N^2) \lesssim 1$. This verifies (3.13), and hence completes the proof of the theorem.

3.2. *Proof of Theorem* 1.4. We will now prove Theorem 1.4 using Lemmas 2.1 and 2.2 whose proofs have been deferred to Section 5.

PROOF. Without loss of generality we can assume that the RHS of (1.12) is bounded as before. As in the proof of the previous theorems, it suffices to bound the RHS of (3.6), but with $t = M(\boldsymbol{\sigma}) = 0$ which implies $H_3(T_N) = 0$. To begin, use (3.12) to get

$$\mathbb{E}\left|\mathbb{E}\left[1 - \frac{N}{2}(T_N - T_N')^2\middle| T_N\right]\right| \lesssim \frac{\|A_N\|_F^2}{N} + \frac{1}{\sqrt{N}}, \tag{3.14}$$

using (2.1), which allows us to replace η_N in the previous proof by $\|A_N\|_F^2$. Proceeding to bound $\mathbb{E}|H_1(T_N)|$, use the first equality of (3.7) along with Cauchy–Schwarz inequality to note that

$$\begin{aligned} N\sqrt{N}\mathbb{E}\big|H_1(T_N)\big| &\lesssim \mathbb{E}\max_{1\le i\le N}\big|m_i(\boldsymbol{\sigma})\big|\sum_{i=1}^{N} m_i(\boldsymbol{\sigma})^2 \\ &\le \sqrt{\mathbb{E}\max_{1\le i\le N} m_i(\boldsymbol{\sigma})^2}\sqrt{\mathbb{E}\left(\sum_{i=1}^{N} m_i(\boldsymbol{\sigma})^2\right)^2} \lesssim \|A_N\|_F^2\sqrt{\alpha_N\log N}, \end{aligned} \tag{3.15}$$

where the last inequality uses part (a) of Lemma 2.2. Finally, for $\mathbb{E}|H_2(T_N)|$ we have

$$N\sqrt{N}\mathbb{E}\big|H_2(T_N)\big| \le \mathbb{E}\left|\sum_{i=1}^{N}(R_i-1)\sigma_i\right| \lesssim \sqrt{\left(\sum_{i=1}^{N}(R_i-1)^2\right)\big[1 + \|A_N\|\alpha_N\log N\big]},$$

where we use part (b) of Lemma 2.2. Plugging in the above bounds in (3.6), we have

$$d_{\mathrm{KS}}(T_N, Z_\tau) \lesssim \frac{1+\mathbb{E}(T_N^2)}{\sqrt{N}} + \frac{\|A_N\|_F^2\sqrt{\alpha_N\log N}}{\sqrt{N}} + \big[1 + \|A_N\|\alpha_N\log N\big]\sqrt{\frac{\sum_{i=1}^{N}(R_i-1)^2}{N}},$$

with Z_τ defined as in Lemma 1.2. The claimed bound follows immediately, if we can verify (3.13). But the proof of this is the same as in the previous theorem, and so we are done. □

3.3. *Proof of Theorem* 1.3. In this section, we will use Lemmas 2.1 and 2.3 to prove Theorem 1.3. The proofs of the aforementioned lemmas are presented in Section 5.

PROOF. With $(\boldsymbol{\sigma}, \boldsymbol{\sigma}')$ the usual exchangeable pair, setting $T_N := N^{1/4}\overline{\boldsymbol{\sigma}}$ and $T_N := N^{1/4}\overline{\boldsymbol{\sigma}}'$ we have

$$\begin{aligned} \mathbb{E}\big[T_N - T_N'|\boldsymbol{\sigma}\big] &= N^{-3/4}\big(\overline{\boldsymbol{\sigma}} - \tanh(\overline{\boldsymbol{\sigma}})\big) + N^{-3/4}\big(\tanh(\overline{\boldsymbol{\sigma}}) - \tanh\big(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\big) \\ &\quad + N^{-7/4}\sum_{i=1}^{N}\big(\tanh\big(m_i(\boldsymbol{\sigma})\big) - \tanh\big(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\big). \end{aligned}$$

Using Taylor's expansion, this gives

$$\begin{aligned}
&\left|\mathbb{E}[T_N - T_N'|\boldsymbol{\sigma}] - N^{-3/4}(\overline{\boldsymbol{\sigma}} - \tanh(\overline{\boldsymbol{\sigma}}))\right| \\
&\lesssim N^{-3/4}|\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})| + N^{-7/4}|\overline{\boldsymbol{\sigma}}| \sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2 \\
&\quad + N^{-7/4}\left|\sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^3\right|,
\end{aligned} \tag{3.16}$$

and so we have $\mathbb{E}[T_N - T_N'|T_N] = g(T_N) + H(T_N)$ where $g(x) = N^{-3/2}x^3/3$, and $H(T_N)$ satisfies

$$\begin{aligned}
\mathbb{E}[|H(T_N)|] &\lesssim N^{-2}\mathbb{E}[|T_N|^5] + N^{-3/4}\mathbb{E}[|\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|] \\
&\quad + N^{-2}\mathbb{E}\left[|T_N| \sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2\right] \\
&\quad + N^{-7/4}\mathbb{E}\left[\left|\sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^3\right|\right].
\end{aligned}$$

Invoking [14], Theorem 1.2, with $G(x) := \int_0^x g(t)\,dt = N^{-3/2}x^4/12$ we have

$$d_{\mathrm{KS}}(T_N, W) \lesssim \mathbb{E}\left|1 - \frac{N^{3/2}}{2}\mathbb{E}[(T_N - T_N')^2|T_N]\right| + N^{3/2}\mathbb{E}[|H(T_N)|] + N^{-3/4}\mathbb{E}|T_N|^3. \tag{3.17}$$

By part (c) of Lemma 2.3 we have $\mathbb{E}[|T_N|^5] \lesssim 1$. Set

$$\delta_N := \sum_{i=1}^{N} (R_i - 1)^2 + N^{-1/2}\left[\sum_{i=1}^{N} (R_i - 1)\right]^2,$$

and use part (b) of Lemma 2.3 and the Cauchy–Schwarz inequality to get

$$\mathbb{E}[|\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|] \lesssim \sqrt{\mathbb{E}(\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2} \lesssim N^{-1}(\log N)^2\sqrt{\delta_N}.$$

Similarly, by the Cauchy–Schwarz inequality and part (c) of Lemma 2.1 along with part (a) of Lemma 2.3, we get

$$\begin{aligned}
\mathbb{E}\left[|T_N| \sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2\right] &\le \sqrt{\mathbb{E}(T_N)^2}\sqrt{\mathbb{E}\left(\sum_{i=1}^{n} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2\right)^2} \\
&\lesssim \varepsilon_N, \\
\mathbb{E}\left[\sum_{i=1}^{N} |m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|^3\right] &\le \sqrt{\mathbb{E}\max_{1 \le i \le N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2}\sqrt{\mathbb{E}\left(\sum_{i=1}^{N} (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2\right)^2} \\
&\lesssim r_N\varepsilon_N,
\end{aligned}$$

where ε_N is as in the statement of Theorem 1.3. Combining the above observations, we get

$$N^{3/2}\mathbb{E}[|H(T_N)|] \lesssim N^{-1/2} + N^{-1/4}(\log N)^2\sqrt{\delta_N} + N^{-1/4}r_N\varepsilon_N. \tag{3.18}$$

Finally, we have

$$
\begin{aligned}
&\mathbb{E}\left|1-\frac{N^{3/2}}{2}\mathbb{E}[(T_N-T_N')^2|T_N]\right| \\
&\lesssim \frac{1}{N}\mathbb{E}\left|\sum_{i=1}^N \sigma_i \tanh m_i(\boldsymbol{\sigma})\right| \\
&\lesssim \frac{1}{N}\mathbb{E}\left|\sum_{i=1}^N (\sigma_i-\tanh m_i(\boldsymbol{\sigma}))\tanh m_i(\boldsymbol{\sigma})\right| + \frac{1}{N}\mathbb{E}\sum_{i=1}^N \left(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\right)^2 + \mathbb{E}\overline{\boldsymbol{\sigma}}^2 \\
&\lesssim \frac{1}{\sqrt{N}} + \frac{\varepsilon_N}{N} + \frac{1}{\sqrt{N}},
\end{aligned} \tag{3.19}
$$

where the last inequality follows from part (a) of Lemma 3.1, part (c) of Lemma 2.1, and part (c) of Lemma 2.3. Combing (3.18) and (3.19) along with (3.17) gives

$$
d_{\mathrm{KS}}(T_N, W) \lesssim \frac{1}{\sqrt{N}} + \frac{\sqrt{\delta_N}(\log N)^2}{N^{1/4}} + \frac{r_N\varepsilon_N}{N^{1/4}},
$$

as desired, with W defined as in Lemma 1.2. □

4. Proofs of Theorems 1.5, 1.6, and Lemma 2.1. We will need the following proposition which expresses the Curie–Weiss model as a mixture of i.i.d. random variables, first shown in [34], Lemma 3.

PROPOSITION 4.1. *Let $\boldsymbol{\sigma}$ be an observation from the Curie–Weiss model in* (1.3). *Given $\boldsymbol{\sigma}$, let W_N be a Gaussian random variable with mean $\overline{\boldsymbol{\sigma}}$ and variance $(N\beta)^{-1}$. Then the following conclusions hold*:

(a) *Given W_N, the random variables $(\sigma_1, \sigma_2, \dots, \sigma_N)$ are i.i.d. with mean $\widetilde{W}_N := \tanh(\beta W_N + B)$.*

(b) *The marginal density of W_N is proportional to* $\exp(-Nf(w))$, *where* $f(w) = \frac{\beta w^2}{2} - \log\cosh(\beta w + B)$.

We state two more lemmas necessary for proving the results of this section, the proofs of which we defer to Section 6. The first lemma is a version of the Hanson–Wright inequality, which controls exponential moment of quadratic forms of binary random variables.

LEMMA 4.1. *Suppose $X_1, X_2, \dots, X_N$, $N \geq 1$ are i.i.d. ± 1 valued random variables such that $\mathbb{E}[X_1] = \mu$ where $\mu \in (-1, 1)$. Define $s_\mu := 2\mu/(\log(1+\mu) - \log(1-\mu))$ with s_0 being 1. Also assume that D_N is a $N \times N$ symmetric matrix such that $s_\mu \limsup_{N\to\infty} \lambda_1(D_N) < 1$. Then, given any vector $\mathbf{c}^\top := (c_1, c_2, \dots, c_N)$, we get*

$$
\log\left\{\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i,j=1}^N D_N(i,j)\widetilde{X}_i\widetilde{X}_j + \sum_{i=1}^N c_i\widetilde{X}_i\right)\right]\right\} \lesssim \mathrm{Tr}^+(D_N) + \|D_N\|_F^2 + \sum_{i=1}^N c_i^2,
$$

where $\widetilde{X}_i = X_i - \mu$ for $1 \leq i \leq N$, and $\mathrm{Tr}^+(D_N) = \sum_{i=1}^N \max(D_N(i,i), 0)$.

The second lemma gives a quantitative estimate which allows us to neglect the region where $\widetilde{W}_N$ is not close to t.

LEMMA 4.2. *Suppose* (1.4), (1.5), *and* (1.7) *holds, and further assume that* $\|A_N\|_F^2 = o(N)$, $\sum_{i=1}(R_i - 1) = o(N)$. *Also, let* V_N *be any random variable such that* $V_N \le cN$ *for some fixed* $c > 0$, *and* $\varepsilon > 0$ *be fixed. Recalling* $\mathcal{A}_N := A_N - \mathbf{1}\mathbf{1}^\top/N$, *for any* $(\beta, B) \in \Theta_2 \cup \Theta_3$ *there exists* $\delta = \delta(\varepsilon, c, \beta) > 0$ *such that*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}^{\mathrm{CW}}\Big[\exp\Big(\delta V_N + \frac{\beta}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\Big)\mathbb{1}\big(\big|\widetilde{W}_N - M(\boldsymbol{\sigma})\big| \ge \epsilon\big)\Big] < 0. \tag{4.1}$$

Additionally, the proofs of Theorems 1.5, 1.6, and Lemma 2.1, require Proposition 6.1 which is stated in Section 6.3.

PROOF OF THEOREM 1.5. (a) To begin, note that

$$\frac{\beta}{2}\boldsymbol{\sigma}^\top A_N \boldsymbol{\sigma} + B\sum_{i=1}^N \sigma_i = \frac{\beta}{2}(\boldsymbol{\sigma} - t)^\top A_N(\boldsymbol{\sigma} - t) + \sum_{i=1}^N (\beta t R_i + B)\sigma_i - (\beta t^2/2)\mathbf{1}^\top A_N \mathbf{1}.$$

Recall that $\mathcal{M}_N(\beta, B) = N\{\frac{\beta t^2}{2} + Bt - I(t)\} + \frac{\beta t^2}{2}\sum_{i=1}^N (R_i - 1)$ as in (1.13). The above display then gives

$$\begin{aligned} &\frac{Z_N(\beta, B)}{\exp(\mathcal{M}_N(\beta, B))} \\ &\quad = \mathbb{E}^{\mathbb{Q}} \exp\Bigg(\frac{\beta}{2}\sum_{i,j=1}^N (\sigma_i - t)A_N(i,j)(\sigma_j - t) + \beta t\sum_{i=1}^N (R_i - 1)(\sigma_i - t)\Bigg), \end{aligned} \tag{4.2}$$

where $\mathbb{Q}$ is the measure induced by N independent ± 1 valued random variables with mean t (as defined in (1.14)). In this case with $D_N = \beta A_N$ we have

$$s_t \limsup_{N\to\infty} \lambda_1(D_N) = \beta s_t \limsup_{N\to\infty} \lambda_1(A_N) \le \beta s_t = \frac{\beta t}{\beta t + B} < 1,$$

for $(\beta, B) \in \Theta_{12}$. If $(\beta, B) \in \Theta_{11}$, then we have $t = 0$, and $s_0 = 1$, and so with $D_N = \beta A_N$ as before, we have $s_t \limsup_{N\to\infty} \lambda_1(D_N) = \beta < 1$. Thus in both cases Lemma 4.1 is applicable with $D_N = \beta A_N$, $c_i = \beta t(R_i - 1)$, which using (4.2) gives

$$\begin{aligned} &\log \mathbb{E}^{\mathbb{Q}} \exp\Bigg(\frac{\beta}{2}\sum_{i,j=1}^N (\sigma_i - t)A_N(i,j)(\sigma_j - t) + \beta t\sum_{i=1}^N (R_i - 1)(\sigma_i - t)\Bigg) \\ &\quad \lesssim \|A_N\|_F^2 + t^2\sum_{i=1}^N (R_i - 1)^2. \end{aligned} \tag{4.3}$$

The conclusion of part (a) follows from this combined with (4.2).

(b) Define

$$\begin{aligned} Y_N :=\ & (\boldsymbol{\sigma} - \widetilde{W}_N)^\top \mathcal{A}_N(\boldsymbol{\sigma} - \widetilde{W}_N) + 2\widetilde{W}_N \sum_{i=1}^N (R_i - 1)(\sigma_i - \widetilde{W}_N) \\ & + (\widetilde{W}_N^2 - t^2)\sum_{i=1}^N (R_i - 1), \end{aligned} \tag{4.4}$$

and note that $\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma} = Y_N + t^2 \sum_{i=1}^N (R_i - 1)$. Using this, with $J_{N,\epsilon} := \{|t| - \varepsilon \le |\widetilde{W}_N| \le |t| + \varepsilon\}$ for some $\epsilon > 0$, by a similar calculation as in part (a) we have

$$\begin{aligned}\frac{Z_N(\beta, B)}{Z_N^{\mathrm{CW}}(\beta, B)} &= \mathbb{E}^{\mathrm{CW}} \exp\left(\frac{\beta}{2} \boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\right) \\ &= \mathbb{E}^{\mathrm{CW}}\left[\exp\left(\frac{\beta}{2} \boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\right) \mathbb{1}(J_{N,\epsilon}^c)\right] \\ &\quad + \exp\left(\frac{\beta}{2} t^2 \sum_{i=1}^N (R_i - 1)\right) \mathbb{E}^{\mathrm{CW}}\left[e^{\frac{\beta}{2} Y_N} \mathbb{1}(J_{N,\epsilon})\right].\end{aligned} \tag{4.5}$$

The first term in the right-hand side of (4.5) is $o(1)$ by invoking Lemma 4.2 with $\delta = 0$. For the second term, by Proposition 4.1, the inner (conditional) expectation is taken with respect to i.i.d. ± 1 valued random variables with mean $\widetilde{W}_N$. In this regime $\beta s_t = \beta t / \beta t = 1$. But since $\limsup_{N \to \infty} \lambda_1(\mathcal{A}_N) < 1$ by (1.7), on the set $J_{N,\varepsilon}$ we have

$$\limsup_{N \to \infty} s_{\widetilde{W}_N} \lambda_1(\beta \mathcal{A}_N) \le \limsup_{N \to \infty} \sup_{\mu \in J_{N,\varepsilon}} s_\mu \lambda_1(\beta \mathcal{A}_N) < 1$$

for ε small enough. Therefore, Lemma 4.1 is applicable with $D_N = \beta \mathcal{A}_N$ and $c_i = 2\widetilde{W}_N(R_i - 1)$ to give

$$\log \mathbb{E}^{\mathrm{CW}}\left(e^{\frac{\beta}{2} Y_N} \mathbb{1}(J_{N,\epsilon}) | \widetilde{W}_N\right) \le C\left\{\|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2\right\} + \frac{\beta}{2}\left|\left(\widetilde{W}_N^2 - t^2\right) \sum_{i=1}^N (R_i - 1)\right|$$

for some $C < \infty$, which on taking another expectation gives

$$\begin{aligned}\log \mathbb{E}^{\mathrm{CW}}\left(e^{\frac{\beta}{2} Y_N} \mathbb{1}(J_{N,\epsilon})\right) &\le C\left\{\|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2\right\} + \log \mathbb{E} e^{\frac{\beta}{2}|(\widetilde{W}_N^2 - t^2) \sum_{i=1}^N (R_i - 1)|} \\ &\lesssim \|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2,\end{aligned} \tag{4.6}$$

where the last step uses part (b) of Proposition 6.1. This along with (4.5) gives

$$\log Z_N(\beta, B) - \log Z_N^{\mathrm{CW}}(\beta, B) - \frac{\beta t^2}{2} \sum_{i=1}^N (R_i - 1) \lesssim \|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2,$$

from which the desired conclusion follows by another application of part (a) of Proposition 6.1 to note that $\log Z_N^{\mathrm{CW}}(\beta, B) - N[\frac{\beta t}{2} + Bt - I(t)] \lesssim 1$.

(c) In this regime we have $t = 0$, and so $s_t = s_0 = 1$, and $\beta s_0 = 1$. As in the proof of part (b), the first term in the RHS of (4.5) is $o(1)$ invoking Lemma 4.2 with $\delta = 0$. For handling the second term, invoking (1.7) gives

$$\limsup_{N \to \infty} s_{\widetilde{W}_N} \lambda_1(\mathcal{A}_N) \le \limsup_{N \to \infty} \sup_{\mu \in J_{N,\varepsilon}} s_\mu \lambda_1(\mathcal{A}_N) < 1$$

for ε small enough. Also Lemma 4.1 with $D_N = \mathcal{A}_N$ and $c_i = 2\widetilde{W}_N(R_i - 1)$ gives

$$\log \mathbb{E}^{\mathrm{CW}}\left(e^{\frac{\beta}{2} Y_N} \mathbb{1}(J_{N,\epsilon}) | \widetilde{W}_N\right) \le C\left\{\|A_N\|_F^2 + \widetilde{W}_N^2 \sum_{i=1}^N (R_i - 1)^2\right\} + \frac{\beta}{2} \widetilde{W}_N^2 \sum_{i=1}^N (R_i - 1)$$

for some $C < \infty$ free of N. This, on taking another expectation along with (4.5) gives

$$
\begin{aligned}
&\log \mathbb{E}^{\mathrm{CW}}\big(e^{\frac{\beta}{2}Y_N}\mathbb{1}(J_{N,\epsilon})\big) \\
&\le C\|A_N\|_F^2 + \log \mathbb{E}\exp\left(C\widetilde{W}_N^2\sum_{i=1}^N (R_i-1)^2 + \frac{\beta}{2}\widetilde{W}_N^2\sum_{i=1}^N (R_i-1)\right) \\
&\lesssim \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i-1) + \sum_{i=1}^N (R_i-1)^2\right]^2,
\end{aligned}
\tag{4.7}
$$

where the last bound uses part (b) of Proposition 6.1. Combining (4.5) and (4.7) gives

$$
\log Z_N(\beta, B) - \log Z_N^{\mathrm{CW}}(\beta, B) \lesssim \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i-1)\right]^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i-1)^2\right]^2.
$$

We incur an additional log factor in the final answer because $\log Z_N^{\mathrm{CW}}(\beta, B) - N[\frac{\beta t}{2} + Bt - I(t)] \lesssim \log N$ by part (a) of Proposition 6.1. □

PROOF OF THEOREM 1.6. (a) Using a similar calculation as in (4.5), we get

$$
\begin{aligned}
&\big(c(N)\big)^{-1}\mathbb{P}(\boldsymbol{\sigma} \in \mathcal{E}_N) \\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\frac{\beta}{2}\sum_{i,j=1}^N (\sigma_i - t)A_N(\sigma_j - t) + \beta t\sum_{i=1}^N (R_i-1)(\sigma_i - t)\right)\mathbb{1}(\boldsymbol{\sigma} \in \mathcal{E}_N)\right],
\end{aligned}
\tag{4.8}
$$

where the deterministic sequence $c(N)$ satisfies

$$
c(N) = \frac{\exp(\beta t^2(\mathbf{1}^\top A_N\mathbf{1} - N))(\exp(\beta t + B) + \exp(-\beta t - B))^N}{Z_N(\beta, B)\exp((\beta t^2/2)\mathbf{1}^\top A_N\mathbf{1})} \le 1,
$$

on invoking the mean-field lower bound (1.13). Next, by using Hölder's inequality with exponent p (to be chosen later), the left-hand side of (4.8) can be bounded above by

$$
\begin{aligned}
&\left\{\mathbb{E}^{\mathbb{Q}}\exp\left(\frac{\beta(1+p)}{2}\sum_{i,j=1}^N (\sigma_i - t)A_N(\sigma_j - t) + \beta t(1+p)\sum_{i=1}^N (R_i-1)(\sigma_i - t)\right)\right\}^{\frac{1}{1+p}} \\
&\times \big(\mathbb{Q}(\mathcal{E}_N)\big)^{\frac{p}{1+p}}.
\end{aligned}
\tag{4.9}
$$

Using arguments similar to the derivation of (4.3) shows that for p small enough we have

$$
\begin{aligned}
&\log \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\frac{\beta(1+p)}{2}\sum_{i,j}(\sigma_i - t)A_N(\sigma_j - t) + \beta t(1+p)\sum_{i=1}^N (R_i-1)(\sigma_i - t)\right)\right] \\
&\lesssim \|A_N\|_F^2 + t^2\sum_{i=1}^N (R_i-1)^2.
\end{aligned}
$$

Combining this along with (4.8) and (4.9) gives the desired conclusion.

(b) With Y_N as in (4.4), using a similar calculation as in the derivation of (4.5) we can bound $P(\boldsymbol{\sigma} \in \mathcal{E}_N)$ by

$$\begin{aligned}
&\frac{Z_N^{\mathrm{CW}}(\beta, B)}{Z_N(\beta, B)} \mathbb{E}^{\mathrm{CW}} e^{\frac{\beta}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}} \mathbb{1}(\boldsymbol{\sigma} \in \mathcal{E}_N) \\
&\quad \le \frac{Z_N^{\mathrm{CW}}(\beta, B)}{Z_N(\beta, B)} \big[\mathbb{E}^{\mathrm{CW}} e^{\frac{\beta}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}} \mathbb{1}\big(\boldsymbol{\sigma} \in J_{N,\varepsilon}^c\big) \\
&\qquad + e^{\frac{\beta t^2}{2}\sum_{i=1}^N (R_i - 1)} \mathbb{E}^{\mathrm{CW}} e^{\frac{\beta}{2} Y_N} \mathbb{1}(\boldsymbol{\sigma} \in \mathcal{E}_N) \mathbb{1}(J_{N,\varepsilon})\big].
\end{aligned} \tag{4.10}$$

For controlling the ratio of partition functions in the RHS of (4.10), use the mean-field approximation (1.13) to get a lower bound for $\log Z_N(\beta, B)$, whereas part (a) of Proposition 6.1 gives $\log Z_N^{\mathrm{CW}}(\beta, B) - N[\frac{\beta t}{2} + Bt - I(t)] \lesssim 1$. Combining these two observations, we get

$$\log Z_N^{\mathrm{CW}}(\beta, B) - \log Z_N(\beta, B) + \frac{\beta t^2}{2}\sum_{i=1}^N (R_i - 1) \lesssim 1. \tag{4.11}$$

Also, the first term inside the parenthesis in the RHS of (4.10) is exponentially small in N, by invoking Lemma 4.2 with $\delta = 0$. Proceeding to control the second term in the RHS of (4.10) we have

$$\mathbb{E}^{\mathrm{CW}} e^{\frac{\beta}{2} Y_N} \mathbb{1}(\boldsymbol{\sigma} \in \mathcal{E}_N) \mathbb{1}(J_{N,\varepsilon}) \le \big[\mathbb{E}^{\mathrm{CW}} e^{\frac{\beta(1+p)}{2} Y_N} \mathbb{1}(J_{N,\varepsilon})\big]^{\frac{1}{1+p}} \big[\mathbb{P}^{\mathrm{CW}}(\boldsymbol{\sigma} \in \mathcal{E}_N)\big]^{\frac{p}{1+p}}, \tag{4.12}$$

where the last step uses Holder's inequality for any $p > 0$. For controlling the first term inside the bracket in the RHS of (4.12), by choosing $p > 0$ small enough and repeating the same argument as in the derivation of (4.6), we get

$$\log \mathbb{E}^{\mathrm{CW}}\big(e^{\frac{\beta(1+p)}{2} Y_N}\big) \mathbb{1}(J_{N,\varepsilon}) \lesssim \|A_N\|_F^2 + \sum_{i=1}^N (R_i - 1)^2. \tag{4.13}$$

Combining (4.10), (4.11), (4.12), and (4.13), the desired conclusion follows.

(c) All steps of part (b) above go through verbatim, except the RHS of (4.11) gets replaced by $\log N$ (by part (a) of Proposition 6.1), and (4.13) is replaced by (cf. (4.7))

$$\log \mathbb{E}^{\mathrm{CW}} e^{\frac{\beta(1+p)}{2} Y_N} \mathbb{1}(J_{N,\varepsilon}) \lesssim \|A_N\|_F^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)^2\right]^2 + \frac{1}{N}\left[\sum_{i=1}^N (R_i - 1)\right]^2. \tag{4.14}$$

Combining this with (4.10) and (4.14) gives the desired conclusion. □

PROOF OF LEMMA 2.1. (a) Invoking Theorem 1.6 and changing δ if necessary, it suffices to show the desired conclusion under $\mathbb{Q}$, where $\mathbb{Q}$ is the i.i.d. measure induced by N ± 1 valued random variables with mean t, as defined in (1.14). A direct calculation shows that $m_i(\boldsymbol{\sigma}) - t$ equals $\sum_{j=1}^N A_N(i, j)(\sigma_j - t) + t(R_i - 1)$, and so

$$\begin{aligned}
\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - t\big)^2 &\le 2\sum_{i=1}^N \left[\sum_{j=1}^N A_N(i, j)(\sigma_j - t)\right]^2 + 2t^2 \sum_{i=1}^N (R_i - 1)^2 \\
&= 2\sum_{i=1}^N \sum_{j=1}^N (A_N^2)(i, j)(\sigma_i - t)(\sigma_j - t) + 2t^2 \sum_{i=1}^N (R_i - 1)^2.
\end{aligned} \tag{4.15}$$

It therefore suffices to control the exponential moment of the first term in the RHS of (4.15). Since $\limsup_{N\to\infty} \lambda_1(A_N^2) \le 1$, for any $\delta \in (0, 1/2)$, using Lemma 4.1 with $D_N = \delta A_N^2$ and

$c_i = 0$ we have

$$\log \mathbb{E}^{\mathbb{Q}} \exp\big(\delta(\boldsymbol{\sigma} - t)^\top A_N^2 (\boldsymbol{\sigma} - t)\big) \lesssim \|A_N^2\|_F^2 = \sum_{i=1}^N \lambda_i^4 \lesssim \sum_{i=1}^N \lambda_i^2 = \|A_N\|_F^2.$$

This gives the desired conclusion.

(c) By invoking Theorem 1.6, it suffices to show the desired conclusion under the Curie–Weiss model. Start by noting that $\overline{\boldsymbol{m}}(\boldsymbol{\sigma}) = \frac{1}{N}\sum_{i=1}^N R_i \sigma_i$, and so

$$m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}) = \sum_{j=1}^N A_N(i,j)(\sigma_j - \widetilde{W}_N) + \frac{1}{N}\sum_{i=1}^N R_i(\sigma_i - \widetilde{W}_N) + \widetilde{W}_N(R_i - \overline{R}).$$

This shows that $\sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2$ is bounded by

$$\begin{aligned} &3\sum_{i=1}^N \Bigg[\sum_{j=1}^N A_N(i,j)(\sigma_j - \widetilde{W}_N)\Bigg]^2 + \frac{3}{N}\Bigg[\sum_{i=1}^N R_i(\sigma_i - \widetilde{W}_N)\Bigg]^2 + 3\widetilde{W}_N^2 \sum_{i=1}^N (R_i - \overline{R})^2 \\ &\quad \le 3\sum_{i,j=1}^N \bigg(\big(A_N^2\big)(i,j) + \frac{3}{N}R_i R_j\bigg)(\sigma_i - \widetilde{W}_N)(\sigma_j - \widetilde{W}_N) + 3\widetilde{W}_N^2 \sum_{i=1}^N (R_i - 1)^2. \end{aligned} \tag{4.16}$$

Conditioning on $\widetilde{W}_N$, we now control the exponential moment of the first term in the RHS of the above display under the Curie–Weiss model. By Proposition 4.1, under the Curie–Weiss model, given $\widetilde{W}_N$, the random vector $(\sigma_1, \dots, \sigma_N)$ are i.i.d. with mean $\widetilde{W}_N$. Note that

$$\limsup_{N\to\infty} \lambda_1\bigg(A_N^2 + \frac{3}{N}\mathbf{R}\mathbf{R}^\top\bigg) \le \limsup_{N\to\infty} \lambda_1\big(A_N^2\big) + \limsup_{N\to\infty} \frac{3}{N}\lambda_1\big(\mathbf{R}\mathbf{R}^\top\big) \lesssim 1,$$

$$\bigg\|\frac{1}{N}\mathbf{R}\mathbf{R}^\top\bigg\|_F^2 = \frac{1}{N^2}\Bigg(\sum_{i=1}^N R_i^2\Bigg)^2 \lesssim 1,$$

where the last display follows from the assumption that $\max_{1\le i\le N} R_i \lesssim 1$ by (1.5). Based on these observations, on invoking Lemma 4.1 with $c_i = 0$, $D_N = \delta(A_N^2 + \frac{3}{N}\mathbf{R}\mathbf{R}^\top)$ for δ small enough, we get

$$\log \mathbb{E}^{\mathrm{CW}} e^{\delta \sum_{i=1}^N (m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2} - \log \mathbb{E}^{\mathrm{CW}} e^{3\delta \widetilde{W}_N^2 \sum_{i=1}^N (R_i - 1)^2} \lesssim \|A_N^2\|_F^2 + \mathrm{tr}\big(A_N^2\big) \lesssim \|A_N\|_F^2,$$

from which the desired conclusion follows on noting that

$$\log \mathbb{E}^{\mathrm{CW}} e^{3\delta \widetilde{W}_N^2 \sum_{i=1}^N (R_i - 1)^2} \lesssim \frac{1}{N}\Bigg[\sum_{i=1}^N (R_i - 1)^2\Bigg]^2,$$

which follows from part (b) of Proposition 6.1.

(b) To begin, note that

$$\begin{aligned} \sum_{i=1}^N \big(m_i - M(\boldsymbol{\sigma})\big)^2 &\lesssim \sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2 + \frac{1}{N}\Bigg[\sum_{i=1}^N R_i(\sigma_i - \widetilde{W}_N)\Bigg]^2 \\ &\quad + \big(\widetilde{W}_N - M(\boldsymbol{\sigma})\big)^2 \Bigg|\sum_{i=1}^N (R_i - 1)\Bigg|. \end{aligned} \tag{4.17}$$

By Hölder's inequality, it suffices to bound the exponential moments of the three terms of the above display at some $\delta > 0$. Exponential moment of the third term in the RHS of (4.17) is bounded by part (b) of Proposition 6.1, as $\sum_{i=1}^N |R_i - 1| = o(N)$. Proceeding to bound the sum of the first two terms, use (4.16) to get

$$\sum_{i=1}^N \big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2 + \frac{1}{N}\Bigg[\sum_{i=1}^N R_i(\sigma_i - \widetilde{W}_N)\Bigg]^2$$
$$\leq \sum_{i,j=1}^N \bigg(\big(A_N^2\big)(i,j) + \frac{4}{N}R_i R_j\bigg)(\sigma_i - \widetilde{W}_N)(\sigma_j - \widetilde{W}_N) + 3\sum_{i=1}^N (R_i - 1)^2,$$

and so it suffices to bound

$$\log \mathbb{E}^{\mathrm{CW}} \exp\Bigg(\delta \sum_{i,j=1}^N \bigg(\big(A_N^2\big)(i,j) + \frac{4}{N}R_i R_j\bigg)(\sigma_i - \widetilde{W}_N)(\sigma_j - \widetilde{W}_N)\Bigg)$$

for δ small enough. But this follows on invoking Lemma 4.1 with $D_N = \delta(A_N^2 + \frac{4}{N}\mathbf{R}\mathbf{R}^\top)$ and $c_i = 0$ to get

$$\log \mathbb{E}^{\mathrm{CW}} \exp\Bigg(\delta \sum_{i,j=1}^N \bigg(\big(A_N^2\big)(i,j) + \frac{4}{N}R_i R_j\bigg)(\sigma_i - \widetilde{W}_N)(\sigma_j - \widetilde{W}_N)\Bigg)$$
$$\lesssim \|A_N^2\|_F^2 + \mathrm{tr}\big(A_N^2\big) \lesssim \|A_N\|_F^2,$$

which completes the proof of part (b). $\square$

5. Proofs of Lemmas 2.2 and 2.3.

PROOF OF LEMMA 2.2. (a) To begin, note that it suffices to prove the bound for λ large enough. To this effect, using part (b) of Lemma 3.1 we have the existence of a constant M free of N, such that for all $\lambda > 0$ we have

$$\mathbb{P}\Bigg(\bigg|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(i,j)\tanh\big(\beta m_j(\boldsymbol{\sigma})\big)\bigg| > \lambda\sqrt{\log N \sum_{j=1}^N A_N(i,j)^2}\Bigg) \leq 2e^{-\frac{\lambda^2 \log N}{M}},$$

which on using a union bound with $\alpha_N = \max_{1\leq i\leq N}\sum_{j=1}^N A_N(i,j)^2$ (as in Theorem 1.4) gives

$$\mathbb{P}\Bigg(\max_{1\leq i\leq N}\bigg|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(i,j)\tanh\big(\beta m_j(\boldsymbol{\sigma})\big)\bigg| > \lambda\sqrt{\alpha_N \log N}\Bigg) \leq 2Ne^{-\frac{\lambda^2 \log N}{M}}.$$

On the set $\{\max_{1\leq i\leq N}|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(i,j)\tanh(\beta m_j(\boldsymbol{\sigma}))| \leq \lambda\sqrt{\alpha_N \log N}\}$ using the bound $|\tanh(x)| \leq |x|$ we have

$$\max_{1\leq i\leq N}\big|m_i(\boldsymbol{\sigma})\big| \leq \sqrt{\alpha_N \log N} + \beta \max_{1\leq i\leq N} R_i \max_{1\leq i\leq N}\big|m_i(\boldsymbol{\sigma})\big|,$$

which on using the fact that $\max_{1\leq i\leq N} R_i \to 1$ (see (1.11)) gives $\max_{1\leq i\leq N}|m_i(\boldsymbol{\sigma})| \lesssim \sqrt{\alpha_N \log N}$. Thus there exists a constant c' such that

$$\mathbb{P}\Big(\max_{1\leq i\leq N}\big|m_i(\boldsymbol{\sigma})\big| > c'\lambda\sqrt{\alpha_N \log N}\Big) \leq 2Ne^{-\frac{\lambda^2 \log N}{M}},$$

from which the desired conclusion follows for all λ large enough.

(b) More generally, we will show that for any vector $\mathbf{c} \in \mathbb{R}^N$ we have

$$\mathbb{E}\left(\sum_{i=1}^N c_i\sigma_i\right)^2 \lesssim (\log N)^{3/2}\sum_{i=1}^N c_i^2. \tag{5.1}$$

To this effect, for every nonnegative integer ℓ set $\mathbf{c}^{(\ell)} := \beta^\ell A_N^\ell \mathbf{c}$, and $x_\ell := \mathbb{E}[(\sum_i c_i^{(\ell)}\sigma_i)^2]$, and note that $\mathbf{c}^{(0)} = \mathbf{c}$, and the LHS of (5.1) is just x_0. Now for any $\ell \geq 0$ we can write

$$x_\ell = T_{1,\ell} + T_{2,\ell} + T_{3,\ell}, \tag{5.2}$$

where

$$T_{1,\ell} := \mathbb{E}\left[\left(\sum_{i=1}^N c_i^{(\ell)}\big(\sigma_i - \tanh\big(\beta m_i(\boldsymbol{\sigma})\big)\big)\right)^2\right], \qquad T_{2,\ell} := \mathbb{E}\left[\left(\sum_{i=1}^N c_i^{(\ell)} \tanh\big(\beta m_i(\boldsymbol{\sigma})\big)\right)^2\right],$$

$$T_{3,\ell} = 2\mathbb{E}\left[\left(\sum_{i\neq j} c_i^{(\ell)}c_j^{(\ell)}\big(\sigma_i - \tanh\big(\beta m_i(\boldsymbol{\sigma})\big)\big)\tanh\big(\beta m_j(\boldsymbol{\sigma})\big)\right)\right].$$

For controlling $T_{3,\ell}$, setting $m_i^j(\boldsymbol{\sigma}) := \sum_{k=1,k\neq j}^N A_N(i,k)\sigma_k\sigma_j$ we have

$$\begin{aligned}|T_{3,\ell}| &= 2\left|\sum_{i\neq j}^N c_i^{(\ell)}c_j^{(\ell)}\mathbb{E}\big[\big(\sigma_i - \tanh\big(\beta m_i(\boldsymbol{\sigma})\big)\big)\big(\tanh\big(\beta m_j(\boldsymbol{\sigma})\big) - \tanh\big(\beta m_j^i(\boldsymbol{\sigma})\big)\big)\big]\right| \\ &\lesssim \sum_{i\neq j}^N |c_i^{(\ell)}||c_j^{(\ell)}|A_N(i,j) \lesssim \|\mathbf{c}^{(\ell)}\|_2^2,\end{aligned} \tag{5.3}$$

where, in the first line, we use $\mathbb{E}[\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma}))|(\sigma_j, j\neq i)] = 0$ and consequently $\mathbb{E}[(\sigma_i - \tanh(\beta m_i(\boldsymbol{\sigma})))\tanh(\beta m_j^i(\boldsymbol{\sigma}))] = 0$ for $i \neq j$. The bound $|\tanh(\beta m_i(\boldsymbol{\sigma})) - \tanh(\beta m_i^j(\boldsymbol{\sigma}))| \lesssim A_N(i,j)$ is used in the second line.

Proceeding to bound $T_{2,\ell}$, use a Taylor series expansion to get $\tanh(\beta m_i(\boldsymbol{\sigma})) = \beta m_i(\boldsymbol{\sigma}) + \xi_i m_i(\boldsymbol{\sigma})^3$ for random variables $\{\xi_i\}_{1\leq i\leq N}$ uniformly bounded by 1 in absolute value. Also note that

$$x_{\ell+1} = \mathbb{E}\big[\big((\mathbf{c}^{(\ell+1)})^\top\boldsymbol{\sigma}\big)^2\big] = \mathbb{E}\big[\big(\beta(\mathbf{c}^{(\ell)})^\top A_N\boldsymbol{\sigma}\big)^2\big] = \mathbb{E}\left[\left(\beta\sum_{i=1}^N c_i^\ell m_i(\boldsymbol{\sigma})\right)^2\right].$$

Consequently,

$$\begin{aligned}T_{2,\ell} - x_{\ell+1} &= \mathbb{E}\left[\left(\sum_{i=1}^N c_i^{(\ell)}\{m_i(\boldsymbol{\sigma})\beta + \xi_i m_i(\boldsymbol{\sigma})^3\}\right)^2\right] - \mathbb{E}\left[\left(\beta\sum_{i=1}^N c_i^{(\ell)}m_i(\boldsymbol{\sigma})\right)^2\right] \\ &\leq 2\sqrt{x_{\ell+1}}\|\mathbf{c}^{(\ell)}\|_2\sqrt{\mathbb{E}\left[\sum_i m_i(\boldsymbol{\sigma})^6\right]} + \|\mathbf{c}^{(\ell)}\|_2^2\mathbb{E}\left[\sum_i m_i(\boldsymbol{\sigma})^6\right].\end{aligned} \tag{5.4}$$

Finally, using the Cauchy–Schwarz inequality gives

$$\mathbb{E}\left(\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6\right) \leq \sqrt{\mathbb{E}\left(\sum_{i=1}^N m_i^2\right)^2}\sqrt{\mathbb{E}\max_{1\leq i\leq N}|m_i(\boldsymbol{\sigma})|^8} \leq C^2\|A_N\|_F^2\alpha_N^2(\log N)^2 \tag{5.5}$$

for some C free of N, where the last inequality uses part (a) of this lemma and part (b) of Lemma 2.1. Noting that $T_{1,\ell} \lesssim \|\mathbf{c}^{(\ell)}\|_2^2$ by part (b) of Lemma 3.1, combining (5.3), (5.4), and (5.5) along with (5.2) gives the existence of a constant D free of N, ℓ such that

$$x_\ell \leq x_{\ell+1} + 2\sqrt{x_{\ell+1}}\|\mathbf{c}\|_2\beta_N^\ell\delta_N + \|\mathbf{c}\|_2^2\beta_N^{2\ell}\delta_N^2 + D\beta_N^{2\ell}\|\mathbf{c}\|_2^2, \tag{5.6}$$

where we have also used the bound $\|\mathbf{c}^{(\ell)}\|_2 \leq \beta_N^\ell \|\mathbf{c}\|_2$ with $\beta_N := \beta\|A_N\|_2$, and we set $\delta_N := \max(1, C\|A_N\|\alpha_N \log N)$. Since $\beta_N \to \beta < 1$, for all N large we have $\beta_N \leq \beta_0$ for some $\beta_0 < 1$. Given constants $\beta_0 \in (0,1)$, $D > 0$, there exists M large enough such that $M > (\beta_0\sqrt{M}+1)^2 + D$. With this M, β_0 we claim that for all ℓ, we have

$$x_\ell \leq M\|\mathbf{c}\|_2^2 \beta_0^{2\ell}\delta_N^2, \tag{5.7}$$

from which (5.1) is immediate on setting $\ell = 0$. For proving (5.7) we use backwards induction on ℓ. Using Cauchy–Schwarz inequality gives

$$x_\ell \leq N\|\mathbf{c}^{(\ell)}\|_2^2 \leq N\beta_N^\ell\|\mathbf{c}\|_2^2,$$

and so (5.7) holds for all ℓ large enough, as $\beta_N < \beta_0$. Assume that the result holds for $x_{\ell+1}$ for some ℓ, that is, $x_{\ell+1} \leq M\|\mathbf{c}\|_2^2\beta_0^{2\ell+2}\delta_N^2$. Using (5.6) gives

$$x_\ell \leq \|\mathbf{c}\|_2^2\beta_0^{2\ell}\delta_N^2\big(M\beta_0^2 + 2\sqrt{M}\beta_0 + 1 + D\big) \leq M\|\mathbf{c}\|_2^2\beta_0^{2\ell}\delta_N^2,$$

where the last step uses the choice of M. This verifies the claim for ℓ, and hence proves (5.7) by backward induction, for all $\ell \geq 0$. □

PROOF OF LEMMA 2.3. (a) As in the proof of part (a) of Lemma 2.2, it suffices to prove the result for λ large. To this effect, define an $N \times N$ matrix $\widetilde{A}_N$ by setting $\widetilde{A}_N(i,j) := A_N(i,j)/R_{\max}$ for $i \neq j$ and $\widetilde{A}_N(i,i) := 1 - R_i/R_{\max}$ where $R_{\max} = \max_{1\leq i\leq N} R_i$. Observe that $\mathbf{1}^\top\widetilde{A}_N = \mathbf{1}^\top$, and so

$$\begin{aligned}
&\left|\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big) - \sum_{j=1}^N \widetilde{A}_N(i,j)\big(m_j(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\quad= \left|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N \widetilde{A}_N(i,j)m_j(\boldsymbol{\sigma})\right| \\
&\quad\leq \left|m_i - \sum_{j=1}^N A_N(i,j)m_j(\boldsymbol{\sigma})\right| + \sum_{j=1}^N\big|A_N(i,j) - \widetilde{A}_N(i,j)\big| \\
&\quad\lesssim \left|m_i - \sum_{j=1}^N A_N(i,j)\tanh\big(m_j(\boldsymbol{\sigma})\big)\right| + \max_{1\leq i\leq N}\big|m_i(\boldsymbol{\sigma})\big|^3 + \max_{1\leq i\leq N}|R_i - 1|.
\end{aligned}\tag{5.8}$$

Using part (b) of Lemma 3.1, a union bound as in the proof of part (a) of Lemma 2.2 shows that for all $\lambda > 0$ we have $\mathbb{P}(E_N^c) \leq 2e^{-c\lambda^2}$ for some constant $c > 0$ free of N, where

$$E_N := \left\{\max_{1\leq i\leq N}\left|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(i,j)\tanh\big(m_j(\boldsymbol{\sigma})\big)\right| \leq \lambda\sqrt{\alpha_N\log N}\right\} \tag{5.9}$$

for some constant c free of N, with $\alpha_N = \max_{1\leq i\leq N}\sum_{j=1}^N A_N(i,j)^2$ as in Theorem 1.4. Proceeding to bound the second term in the RHS of (5.8), note that, with $K := \operatorname{argmax}_{1\leq i\leq N}|m_i(\boldsymbol{\sigma})|$ and assuming $m_K(\boldsymbol{\sigma}) \geq 0$ without loss of generality, we have

$$m_K(\boldsymbol{\sigma})^3 \lesssim m_K(\boldsymbol{\sigma}) - \tanh\big(m_K(\boldsymbol{\sigma})\big) \leq m_K(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(K,j)\tanh\big(m_j(\boldsymbol{\sigma})\big) + \max_{1\leq i\leq N}|R_i - 1|.$$

By a symmetric argument, we get

$$\max_{1\leq i\leq N}\big|m_i(\boldsymbol{\sigma})\big|^3 \lesssim \max_{1\leq i\leq N}\left|m_i(\boldsymbol{\sigma}) - \sum_{j=1}^N A_N(i,j)\tanh\big(m_j(\boldsymbol{\sigma})\big)\right| + \max_{1\leq i\leq N}|R_i - 1|. \tag{5.10}$$

Thus, combining (5.8) and (5.10), on the set E_N we have

$$\begin{aligned}
&\max_{1\le i\le N}\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\qquad\le C\Big[\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big]
\end{aligned}\tag{5.11}$$

for some $C<\infty$ free of N. Now, for any integer $\ell\ge 2$ we have

$$\begin{aligned}
&\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N^{\ell}(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\quad\le\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N^{\ell-1}(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\qquad+\left|\sum_{j=1}^{N}\widetilde{A}_N^{\ell-1}(i,j)\left\{\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{k=1}^{N}\widetilde{A}_N(j,k)\big(m_k(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right\}\right| \\
&\quad\le\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N^{\ell-1}(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\qquad+\max_{1\le j\le N}\left|\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{k=1}^{N}\widetilde{A}_N(j,k)\big(m_k(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right|,
\end{aligned}$$

which, via a recursive argument gives

$$\begin{aligned}
&\max_{1\le i\le N}\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N^{\ell}(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\quad\le\ell\max_{1\le i\le N}\left|\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)-\sum_{j=1}^{N}\widetilde{A}_N(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right| \\
&\quad\le C\ell\Big(\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big),
\end{aligned}\tag{5.12}$$

where the last line uses (5.11) on the set E_N. Using part (a) of Lemma 6.2, we note the existence of D free of N such that for the choice $\ell=D\log N$ we have $\max_{1\le i\le N}A^{\ell}(i,i)\le\frac{3}{N}$. With this choice of ℓ, we have

$$\begin{aligned}
&\mathbb{P}\Big(\max_{1\le i\le N}\big|m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big|\ge 2C\ell\Big[\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big],E_N\Big) \\
&\quad\le\mathbb{P}\left(\max_{1\le i\le N}\left|\sum_{j=1}^{N}\widetilde{A}_N^{\ell}(i,j)\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)\right|\ge C\ell\Big[\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big]\right) \\
&\quad\le\mathbb{P}\left(\sum_{j=1}^{N}\big(m_j(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\ge\frac{C^2\ell^2N}{2}\Big[\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big]^2\right),
\end{aligned}$$

where the last line uses Cauchy–Schwarz inequality. Fixing δ small enough and using part (c) of Lemma 2.1, this gives

$$\log \mathbb{P}\Big(\max_{1\le i\le N}|m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})|\ge 2C\ell\Big[\lambda\sqrt{\alpha_N\log N}+\max_{1\le i\le N}|R_i-1|\Big],E_N\Big)$$
$$\lesssim -N\alpha_N(\log N)^3\lambda^2-N(\log N)^2\max_{1\le i\le N}|R_i-1|^2+\log\mathbb{E}e^{\delta\sum_{i=1}^N(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2}$$
$$\lesssim -N\alpha_N(\log N)^3\lambda^2-N(\log N)^2\max_{1\le i\le N}|R_i-1|^2+\|A_N\|_F^2+\frac{1}{N}\Bigg[\sum_{i=1}^N(R_i-1)^2\Bigg]^2$$
$$+\frac{1}{N}\Bigg[\sum_{i=1}^N(R_i-1)\Bigg]^2+\log N,$$

from which the desired conclusion follows for λ large enough on noting the inequality $N\alpha_N\ge\|A_N\|_F^2\gtrsim 1$. □

In order to prove Lemma 2.3, parts (b) and (c), we need the following lemma whose proof we defer to the end of this section.

LEMMA 5.1. *Assume that* (1.4), (1.5), (1.9) *holds, and the RHS of* (1.10) *is bounded. Then, setting* $\nu_N:=\mathbb{E}_1^N[(N^{1/4}\overline{\boldsymbol{\sigma}})^6]$ *the following conclusions hold*:

$$\nu_N\lesssim \nu_N^{2/3}+\nu_N^{1/3}+\nu_N^{1/2}+\nu_N^{1/2}\sqrt{\frac{\mathbb{E}[\sum_{i=1}^N(R_i-1)\sigma_i]^2}{N^{1/2}}},\tag{5.13}$$

$$\mathbb{E}\Bigg[\sum_{i=1}^N(R_i-1)\sigma_i\Bigg]^2\lesssim(\log N)^4\Bigg(\sum_{i=1}^N(R_i-1)^2+N^{-1/2}\Bigg[\sum_{i=1}^N(R_i-1)\Bigg]^2\Bigg)\times\big(1+\mathbb{E}[(N^{1/4}\overline{\boldsymbol{\sigma}})^2]\big).\tag{5.14}$$

PROOF OF LEMMA 2.3, PARTS (b) AND (c). Use (5.14) and the fact that the RHS of (1.10) is bounded to get

$$\mathbb{E}\Bigg[\sum_{i=1}^N(R_i-1)\sigma_i\Bigg]^2\lesssim\sqrt{N}\big(1+\mathbb{E}[(N^{1/4}\overline{\boldsymbol{\sigma}})^2]\big)\lesssim\sqrt{N}\big(1+\nu_N^{1/3}\big).$$

Along with (5.13), this gives $\nu_N\lesssim\nu_N^{2/3}+\nu_N^{1/3}+\nu_N^{1/2}(1+\nu_N^{1/3})+1$, and so ν_N must be bounded, thereby proving part (b). Now, part (c) is an immediate consequence of part (b) and (5.14). □

PROOF OF LEMMA 5.1. (a) Proof of (5.13).

To begin, borrowing notation from the proof of Theorem 1.3 and using (3.16) gives the existence of $C<\infty$ such that

$$\big|\mathbb{E}[T_N-T_N'|\boldsymbol{\sigma}]-N^{-3/2}T_N^3/3\big|$$
$$\le\frac{2}{15}N^{-2}|T_N|^5+C\Bigg\{N^{-3/4}|\overline{\boldsymbol{\sigma}}-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})|+N^{-2}|T_N|\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2$$
$$+N^{-7/4}\Bigg|\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^3\Bigg|\Bigg\}.$$

On multiplying both sides of the above inequality by $N^{3/2}|T_N|^3$ and taking expectation gives

$$\begin{aligned}\mathbb{E}[T_N^6] \le{}& (2/5)N^{-1/2}\mathbb{E}|T_N|^8 + 3C\Bigg\{N^{3/4}\mathbb{E}\big[|T_N|^3|\bar{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|\big] \\ &+ N^{-1/2}\mathbb{E}\Bigg[|T_N|^4\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\Bigg] \\ &+ N^{-1/4}\mathbb{E}\Bigg[|T_N|^3\Bigg|\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^3\Bigg|\Bigg]\Bigg\} + 3N^{3/2}|\mathbb{E}(T_N - T_N')T_N^3|.\end{aligned} \tag{5.15}$$

We will now bound each of the terms in the RHS of (5.15). To begin, note that $|T_N - T_N'| \le 2N^{-3/4}$ and $\mathbb{E}[T_N] = \mathbb{E}[T_N']$. This, along with the fact that (T_N, T_N') is an exchangeable pair gives

$$\begin{aligned}\mathbb{E}(T_N - T_N')T_N^3 &= (1/2)\mathbb{E}(T_N - T_N')T_N^3 - (1/2)\mathbb{E}(T_N - T_N')(T_N')^3 \\ &= (1/2)\mathbb{E}\big[(T_N - T_N')^2\big(T_N^2 + T_NT_N' + (T_N')^2\big)\big] \\ &\le 6N^{-3/2}\mathbb{E}[T_N^2] \le 6N^{-3/2}\nu_N^{1/6},\end{aligned} \tag{5.16}$$

where $\nu_N = \mathbb{E}[(N^{1/4}\bar{\boldsymbol{\sigma}})^6]$ as in the statement of the lemma. Also with ε_N, r_N as in the statement of Theorem 1.3, use part (c) of Lemma 2.1, and part (a) of Lemma 2.3 to get that for any positive integer p, we have

$$\mathbb{E}\Bigg[\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\Bigg]^p \lesssim \varepsilon_N^p, \qquad \mathbb{E}\max_{1\le i\le N}|m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|^p \lesssim r_N^p. \tag{5.17}$$

Finally, since the RHS of (1.10) is bounded, we have

$$\varepsilon_N \lesssim \sqrt{N}, \qquad \varepsilon_N r_N \lesssim N^{1/4}, \qquad \|\mathbf{c}\|_2^2 + N^{-1/2}\Bigg[\sum_{i=1}^N c_i\Bigg]^2 \lesssim \sqrt{N}. \tag{5.18}$$

Armed with these estimates and proceeding to bound the second, third, and fourth terms in (5.15), use Hölder's inequality to get

$$N^{3/4}\mathbb{E}\big[|T_N|^3|\bar{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})|\big] \le N^{-1/4}\sqrt{\nu_N}\sqrt{\mathbb{E}\Bigg[\sum_{i=1}^N(R_i - 1)\sigma_i\Bigg]^2}, \tag{5.19}$$

$$\begin{aligned}\mathbb{E}\Bigg[T_N^4\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\Bigg] &\le \nu_N^{2/3}\Bigg(\mathbb{E}\Bigg[\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}\big)^2\Bigg]^3\Bigg)^{1/3} \\ &\lesssim \nu_N^{2/3}\varepsilon_N \lesssim \nu_N^{2/3}\sqrt{N},\end{aligned} \tag{5.20}$$

$$\begin{aligned}&\mathbb{E}\Bigg[|T_N|^3\Bigg|\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^3\Bigg|\Bigg] \\ &\quad\le \sqrt{\nu_N}\Bigg(\mathbb{E}\Bigg[\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\Bigg]^4\Bigg)^{1/4}\Big(\mathbb{E}\Big[\max_{1\le i\le N}\big(m_i(\boldsymbol{\sigma}) - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^4\Big]\Big)^{1/4} \\ &\quad\lesssim \sqrt{\nu_N}\varepsilon_N r_N \lesssim \sqrt{\nu_N}N^{1/4},\end{aligned} \tag{5.21}$$

where the last bounds in (5.20) and (5.21) use (5.17) and (5.18). Finally, for the fifth term in the RHS of (5.15), note that $|T_N| \le N^{1/4}$, and so the first term in the RHS of (5.15) is bounded by $(2/5)\mathbb{E}[T_N^6]$. Combining this along with (5.15), (5.16), (5.19), (5.20), and (5.21) gives

$$\nu_N \lesssim \nu_N^{1/3} + \nu_N^{1/2} + \nu_N^{2/3} + \nu_N^{1/2}\sqrt{\frac{\mathbb{E}[\sum_{i=1}^N (R_i - 1)\sigma_i]^2}{N^{1/2}}}$$

which completes the proof of (5.13)

(b) Proof of (5.14).

To begin, for any vector $\mathbf{h} := (h_1, \dots, h_N)$ write

$$\sum_{i=1}^N h_i\sigma_i = \sum_{i=1}^N h_i\big(\sigma_i - \tanh(m_i(\boldsymbol{\sigma}))\big) + \sum_{i=1}^N h_i\big(\tanh(m_i(\boldsymbol{\sigma})) - \tanh(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))\big) + \tanh(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))\sum_{i=1}^N h_i,$$

which using part (b) of Lemma 3.1 gives

$$\mathbb{E}\left[\sum_{i=1}^N h_i\sigma_i\right]^2 \lesssim \|\mathbf{h}\|_2^2 + \|\mathbf{h}\|_2^2\varepsilon_N + \left[\sum_{i=1}^N h_i\right]^2 \mathbb{E}\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2 \lesssim \|\mathbf{h}\|_2^2\varepsilon_N + \left[\sum_{i=1}^N h_i\right]^2 \mathbb{E}\big(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2\big), \tag{5.22}$$

where the second line uses part (c) of Lemma 2.1, and ε_N equals the RHS of (2.2). Setting $\mathbf{c} = \mathbf{R} - \mathbf{1}$ and using (5.22) with $\mathbf{h} = \mathbf{c}$ gives

$$\mathbb{E}\left[\sum_{i=1}^N c_i\sigma_i\right]^2 \lesssim \|\mathbf{c}\|_2^2 + \|\mathbf{c}\|_2^2\varepsilon_N + \left[\sum_{i=1}^N c_i\right]^2 \mathbb{E}\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2 \lesssim N, \tag{5.23}$$

where the last line uses (5.18). Along with (5.13) this gives $\nu_N \lesssim \nu_N^{1/3} + \nu_N^{1/2} + \nu_N^{2/3} + \nu_N^{1/2}\sqrt{N}$, and so

$$\nu_N \lesssim \sqrt{N} \quad \Rightarrow \quad \mathbb{E}\overline{\boldsymbol{\sigma}}^6 \lesssim N^{-1}. \tag{5.24}$$

Also, an argument similar to the derivation of (5.23) shows that for any positive integer p, we have

$$\mathbb{E}\big(\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^{2p} = N^{-2p}\mathbb{E}\left[\sum_{i=1}^N c_i\sigma_i\right]^{2p} \tag{5.25}$$

$$\lesssim N^{-2p}\left(\|\mathbf{c}\|_2^{2p} + \|\mathbf{c}\|_2^{2p}\varepsilon_N^p + \left(\sum_{i=1}^N c_i\right)^{2p} \mathbb{E}\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2\right) \lesssim N^{-p}, \tag{5.26}$$

where the last bound uses (5.18). Combining we have the following conclusions:

$$\mathbb{E}\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^6 \lesssim \mathbb{E}(\overline{\boldsymbol{\sigma}})^6 + \mathbb{E}\big(\overline{\boldsymbol{\sigma}} - \overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^6 \lesssim \frac{1}{N}, \tag{5.27}$$

$$\mathbb{E}\left(\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6\right) \lesssim N\mathbb{E}\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^6 + \sqrt{\mathbb{E}\max_{1\le i\le N}\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^8}\sqrt{\mathbb{E}\left[\sum_{i=1}^N\big(m_i(\boldsymbol{\sigma})-\overline{\boldsymbol{m}}(\boldsymbol{\sigma})\big)^2\right]^2} \lesssim 1 + r_N^4\varepsilon_N \lesssim 1, \tag{5.28}$$

where (5.27) uses (5.24) and (5.25) with $p=3$, and (5.28) uses (5.27) along with (5.17) and (5.18). Armed with these estimates, we now focus on deriving (5.14).

Let $\widetilde{A}_N$ be as defined in the proof of part (a) of Lemma 2.3, and set $\mathbf{c}^{(\ell)} := \mathbf{c}^\top \widetilde{A}_N^\ell$ and $x_\ell := \mathbb{E}[\sum_{i=1}^N c_i^{(\ell)}\sigma_i]^2$ for $\ell \ge 0$. As in the proof of part (b) of Lemma 2.2, we can write $x_\ell = T_{1,\ell}+T_{2,\ell}+T_{3,\ell}$, where

$$T_{1,\ell} := \mathbb{E}\left[\sum_{i=1}^N c_i^{(\ell)}\big(\sigma_i - \tanh m_i(\boldsymbol{\sigma})\big)\right]^2, \qquad T_{2,\ell} := \mathbb{E}\left[\sum_{i=1}^N c_i^{(\ell)}\tanh m_i(\boldsymbol{\sigma})\right]^2,$$

$$T_{3,\ell} := 2\mathbb{E}\left[\sum_{i\ne j} c_i^{(\ell)}c_j^{(\ell)}\big(\sigma_i - \tanh m_i(\boldsymbol{\sigma})\big)\tanh m_j(\boldsymbol{\sigma})\right].$$

By the argument presented in the proof of part (b) of Lemma 2.2 we have $T_{1,\ell} \lesssim \|\mathbf{c}^{(\ell)}\|_2^2 \le \|\mathbf{c}\|_2^2$, and $T_{3,\ell} \lesssim \|\mathbf{c}\|_2^2$. Next, using the Taylor series expansion, we can write $\tanh(m_i(\boldsymbol{\sigma})) = m_i(\boldsymbol{\sigma}) + \xi_i m_i(\sigma)^3$ for random variables $\{\xi_i\}_{1\le i\le N}$ which are uniformly bounded by 1 in absolute value. Consequently,

$$\begin{aligned}
T_{2,\ell} - x_{\ell+1} &= \mathbb{E}\left[\boldsymbol{\sigma}^\top A_N\mathbf{c}^{(\ell)} + \sum_{i=1}^N c_i^{(\ell)}\xi_i m_i(\boldsymbol{\sigma})^3\right]^2 - \mathbb{E}[\boldsymbol{\sigma}^\top\widetilde{A}_N\mathbf{c}^{(\ell)}]^2 \\
&\le 2\sqrt{x_{\ell+1}}\sqrt{\mathbb{E}\left[\sum_{i=1}^N |c_i^{(\ell)}m_i(\boldsymbol{\sigma})^3|\right]^2} \\
&\quad + 2\sqrt{x_{\ell+1}}\sqrt{\mathbb{E}[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2} + \mathbb{E}[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2 \\
&\quad + \mathbb{E}\left[\sum_{i=1}^N |c_i^{(\ell)}m_i(\boldsymbol{\sigma})^3|\right]^2 + 2\sqrt{\mathbb{E}[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2}\sqrt{\mathbb{E}\left[\sum_{i=1}^N |c_i^{(\ell)}m_i(\boldsymbol{\sigma})^3|\right]^2} \qquad (5.29)\\
&\le 2\sqrt{x_{\ell+1}}\|\mathbf{c}\|_2^2\sqrt{\mathbb{E}\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6} \\
&\quad + 2\sqrt{x_{\ell+1}}\sqrt{E[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2} + \mathbb{E}[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2 \\
&\quad + \|\mathbf{c}\|_2^2\mathbb{E}\left[\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6\right] + 2\|\mathbf{c}\|_2^2\sqrt{\mathbb{E}[\mathbf{c}^{(\ell)}H_N\boldsymbol{\sigma}]^2}\sqrt{\mathbb{E}\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6}.
\end{aligned}$$

Proceeding to bound the RHS of (5.29), use (1.9) and (5.25) respectively to note that $\|H_N\|_{\mathrm{op}} \lesssim N^{-1/4}$, and $N\mathbb{E}(\overline{\boldsymbol{m}}(\boldsymbol{\sigma}))^2 \lesssim N\mathbb{E}(\overline{\boldsymbol{\sigma}})^2 + 1$, and an application of (5.22) with $\mathbf{h} = H_N\mathbf{c}^{(\ell)}$) gives

$$\mathbb{E}[(\mathbf{c}^{(\ell)})^\top H_N\boldsymbol{\sigma}]^2 \lesssim \|\mathbf{c}^{(\ell)}\|_2^2\|H_N\|_{\mathrm{op}}^2\big(\varepsilon_N + N\mathbb{E}(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2)\big) \lesssim \|\mathbf{c}\|_2^2\mu_N, \tag{5.30}$$

with $\mu_N := 1 + \mathbb{E}(N^{1/4}\overline{\boldsymbol{\sigma}})^2$, where the second inequality uses (5.18) and (5.25). We now claim that there exists a constant $D > 0$ such that

$$x_{D(\log N)^2} \lesssim \mu_N \left(\|\mathbf{c}\|_2^2 + N^{-1/2} \left[\sum_{i=1}^N c_i \right]^2 \right)^2. \tag{5.31}$$

Given this claim, we have the existence of a constant C free of N such that

$$x_{D(\log N)^2} \le C^2 \mu_N \left(\|\mathbf{c}\|_2^2 + N^{-1/2} \left[\sum_{i=1}^N c_i \right]^2 \right)^2. \tag{5.32}$$

Also, using (5.30) and (5.28), and making C bigger if needed, for all $\ell \ge 0$ we have

$$x_\ell \le x_{\ell+1} + 2C\sqrt{x_{\ell+1}}\|\mathbf{c}\|_2\sqrt{\mu_N} + C^2\|\mathbf{c}\|_2^2\mu_N. \tag{5.33}$$

With $L = D(\log N)^2$, we will now show that the bound

$$x_\ell \le (L - \ell + 1)^2 C^2 \left[\|\mathbf{c}\|_2^2 + N^{-1/2} \left(\sum_{i=1}^N c_i \right)^2 \right] \tag{5.34}$$

holds for all $\ell \in [0, L]$ by backwards induction. By (5.32) we have that (5.34) holds for $\ell = L$. Suppose (5.34) holds for $\ell + 1$ for some $\ell \in [0, L - 1]$. Using (5.33) gives

$$x_\ell \le C^2 \mu_N \|\mathbf{c}\|_2^2 [(L - \ell)^2 + 2(L - \ell) + 1] = (L - \ell + 1)^2 C^2 \mu_N \|\mathbf{c}\|_2^2,$$

verifying (5.34) for ℓ, and thus verifying (5.34) for all $\ell \in [0, L]$ by induction. Setting $\ell = 0$ in (5.34) we get the bound

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^N c_i \sigma_i \right)^2 &\le L^2 C^2 \mu_N \left[\sum_{i=1}^N c_i^2 + N^{-1/2} \left(\sum_{i=1}^N c_i \right)^2 \right] \\
&\le C^2 D^2 \mu_N (\log N)^4 \left[\sum_{i=1}^N c_i^2 + N^{-1/2} \left(\sum_{i=1}^N c_i \right)^2 \right],
\end{aligned}$$

which verifies (5.14), as desired.

It thus remains to verify (5.31), for which using spectral decomposition write $\widetilde{A}_N = \sum_{i=1}^N \widetilde{\lambda}_i \widetilde{\mathbf{q}}_i \widetilde{\mathbf{q}}_i^\top$, where we set $\widetilde{\lambda}_i := \lambda_i(\widetilde{A}_N)$ for convenience of notation. With $L = D(\log N)^2$, this gives

$$\begin{aligned}
\mathbf{c}^\top \widetilde{A}_N^L \boldsymbol{\sigma} &= \overline{\boldsymbol{\sigma}} \sum_{i=1}^N c_i + \widetilde{\lambda}_N^L \mathbf{c}^\top \widetilde{\mathbf{q}}_N \widetilde{\mathbf{q}}_N^\top \boldsymbol{\sigma} + \sum_{i=2}^{N-1} \widetilde{\lambda}_i^L \mathbf{c}^\top \widetilde{\mathbf{q}}_i \widetilde{\mathbf{q}}_i^\top \boldsymbol{\sigma} \\
&= \overline{\boldsymbol{\sigma}} \sum_{i=1}^N c_i + \widetilde{\lambda}_N^L \mathbf{c}^\top \widetilde{\mathbf{q}}_N \widetilde{\mathbf{q}}_N^\top \boldsymbol{\sigma} + O(N^{-cD+2}),
\end{aligned}$$

where the last equality uses Lemma 6.2 to get

$$\max_{2 \le i \le N-1} |\widetilde{\lambda}_i|^L \le \left(1 - \frac{c}{\log N} \right)^\ell \le N^{-cD}$$

for some $c > 0$. Consequently for D large enough we have

$$\mathbb{E}[\mathbf{c}^\top \widetilde{A}_N^L \boldsymbol{\sigma}]^2 \lesssim \left[\sum_{i=1}^N c_i \right]^2 \mathbb{E}[\overline{\boldsymbol{\sigma}}^2] + \|\mathbf{c}\|_2^2 \mathbb{E}[(\widetilde{\mathbf{q}}_N^\top \boldsymbol{\sigma})^2]. \tag{5.35}$$

Since $\widetilde{\boldsymbol{q}}_N^\top \widetilde{A}_N = \lambda_N \widetilde{\boldsymbol{q}}_N^\top$ where $\widetilde{\lambda}_N$ is bounded away from 1 by (1.7), we have

$$
\begin{aligned}
&(1-\widetilde{\lambda}_N)\sum_{i=1}^N \widetilde{q}_N(i)\sigma_i \\
&\quad = \sum_{i=1}^N \widetilde{q}_N(i)\big(\sigma_i - m_i(\boldsymbol{\sigma})\big) + \widetilde{\mathbf{q}}_N^\top H_N \sigma \\
&\quad = \sum_{i=1}^N \widetilde{q}_N(i)\big(\sigma_i - \tanh\big(m_i(\boldsymbol{\sigma})\big)\big) + \sum_{i=1}^N \widetilde{q}_N(i)\big(\tanh\big(m_i(\boldsymbol{\sigma})\big) - m_i(\boldsymbol{\sigma})\big) + \widetilde{\mathbf{q}}_N^\top H_N \sigma.
\end{aligned}
$$

This immediately gives

$$
\begin{aligned}
&(1-\widetilde{\lambda}_N)^2 \mathbb{E}\Bigg[\sum_{i=1}^N \widetilde{q}_N(i)\sigma_i\Bigg]^2 \\
&\quad \lesssim \mathbb{E}\Bigg[\sum_{i=1}^N \widetilde{q}_N(i)\big(\sigma_i - \tanh\big(m_i(\boldsymbol{\sigma})\big)\big)\Bigg]^2 + \mathbb{E}\Bigg[\sum_{i=1}^N |\widetilde{q}_N(i)|\big|m_i(\boldsymbol{\sigma})\big|^3\Bigg]^2 + \mathbb{E}\big[\widetilde{\mathbf{q}}_N^\top H_N \boldsymbol{\sigma}\big]^2 \\
&\quad \le \sum_{i=1}^N \widetilde{q}_N(i)^2 + \sqrt{\sum_{i=1}^N \widetilde{q}_N(i)^2}\sqrt{\mathbb{E}\Bigg[\sum_{i=1}^N m_i(\boldsymbol{\sigma})^6\Bigg]} + \mathbb{E}\big[\widetilde{\mathbf{q}}_N^\top H_N \boldsymbol{\sigma}\big]^2 \\
&\quad \lesssim 1 + \|H_N\|_{\mathrm{op}}^2\big[\varepsilon_N + N\mathbb{E}\big(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2\big)\big],
\end{aligned}
$$

where the last bound uses (5.22) with $\mathbf{h} = \widetilde{\mathbf{q}}_N$. Since $N\mathbb{E}(\overline{\boldsymbol{m}}(\boldsymbol{\sigma})^2) \lesssim N\mathbb{E}(\overline{\boldsymbol{\sigma}}^2) + 1 \lesssim \sqrt{N}\mu_N$, using the last bound along with (5.35) gives

$$
\mathbb{E}\big(\mathbf{c}^\top \widetilde{A}_N^L \boldsymbol{\sigma}\big)^2 \lesssim \mu_N\Bigg(N^{-1/2}\Bigg[\sum_{i=1}^N c_i\Bigg]^2 + \sum_{i=1}^N c_i^2\Bigg),
$$

thus verifying (5.31), and hence completing the proof of the lemma. □

REMARK 5.1. As in the proofs of part (b) of Lemmas 2.2 and 2.3, the above argument can be modified to bound the moments of general linear combinations $\sum_{i=1}^N c_i\sigma_i$ for any $\mathbf{c} \in \mathbb{R}^N$.

6. Supplementary lemmas and proofs.

6.1. *Proofs of Lemma 4.1 and Lemma 4.2.*

PROOF OF LEMMA 4.1. Noting the presence of $\mathrm{Tr}^+(D_N)$ in the RHS of the bound, it suffices to prove the result for D_N with all diagonal entries set to 0. Let $(Z_1, Z_2, \dots, Z_N)$ be i.i.d. $\mathcal{N}(0, 1)$ random variables. We claim that

$$
\begin{aligned}
&\mathbb{E}\Bigg[\exp\Bigg(\frac{1}{2}\sum_{i,j=1}^N D_N(i,j)\widetilde{X}_i\widetilde{X}_j + \sum_{i=1}^N c_i\widetilde{X}_i\Bigg)\Bigg] \\
&\quad \le \mathbb{E}\Bigg[\exp\Bigg(\frac{s_\mu}{2}\sum_{i,j=1}^N D_N(i,j)Z_iZ_j + \sqrt{s_\mu}\sum_{i=1}^N c_iZ_i\Bigg)\Bigg].
\end{aligned}
\tag{6.1}
$$

Indeed, to see this, recall that the sub-Gaussian norm of $\widetilde{X}_i$ is given by s_μ for $1 \le i \le n$ (see, e.g., [28], Lemma 1 and [37], Theorem 2.1). Consequently, for every $\theta \in \mathbb{R}$ we have

$\mathbb{E}[\exp(\theta \widetilde{X}_i)] \le \mathbb{E}[\exp(\theta \sqrt{s_\mu} Z_i)]$. Using this, (6.1) can be obtained by inductively replacing each $\widetilde{X}_i$ on the left-hand side of (6.1) with $\sqrt{s_\mu} Z_i$.The RHS of (6.1) can be computed directly to get

$$\log\left\{\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i,j=1}^N s_\mu D_N(i,j) Z_i Z_j + \sqrt{s_\mu}\sum_{i=1}^N c_i Z_i\right)\right]\right\}$$
$$= -(1/2)\log\det(I_N - s_\mu D_N) + (1/2) s_\mu \sum_{i=1}^N c_i^2.$$

Finally, by noting the existence of $\rho \in (s_\mu \limsup_{N\to\infty} \lambda_1(D_N), 1)$, and using the bound $-\log(1-x) \lesssim x$ for $x \in [0, 1-\rho]$, the desired conclusion follows. □

PROOF OF LEMMA 4.2. By Hölder's inequality, for any $p > 0$ the left-hand side of (4.1) can be bounded by

$$\left(\mathbb{E}^{\mathrm{CW}}\left[\exp\left(\frac{\beta(1+p)}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\right)\right]\right)^{1/(1+p)} \mathbb{P}\big(\big|\widetilde{W}_N - M(\boldsymbol{\sigma})\big| \ge \varepsilon\big)^{\frac{p}{1+p}}.$$

Since $\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}(\widetilde{W}_N - M(\boldsymbol{\sigma})| > \varepsilon) < 0$ by part (b) of Proposition 6.1, it suffices to show the existence of $p > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}^{\mathrm{CW}}\left[\exp\left(\frac{\beta(1+p)}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\right)\right] \le 0. \tag{6.2}$$

To this effect, setting

$$g_p(\boldsymbol{\sigma}) := \frac{\beta}{2}\boldsymbol{\sigma}^\top A_N \boldsymbol{\sigma} + \frac{\beta p}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}$$

note that

$$\begin{aligned} &\log \mathbb{E}^{\mathrm{CW}}\left[\exp\left(\frac{\beta(1+p)}{2}\boldsymbol{\sigma}^\top \mathcal{A}_N \boldsymbol{\sigma}\right)\right] \\ &\quad = \sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left\{g_p(\boldsymbol{\sigma}) - \sum_{i=1}^N I(\sigma_i)\right\} - \log Z_N^{\mathrm{CW}}(\beta, B) + o(N), \end{aligned} \tag{6.3}$$

where the last line uses [4], Theorem 1.1, along with the observation $\mathrm{Tr}((A_N + \mathcal{A}_N)^2) = o(N)$. Using spectral theorem we have $A_N = \sum_{i=1}^N \lambda_i \mathbf{q}_i \mathbf{q}_i^\top$ with $\lambda_i = \lambda_i(A_N)$, and so

$$\begin{aligned} &\sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left(g_p(\boldsymbol{\sigma}) - \frac{\beta}{2}\sum_{i=1}^N \sigma_i^2\right) \\ &\quad = \sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left[\frac{\beta}{2}\sum_{i=1}^N (\lambda_i - 1)\boldsymbol{\sigma}^\top \mathbf{q}_i \mathbf{q}_i^\top \boldsymbol{\sigma} + \frac{\beta p}{2}\boldsymbol{\sigma}^\top\left(\lambda_1 \mathbf{q}_1 \mathbf{q}_1^\top - \frac{\mathbf{1}\mathbf{1}^\top}{N}\right)\boldsymbol{\sigma} \right. \\ &\qquad \left. + \frac{\beta p}{2}\sum_{i=2}^N \lambda_i \boldsymbol{\sigma}^\top \boldsymbol{q}_i \boldsymbol{q}_i^\top \boldsymbol{\sigma}\right] \\ &\quad \lesssim o(N) + \sum_{i=2}^N (\boldsymbol{\sigma}^\top \boldsymbol{q}_i \boldsymbol{q}_i^\top \boldsymbol{\sigma})\left(-\frac{\beta}{2}(1-\lambda_i) + \frac{\beta p}{2}\lambda_i\right), \end{aligned}$$

where the bound in the last line uses (1.5), and Lemma 6.1. Finally note that (1.7) shows the existence of $\rho < 1$ such that $\max_{2\le i\le N}\lambda_i \le \rho$, and so there exists $p = p(\rho)$ such that $\max_{2\le i\le N}(-\frac{\beta}{2}(1-\lambda_i)+\frac{\beta p}{2}\lambda_i) \le 0$. Combining we have

$$\sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left(g_p(\boldsymbol{\sigma}) - \frac{\beta}{2}\sum_{i=1}^N \sigma_i^2\right) \le o(N),$$

and so

$$\begin{aligned}\sup_{\boldsymbol{\sigma}\in[-1,1]^N}\bigl(g_p(\boldsymbol{\sigma}) - I(\boldsymbol{\sigma})\bigr) &\le \sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left(g_p(\boldsymbol{\sigma}) - \frac{\beta}{2}\sum_{i=1}^N \sigma_i^2\right) + \sup_{\boldsymbol{\sigma}\in[-1,1]^N}\left(\frac{\beta}{2}\sum_{i=1}^N \sigma_i^2 - I(\boldsymbol{\sigma})\right)\\ &= o(N) + \mathcal{M}_N(\beta, B),\end{aligned}$$

where $\mathcal{M}_N(\beta, B)$ is the mean-field prediction defined in (1.13). Since $|\log Z_N^{\mathrm{CW}}(\beta, B) - \mathcal{M}_N(\beta, B)| \lesssim \log N$ by part (a) of Proposition 6.1, (6.2) follows, thus completing the proof of the lemma. □

6.2. *Some results on matrices.*

LEMMA 6.1. *Let $\sum_{i=1}^N \lambda_i(A_N)\mathbf{q}_i\mathbf{q}_i^\top$ be the spectral decomposition of A_N. Suppose that* (1.5) *and* (1.7) *hold, and $\sum_{i=1}^N (R_i - 1) = o(N)$.*

(a) *Then $\|\mathbf{q}_1 - \mathbf{e}\|_2 = o(1)$, where $\mathbf{e} := N^{-1/2}\mathbf{1}$.*
(b) *Further we have $\limsup_{N\to\infty}\lambda_1(\mathcal{A}_N) < 1$, where $\mathcal{A}_N = A_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top$.*

PROOF. (a) Write $\mathbf{e} = \sum_{i=1}^N c_i\mathbf{q}_i$ with $c_1 > 0$ by Perron–Frobenius theorem, and note that

$$1 + o(1) = \frac{1}{N}\sum_{i=1}^N R_i = \mathbf{e}^\top A_N\mathbf{e} = \sum_{i=1}^N c_i^2\lambda_i(A_N) \le \lambda_1(A_N)c_1^2 + \lambda_2(A_N)\bigl(1 - c_1^2\bigr).$$

Along with (1.5) and (1.7), this gives $c_1^2 = 1 + o(1)$, and so $\langle\mathbf{q}_1, \mathbf{e}\rangle = c_1 = 1 + o(1)$, thus completing the proof of part (a).

(b) This follows on using part (a) to note that

$$\|\mathcal{A}_N\|_2 \le \left\|\sum_{i=2}^N \lambda_i(A_N)\mathbf{q}_i\mathbf{q}_i^\top\right\|_2 + \|\lambda_1(A_N)\mathbf{q_1}\mathbf{q_1}^\top - \mathbf{e}\mathbf{e}^\top\|_2 \le \lambda_2(A_N) + o(1),$$

and using (1.7). □

LEMMA 6.2. *Let Γ_N be an $N\times N$ symmetric matrix with nonnegative entries, such that $\mathbf{1}^\top\Gamma_N = \mathbf{1}^\top$ and Γ_N satisfies* (1.7). *Then the following conclusions hold*:

(a) *There exists $c > 0$ such that for all $\ell \ge 1$ and N large we have*

$$\max_{1\le i\le N}\Gamma_N^\ell(i,i) \le \frac{2}{N} + \frac{2}{e^{c\ell}}.$$

(b) *There exists $\delta > 0$ such that for all N large enough we have*

$$\max_{2\le i\le N-1}\bigl|\lambda_i(\Gamma_N)\bigr| \le 1 - \frac{\delta}{\log N}.$$

PROOF. (a) Setting $\lambda_i := \lambda_i(\Gamma_N)$ for simplicity of notation, let $\mathcal{J}_+ := \{j \in [2, N] : \lambda_j > 0\}$ and $\mathcal{J}_- := \{j \in [2, N] : \lambda_j < 0\}$, and use the spectral theorem to note that for any positive integer ℓ we have

$$\Gamma_N^\ell = \frac{1}{N}\mathbf{1}\mathbf{1}^\top + \sum_{j\in\mathcal{J}_+} |\lambda_j|^\ell \mathbf{q}_j\mathbf{q}_j^\top + (-1)^\ell \sum_{j\in\mathcal{J}_-} |\lambda_j|^\ell \mathbf{q}_j\mathbf{q}_j^\top,$$

where $(\mathbf{q}_1, \dots, \mathbf{q}_N)$ are the eigenvectors of Γ_N. To begin, use (1.7) to note the existence of $c > 0$ such that for all N large enough we have $\lambda_2 \le e^{-c}$, which gives

$$\sum_{j\in\mathcal{J}_+} |\lambda_j|^\ell q_{ij}^2 \le \lambda_2^\ell \le e^{-c\ell}, \tag{6.4}$$

where q_{ij} denotes the ith entry of the vector $\mathbf{q}_j$.

For ℓ odd, noting that $\Gamma_N^\ell(i,i) \ge 0$ gives

$$\sum_{j\in\mathcal{J}_-} |\lambda_j|^\ell q_{ij}^2 \le \frac{1}{N} + \sum_{j\in\mathcal{J}_+} |\lambda_j|^\ell q_{ij}^2 \le \frac{1}{N} + \lambda_2^\ell \le \frac{1}{N} + e^{-c\ell},$$

where the last inequality uses (6.4). Using the fact that $\max_{2\le i\le N} |\lambda_i| \le 1$, for $\ell \ge 2$ we have

$$\sum_{j\in\mathcal{J}_-} |\lambda_j|^\ell q_{ij}^2 \le \sum_{j\in\mathcal{J}_-} |\lambda_j|^{\ell-1} q_{ij}^2 \le \frac{1}{N} + e^{-c\ell}.$$

Combining these two bounds, for all $\ell \ge 1$ we have

$$\left|\Gamma_N^\ell(i,i)\right| \le \frac{1}{N} + \sum_{j\in\mathcal{J}_+} |\lambda_j|^\ell q_{ij}^2 + \sum_{j\in\mathcal{J}_-} |\lambda_j|^\ell q_{ij}^2 \le \frac{2}{N} + \frac{2}{e^{c\ell}},$$

thus completing the proof of part (a).

(b) Let $\delta > 0$ be such that $3e^{-2\delta/c} > 2$. Using part (a) with $\ell = \frac{2\log N}{c}$ and even, we have

$$\sum_{i=1}^N |\lambda_i|^\ell = \sum_{i=1}^N \Gamma_N^\ell(i,i) \le 2 + 2Ne^{-2\log N} \to 2.$$

On the other hand, if $\max_{2\le i\le N-1} |\lambda_i| > 1 - \frac{\delta}{\log N}$, then

$$\sum_{i=1}^N |\lambda_i|^\ell \ge 3\left(1 - \frac{\delta}{\log N}\right)^{\frac{2\log N}{c}} \to 3e^{-\frac{2\delta}{c}}.$$

These two together imply $3e^{-2\delta/c} \le 2$, a contradiction. □

REMARK 6.1. Note that if Γ_N is the adjacency matrix of a d_N regular bipartite graph scaled by the degree d_N, which satisfies the spectral gap condition, see (1.7)), then our lemma implies

$$\lim_{N\to\infty} \max_{1\le i\le N} \left|N\Gamma_N^{2\ell}(i,i) - 2\right| = 0$$

for $\ell = D\log N$ with D large enough. This highlights the asymptotic optimality of the bound obtained in part (a) of Lemma 6.2. Part (b) quantifies the graph theoretic fact that for a connected d_N regular graph, say G_N, the multiplicity of the eigenvalue $-d_N$ can be at most 1. It is easy to check that if $-d_N$ happens to be an eigenvalue the graph must be a bipartite graph, and all other eigenvalues will be strictly larger than $-d_N$ (i.e., there is a unique bipartition for a connected bipartite graph). In fact, our proof can be modified to show the stronger conclusion that for a d_N regular bipartite graphs satisfying the spectral gap condition, the second last eigenvalue is bounded away from -1, that is,

$$\liminf_{N\to\infty} \frac{\lambda_{N-1}(G_N)}{d_N} > -1.$$

6.3. *Some results for the Curie–Weiss model.* The following proposition collects all the results for the Curie–Weiss model which we have used previously.

PROPOSITION 6.1. *Suppose $\boldsymbol{\sigma}$ is drawn from the Curie–Weiss model. With $\widetilde{W}_N$ as in Proposition 4.1, the following conclusions hold*:

(a)

$$\log Z_N^{\mathrm{CW}}(\beta, B) - N\left\{\frac{\beta}{2}t^2 + Bt - I(t)\right\} \lesssim 1 \quad \text{if } (\beta, B) \in \Theta_1 \cup \Theta_2$$
$$\lesssim \log N \quad \text{if } (\beta, B) \in \Theta_3.$$

(b) *For any $\lambda > 0$, we have*

$$\log \mathbb{P}^{\mathrm{CW}}\big(\big|\widetilde{W}_N - M(\boldsymbol{\sigma})\big| \geq \lambda\big) \lesssim -N\lambda^2 \quad \text{if } (\beta, B) \in \Theta_1 \cup \Theta_2$$
$$\lesssim -N\min\big(\lambda^2, \lambda^4\big) \quad \text{if } (\beta, B) \in \Theta_3.$$

Consequently for any sequence $\delta_N = o(N)$ we have

$$\log \mathbb{E}^{\mathrm{CW}} e^{\delta_N(\widetilde{W}_N - M(\boldsymbol{\sigma}))^2} \lesssim 1 \quad \text{if } (\beta, B) \in \Theta_1 \cup \Theta_2$$
$$\lesssim \frac{\delta_N^2}{N} \quad \text{if } (\beta, B) \in \Theta_3.$$

(c) *For $(\beta, B) \in \Theta_2$, we have*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}^{\mathrm{CW}}\left(\sum_{i=1}^N \sigma_i \in \{-2, -1, 0, 1, 2\}\right) < 0.$$

PROOF. (a) With $f(w) = \frac{\beta w^2}{2} - \log\cosh(\beta w + B)$ as in 4.1, a direct computation gives $Z_N^{\mathrm{CW}}(\beta, B) = e^{-\beta/2}\sqrt{\frac{n\beta}{2\pi}}\int_{\mathbb{R}} e^{-nf(w)}\,dw$, where the function $f(w)$ has a unique global minimum at $w = t$ for $(\beta, B) \in \Theta_1 \cup \Theta_3$, and two global minima at $\pm t$ for $(\beta, B) \in \Theta_2$. Also, it is easy to verify that

$$\begin{aligned} & f(w) - f(t) \cong (w - t)^2 \quad \text{for all } w \in \mathbb{R}, \text{if } (\beta, B) \in \Theta_1, \\ (6.5) \qquad & f(w) - f(t) \cong (w - t)^2 \quad \text{for all } w > 0, \text{if } (\beta, B) \in \Theta_2, \\ & f(w) - f(t) \cong \min\big[(w - t)^2, (w - t)^4\big] \quad \text{for all } w \in \mathbb{R}, \text{if } (\beta, B) \in \Theta_3. \end{aligned}$$

The desired estimates follow from these bounds and using the Laplace method for approximating integrals.

(b) Noting that

$$\big|\widetilde{W}_N - M(\boldsymbol{\sigma})\big| = \big|\tanh(\beta W_N + B) - \tanh\big(\beta M(\boldsymbol{\sigma}) + B\big)\big| \leq \beta\big|W_N - M(\boldsymbol{\sigma})\big|,$$

it suffices to prove the desired bounds W_N, which follows from straightforward computations on using (6.5).

(c) This follows on using part (b) to note that, when $(\beta, B) \in \Theta_2$, the random variable W_N has an exponential concentration near the points $\pm t$, none of which are near 0. □

6.4. *Proof of* (3.6). In this section, we will prove (3.6) using [14], Theorem 1.2, and a soft change of measure argument. Throughout this proof, $c > 0$ will denote constants free of N that might change from one line to the next.

PROOF. Define the set $\widetilde{\mathcal{J}} := \{\boldsymbol{\sigma} \in \{-1,1\}^N : |\sum_{i=1}^N \sigma_i| \geq 3\}$ and $\mathcal{J} := \{\boldsymbol{\sigma} \in \{-1,1\}^N : |\sum_{i=1}^N \sigma_i| \geq 4\}$. Recall the definition of $\mathbb{P}$ from (1.1) and note that, by part (c) of Proposition 6.1 and part (b) of Theorem 1.6, we get

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\boldsymbol{\sigma} \in \widetilde{\mathcal{J}}^c) \leq \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\boldsymbol{\sigma} \in \mathcal{J}^c) < 0. \tag{6.6}$$

Next, let $\mathbb{Q}$ denote the probability measure induced by $\mathbb{P}$ conditioned on the event $\boldsymbol{\sigma} \in \widetilde{\mathcal{J}}$, that is, $\mathbb{Q}(\cdot) := \mathbb{P}(\cdot | \boldsymbol{\sigma} \in \widetilde{\mathcal{J}})$. Therefore, for any $B \subseteq \{-1,1\}^N$, we have

$$\mathbb{Q}(\boldsymbol{\sigma} \in B) = \frac{\mathbb{P}(\boldsymbol{\sigma} \in B \cap \widetilde{\mathcal{J}})}{\mathbb{P}(\boldsymbol{\sigma} \in \widetilde{\mathcal{J}})}.$$

Once again, by part (c) of Proposition 6.1 and part (b) of Theorem 1.6, we get

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{Q}(\boldsymbol{\sigma} \in \mathcal{J}^c) \leq \limsup_{N\to\infty} \frac{1}{N} \log \frac{\mathbb{P}(\boldsymbol{\sigma} \in \mathcal{J}^c)}{\mathbb{P}(\boldsymbol{\sigma} \in \widetilde{\mathcal{J}})} < 0. \tag{6.7}$$

Suppose that we draw $\boldsymbol{\sigma}_{\mathbb{P}} \sim \mathbb{P}$ and $\boldsymbol{\sigma}_{\mathbb{Q}} \sim \mathbb{Q}$. Define $T_{N,\mathbb{Q}} := \sqrt{N}(\overline{\boldsymbol{\sigma}_{\mathbb{Q}}} - M(\boldsymbol{\sigma}_{\mathbb{Q}}))$. We will write $(T_{N,\mathbb{P}}, T'_{N,\mathbb{P}}) \equiv (T_N, T'_N)$ under the law of $\mathbb{P}$ (recall the construction of T'_N from the proof of Theorems 1.4 and 1.2 in Section 3.1). Construct $T'_{N,\mathbb{Q}}$ similar to $T'_{N,\mathbb{P}}$ as follows: Sample I uniformly from the set $\{1,2,\ldots,N\}$. Given $I = i$, replace $\sigma_{\mathbb{Q},i}$ with an independent ± 1 valued random variable $\sigma'_{\mathbb{Q},i}$ with mean $\mathbb{E}_{\mathbb{Q}}[\sigma_{\mathbb{Q},i} | (\sigma_{\mathbb{Q},j}, j \neq i)]$, and set $\boldsymbol{\sigma}'_{\mathbb{Q}} := (\sigma_{\mathbb{Q},1}, \ldots, \sigma_{\mathbb{Q},i-1}, \sigma'_{\mathbb{Q},i}, \sigma_{\mathbb{Q},i+1}, \ldots, \sigma_{\mathbb{Q},N})$, $T'_{N,\mathbb{Q}} := \sqrt{N}(\overline{\boldsymbol{\sigma}'_{\mathbb{Q}}} - M(\boldsymbol{\sigma}'_{\mathbb{Q}}))$.

By construction $(T'_{N,\mathbb{Q}}, T_{N,\mathbb{Q}})$ forms an exchangeable pair under $\mathbb{Q}$. Moreover,

$$\mathbb{Q}(\boldsymbol{\sigma}'_{\mathbb{Q}} | \boldsymbol{\sigma}_{\mathbb{Q}} = \boldsymbol{\sigma}) = \mathbb{P}(\boldsymbol{\sigma}'_{\mathbb{P}} | \boldsymbol{\sigma}_{\mathbb{P}} = \boldsymbol{\sigma}) \quad \text{for } \boldsymbol{\sigma} \in \mathcal{J}, \tag{6.8}$$

and

$$\mathbb{Q}(|T'_{N,\mathbb{Q}} - T_{N,\mathbb{Q}}| \leq 2N^{-1/2}) = 1,$$

which follows by observing that $\max_{i=1}^N |\sigma'_{\mathbb{Q},i} - \sigma_{\mathbb{Q},i}| \leq 2$.

Define $\delta := 2N^{-1/2}$. By using the above display, coupled with [14], Theorem 1.2, we get

$$\begin{aligned}
&\sup_{z\in\mathbb{R}} \big|\mathbb{Q}(T_{N,\mathbb{Q}} \leq z) - \mathbb{P}(Z_\tau \leq z)\big| \\
&\quad \lesssim \mathbb{E}_{\mathbb{Q}}\big|1 - (c_0/2)\mathbb{E}_{\mathbb{Q}}\big((T_{N,\mathbb{Q}} - T'_{N,\mathbb{Q}})^2 | T_{N,\mathbb{Q}}\big)\big| \\
&\qquad + c_1 \max\{(1, c_3)\}\delta + (c_0/c_1)\mathbb{E}_{\mathbb{Q}}\big|r(T_{N,\mathbb{Q}})\big| \\
&\qquad + \delta^3 c_0\{(2 + c_3/2)\mathbb{E}_{\mathbb{Q}}\big|c_0 g(T_{N,\mathbb{Q}})\big| + c_1 c_3/2\},
\end{aligned} \tag{6.9}$$

where Z_τ is defined as in Lemma 1.2 for $(\beta, B) \in \Theta_2$, $r(\cdot) := \sum_{a=1}^3 H_a(\cdot)$ with $g(\cdot)$, $\{H_a(\cdot)\}_{a=1,2,3}$ from (3.3).

In the remainder, we will quantify the cost of moving between the probability measures $\mathbb{P}$ and $\mathbb{Q}$ in (6.9). First, we present a claim which will be used to prove (3.6). The proof of this claim is deferred to the end of the proof.

Claim: Given any function $v(\cdot) : \{-1,1\}^N \to \mathbb{R}$, such that $\sup_{\boldsymbol{\sigma}\in\{-1,1\}^n} |v(\boldsymbol{\sigma})| \leq aN^b$ for constants a, b free of N,

$$\big|\mathbb{E}_{\mathbb{Q}} v(\boldsymbol{\sigma}_{\mathbb{Q}}) - \mathbb{E}_{\mathbb{P}} v(\boldsymbol{\sigma}_{\mathbb{P}})\big| \leq \exp(-cN), \tag{6.10}$$

where c depends only on a, b, and the implied constant in (6.6). We will now complete the rest of the proof assuming Claim (6.10). For any $z \in \mathbb{R}$, with $v_z(\boldsymbol{\sigma}) := \mathbb{1}(\sqrt{N}(\overline{\boldsymbol{\sigma}} - M(\boldsymbol{\sigma})) \le z)$, note that $\sup_{z\in\mathbb{R}} \sup_{\boldsymbol{\sigma}\in\{-1,1\}^N} v_z(\boldsymbol{\sigma}) \le 1$, which by (6.10) yields

$$\sup_{z\in\mathbb{R}} \left|\mathbb{E}_{\mathbb{Q}} v_z(\boldsymbol{\sigma}_{\mathbb{Q}}) - \mathbb{E}_{\mathbb{P}} v_z(\boldsymbol{\sigma}_{\mathbb{P}})\right| = \sup_{z\in\mathbb{R}} \left|\mathbb{Q}(T_{N,\mathbb{Q}} \le z) - \mathbb{P}(T_{N,\mathbb{P}} \le z)\right| \le \exp(-cN). \tag{6.11}$$

A similar computation as in (6.11) further yields

$$\max\{\mathbb{E}_{\mathbb{Q}}\left|r(T_{N,\mathbb{Q}})\right| - \mathbb{E}_{\mathbb{P}}|r(T_{N,\mathbb{P}})||, |\mathbb{E}_{\mathbb{Q}}|g(T_{N,\mathbb{Q}})| - \mathbb{E}_{\mathbb{P}}|g(T_{N,\mathbb{P}})|\} \le \exp(-cN). \tag{6.12}$$

Next, we will focus on the term $\mathbb{E}_{\mathbb{Q}}((T_{N,\mathbb{Q}} - T'_{N,\mathbb{Q}})^2 | T_{N,\mathbb{Q}})$ in (6.9). By (6.8), we have

$$\mathbb{E}_{\mathbb{Q}}[1 - (c_0/2)(T_{N,\mathbb{Q}} - T'_{N,\mathbb{Q}})^2 | \boldsymbol{\sigma}_{\mathbb{Q}} = \boldsymbol{\sigma}] = \mathbb{E}_{\mathbb{P}}[1 - (c_0/2)(T_{N,\mathbb{P}} - T'_{N,\mathbb{P}})^2 | \boldsymbol{\sigma}_{\mathbb{P}} = \boldsymbol{\sigma}] =: u(\boldsymbol{\sigma}),$$

for $\boldsymbol{\sigma} \in \mathcal{J}$. Therefore,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left|\mathbb{E}_{\mathbb{Q}}[u(\boldsymbol{\sigma})\mathbb{1}(\boldsymbol{\sigma} \in \mathcal{J}) | T_{N,\mathbb{Q}}]\right| &= (\mathbb{P}(\boldsymbol{\sigma} \in \mathcal{J}))^{-1} \mathbb{E}_{\mathbb{P}}\left|\mathbb{E}_{\mathbb{P}}[u(\boldsymbol{\sigma})\mathbb{1}(\boldsymbol{\sigma} \in \mathcal{J}) | T_{N,\mathbb{P}}]\right| \\ &= \mathbb{E}_{\mathbb{P}}\left|\mathbb{E}_{\mathbb{P}}[u(\boldsymbol{\sigma})\mathbb{1}(\boldsymbol{\sigma} \in \mathcal{J}) | T_{N,\mathbb{P}}]\right| + r_n,\end{aligned}$$

where $|r_n| \le \exp(-cN)$ by (6.6).

Using the above observation with (6.6) and (6.7), we further get

$$\begin{aligned}&\left|\mathbb{E}_{\mathbb{Q}}\left|\mathbb{E}_{\mathbb{Q}}[1 - (c_0/2)(T_{N,\mathbb{Q}} - T'_{N,\mathbb{Q}})^2 | T_{N,\mathbb{Q}}]\right| - \mathbb{E}_{\mathbb{P}}\left|\mathbb{E}_{\mathbb{P}}[1 - (c_0/2)(T_{N,\mathbb{P}} - T'_{N,\mathbb{P}})^2 | T_{N,\mathbb{P}}]\right|\right| \\ &\quad \lesssim \exp(-cN) + N\mathbb{Q}(\boldsymbol{\sigma} \in \mathcal{J}^c) + N\mathbb{P}(\boldsymbol{\sigma} \in \mathcal{J}^c) \lesssim \exp(-cN).\end{aligned}$$

Combining the above observation with (6.12), (6.11), and (6.9), completes the proof of (3.6).

To complete the proof, it remains to prove (6.10), which is done below.

Proof of Claim (6.10): Observe that

$$\begin{aligned}\left|\mathbb{E}_{\mathbb{Q}} v(\boldsymbol{\sigma}_{\mathbb{Q}}) - \mathbb{E}_{\mathbb{P}} v(\boldsymbol{\sigma}_{\mathbb{P}})\right| &= \left|\frac{\mathbb{E}_{\mathbb{P}}[v(\boldsymbol{\sigma}_{\mathbb{P}})\mathbb{1}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J})]}{\mathbb{P}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J})} - \mathbb{E}_{\mathbb{P}} v(\boldsymbol{\sigma}_{\mathbb{P}})\right| \\ &\le \frac{\mathbb{E}_{\mathbb{P}}[|v(\boldsymbol{\sigma}_{\mathbb{P}})|\mathbb{1}(v(\boldsymbol{\sigma}_{\mathbb{P}}) \in \mathcal{J})]\mathbb{P}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J}^c)}{\mathbb{P}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J})} \\ &\quad + \mathbb{E}_{\mathbb{P}}[|v(\boldsymbol{\sigma}_{\mathbb{P}})|\mathbb{1}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J}^c)] \\ &\lesssim aN^b \mathbb{P}(\boldsymbol{\sigma}_{\mathbb{P}} \in \mathcal{J}^c) \le \exp(-cN),\end{aligned}$$

where the last line follows from (6.6). This establishes Claim (6.10). □

Acknowledgments. The authors would like to thank the Editor, the Associate Editor, and the two anonymous reviewers for their constructive suggestions that helped improve the presentation of this paper.

Funding. SM gratefully acknowledges the partial support of NSF (DMS-1712037) during this research.

REFERENCES

[1] ADAMCZAK, R., KOTOWSKI, M., POLACZYK, B. and STRZELECKI, M. (2019). A note on concentration for polynomials in the Ising model. *Electron. J. Probab.* **24** Paper No. 42, 22 pp. MR3949267 https://doi.org/10.1214/19-EJP280

[2] AUGERI, F. (2021). A transportation approach to the mean-field approximation. *Probab. Theory Related Fields* **180** 1–32. MR4265016 https://doi.org/10.1007/s00440-021-01056-2

[3] BANDEIRA, A. S. and VAN HANDEL, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.* **44** 2479–2506. MR3531673 https://doi.org/10.1214/15-AOP1025

[4] BASAK, A. and MUKHERJEE, S. (2017). Universality of the mean-field for the Potts model. *Probab. Theory Related Fields* **168** 557–600. MR3663625 https://doi.org/10.1007/s00440-016-0718-0

[5] BERTHET, Q., RIGOLLET, P. and SRIVASTAVA, P. (2019). Exact recovery in the Ising blockmodel. *Ann. Statist.* **47** 1805–1834. MR3953436 https://doi.org/10.1214/17-AOS1620

[6] BORGS, C., CHAYES, J. T., COHN, H. and ZHAO, Y. (2018). An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence. *Ann. Probab.* **46** 337–396. MR3758733 https://doi.org/10.1214/17-AOP1187

[7] BORGS, C., CHAYES, J. T., COHN, H. and ZHAO, Y. (2019). An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Trans. Amer. Math. Soc.* **372** 3019–3062. MR3988601 https://doi.org/10.1090/tran/7543

[8] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626 https://doi.org/10.1016/j.aim.2008.07.008

[9] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382 https://doi.org/10.4007/annals.2012.176.1.2

[10] BRESLER, G. and NAGARAJ, D. (2019). Stein's method for stationary distributions of Markov chains and application to Ising models. *Ann. Appl. Probab.* **29** 3230–3265. MR4019887 https://doi.org/10.1214/19-AAP1479

[11] BRODER, A. and SHAMIR, E. (1987). On the second eigenvalue of random regular graphs. In 28*th Annual Symposium on Foundations of Computer Science* (*SFCS* 1987) 286–294. IEEE, New York.

[12] CHATTERJEE, S. (2005). *Concentration Inequalities with Exchangeable Pairs*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)—Stanford Univ. MR2707160

[13] CHATTERJEE, S. and DEMBO, A. (2016). Nonlinear large deviations. *Adv. Math.* **299** 396–450. MR3519474 https://doi.org/10.1016/j.aim.2016.05.017

[14] CHATTERJEE, S. and SHAO, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964 https://doi.org/10.1214/10-AAP712

[15] CHUANG, H. and OMIDI, G. R. (2009). Graphs with three distinct eigenvalues and largest eigenvalues less than 8. *Linear Algebra Appl.* **430** 2053–2062. MR2503952 https://doi.org/10.1016/j.laa.2008.11.028

[16] COMETS, F. and GIDAS, B. (1991). Asymptotics of maximum likelihood estimators for the Curie–Weiss model. *Ann. Statist.* **19** 557–578. MR1105836 https://doi.org/10.1214/aos/1176348111

[17] DEMBO, A. and MONTANARI, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563 https://doi.org/10.1214/09-BJPS027

[18] DESHPANDE, Y., SEN, S., MONTANARI, A. and MOSSEL, E. (2018). Contextual stochastic block models. In *Advances in Neural Information Processing Systems* 8581–8593.

[19] EICHELSBACHER, P. and LÖWE, M. (2010). Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15** 962–988. MR2659754 https://doi.org/10.1214/EJP.v15-777

[20] ELDAN, R. (2020). Taming correlations through entropy-efficient measure decompositions with applications to mean-field approximation. *Probab. Theory Related Fields* **176** 737–755. MR4087482 https://doi.org/10.1007/s00440-019-00924-2

[21] ELLIS, R. S. and NEWMAN, C. M. (1978). The statistics of Curie–Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332 https://doi.org/10.1007/BF01012508

[22] FEIGE, U. and OFEK, E. (2005). Spectral techniques applied to sparse random graphs. *Random Structures Algorithms* **27** 251–275. MR2155709 https://doi.org/10.1002/rsa.20089

[23] GHEISSARI, R., LUBETZKY, E. and PERES, Y. (2018). Concentration inequalities for polynomials of contracting Ising models. *Electron. Commun. Probab.* **23** Paper No. 76, 12 pp. MR3873783 https://doi.org/10.1214/18-ECP173

[24] GIARDINÀ, C., GIBERTI, C., VAN DER HOFSTAD, R. and PRIORIELLO, M. L. (2016). Annealed central limit theorems for the Ising model on random graphs. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 121–161. MR3476210

[25] JAIN, V., RISTESKI, A. and KOEHLER, F. (2019). Mean-field approximation, convex hierarchies, and the optimality of correlation rounding: A unified perspective. In *STOC'*19—*Proceedings of the* 51*st Annual ACM SIGACT Symposium on Theory of Computing* 1226–1236. ACM, New York. MR4003424 https://doi.org/10.1145/3313276.3316299

[26] KABLUCHKO, Z., LÖWE, M. and SCHUBERT, K. (2019). Fluctuations of the magnetization for Ising models on dense Erdős–Rényi random graphs. *J. Stat. Phys.* **177** 78–94. MR4003721 https://doi.org/10.1007/s10955-019-02358-5

[27] KABLUCHKO, Z., LÖWE, M. and SCHUBERT, K. (2022). Fluctuations of the magnetization for Ising models on Erdős–Rényi random graphs—The regimes of low temperature and external magnetic field. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 537–563. MR4394308 https://doi.org/10.30757/alea.v19-21

[28] KEARNS, M. and SAUL, L. (1998). Large deviation methods for approximate probabilistic inference. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* 311–319.

[29] KIRSCH, W. and TOTH, G. (2020). Two groups in a Curie–Weiss model with heterogeneous coupling. *J. Theoret. Probab.* **33** 2001–2026. MR4166190 https://doi.org/10.1007/s10959-019-00933-w

[30] LIU, L. (2017). On the log partition function of Ising model on stochastic block model. Preprint. Available at arXiv:1710.05287.

[31] LOVÁSZ, L. (2012). *Large Networks and Graph Limits*. *American Mathematical Society Colloquium Publications* **60**. Amer. Math. Soc., Providence, RI. MR3012035 https://doi.org/10.1090/coll/060

[32] LÖWE, M. and SCHUBERT, K. (2018). Fluctuations for block spin Ising models. *Electron. Commun. Probab.* **23** Paper No. 53, 12 pp. MR3852267 https://doi.org/10.1214/18-ECP161

[33] MOSSEL, E., NEEMAN, J. and SLY, A. (2012). Stochastic block models and reconstruction. Preprint. Available at arXiv:1202.1499.

[34] MUKHERJEE, R., MUKHERJEE, S. and YUAN, M. (2018). Global testing against sparse alternatives under Ising models. *Ann. Statist.* **46** 2062–2093. MR3845011 https://doi.org/10.1214/17-AOS1612

[35] MUKHERJEE, R. and RAY, G. (2022). On testing for parameters in Ising models. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 164–187. MR4375621 https://doi.org/10.1214/21-aihp1157

[36] MUKHERJEE, S. and XU, Y. (2023). Statistics of the two-star ERGM. *Bernoulli* **29** 24–51. MR4497238 https://doi.org/10.3150/21-BEJ1448

[37] OSTROVSKY, E. and SIROTA, L. (2014). Exact value for subgaussian norm of centered indicator random variable. Available at arXiv:1405.6749.

[38] RAVIKUMAR, P., WAINWRIGHT, M. J. and LAFFERTY, J. D. (2010). High-dimensional Ising model selection using ℓ_1-regularized logistic regression. *Ann. Statist.* **38** 1287–1319. MR2662343 https://doi.org/10.1214/09-AOS691

[39] SLY, A. and SUN, N. (2014). Counting in two-spin models on d-regular graphs. *Ann. Probab.* **42** 2383–2416. MR3265170 https://doi.org/10.1214/13-AOP888