Commun. Math. Phys. 387, 681–728 (2021)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-021-04182-z

Communications in
Mathematical Physics

Fluctuations of the Magnetization in the *p*-Spin Curie–Weiss Model

Somabha Mukherjee[1], **Jaesung Son**[2], **Bhaswar B. Bhattacharya**[3]

[1] Department of Statistics and Data Science, National University of Singapore, Singapore, Singapore.
E-mail: somabhamukherjee@gmail.com
[2] Department of Statistics, Columbia University, New York, USA.
E-mail: js4638@columbia.edu
[3] Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, USA.
E-mail: bhaswar@wharton.upenn.edu

Received: 21 June 2020 / Accepted: 16 July 2021
Published online: 9 September 2021 –

Abstract: In this paper we study the fluctuations of the magnetization in the p-spin Curie–Weiss model, for $p \geqslant 3$. We provide a complete description of the asymptotic distribution of the magnetization in the p-spin Curie–Weiss model, complementing the well-known results in the 2-spin case. Our results unearth various new phase transitions, such as the existence of a certain 'critical' curve in the parameter space, where the limiting distribution of the magnetization is a discrete mixture, with local Gaussian fluctuations around each of the atoms. The number of atoms (mixture components) is either two or three depending on the sign of one of the parameters and the parity of p. Another interesting revelation is the existence of certain 'special' points in the parameter space where the magnetization converges to a non-Gaussian limiting distribution at rate $N^{\frac{1}{4}}$.

1. Introduction

The Ising model is a discrete random field, where the Hamiltonian has a quadratic term designed to capture pairwise interactions between neighboring vertices of a network. This was initially studied almost a century ago as a model for ferromagnetism [24], and has since then emerged as one of the fundamental mathematical tools for understanding interacting spin systems on graphs. Recently, the Ising model has also turned out to be a useful primitive for capturing pairwise dependence among binary attributes with an underlying network structure, which arise naturally in spatial statistics, social networks, computer vision, neural networks, and computational biology, among others (cf. [1,9, 18,20,23,28] and the references therein). However, in many situations, both in modeling interacting spin systems and in real-world network data, dependencies arise not just from pairs, but from interactions between groups of particles or individuals. This leads to the study of p-spin Ising models, where the Hamiltonian is a multilinear polynomial of

Somabha Mukherjee and Jaesung Son have contributed equally to the paper.

degree $p \geqslant 2$, designed for capturing higher-order interactions between the different particles. As in the case of 2-spin models, the p-spin Ising model can be represented as a spin system on a p-uniform hypergraph, where the individual entities represent the vertices of the hypergraph and the p-tuples of interactions are indexed by the hyperedges. Higher-order Ising models arise naturally in the study of multi-atom interactions in lattice gas models, which includes, among others, the square-lattice eight-vertex model, the Ashkin-Teller model, and Suzuki's pseudo-3D anisotropic model (cf. [2,21,29–31,33, 34,36] and the references therein). More recently, higher-order spin systems have been proposed as effective and mathematically tractable models for simultaneously capturing both peer-group effects and individual effects in social networks [8].

In the 2-spin case, one of the most extensively studied models is the classical Curie–Weiss model [6,11,14,16,27], where all the pairwise interactions between the nodes of the network are present (the Ising model on the complete graph). This model preserves several interesting properties of general systems and plays a fundamental role in the understanding of mean-field models with pairwise interactions. The asymptotic distribution of the magnetization (the average of the coordinates of the spin configuration) in the 2-spin Curie–Weiss model is known from the celebrated results of Ellis and Newman [16]. Recently, the fluctuations of the magnetization has also been studied for Ising models on random graphs (cf. [3,4,19,25,26] and the references therein) and general regular graphs [10].

The 2-spin Curie–Weiss model naturally extends to the p-spin Curie–Weiss model, for any $p \geqslant 2$, in which the Hamiltonian has all the possible p-tuples of interactions. More precisely, given an inverse temperature $\beta \geqslant 0$ and a magnetic field $h \in \mathbb{R}$, the p-spin Curie–Weiss model is a spin system on $\mathcal{C}_N := \{-1, 1\}^N$ defined as:

$$\mathbb{P}_{\beta,h,p}(\boldsymbol{\sigma}) = \frac{\exp\left\{\frac{\beta}{N^{p-1}} \sum_{1 \leqslant i_1, i_2, \ldots, i_p \leqslant N} \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p} + h\sum_{i=1}^N \sigma_i\right\}}{2^N Z_N(\beta, h, p)}, \tag{1.1}$$

for $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_N) \in \mathcal{C}_N$. The normalizing constant, also referred to as the partition function, $Z_N(\beta, h, p)$ is determined by the condition $\sum_{\boldsymbol{\sigma} \in \mathcal{C}_N} \mathbb{P}_{\beta,h,p}(\boldsymbol{\sigma}) = 1$, that is,

$$Z_N(\beta, h, p) = \frac{1}{2^N} \sum_{\boldsymbol{\sigma} \in \mathcal{C}_N} \exp\left\{\frac{\beta}{N^{p-1}} \sum_{1 \leqslant i_1, i_2, \ldots, i_p \leqslant N} \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p} + h\sum_{i=1}^N \sigma_i\right\}. \tag{1.2}$$

Denote by $F_N(\beta, h, p) := \log Z_N(\beta, h, p)$ the log-partition function of the model. Hereafter, we will often abbreviate $\mathbb{P}_{\beta,h,p}$, $Z_N(\beta, h, p)$, and $F_N(\beta, h, p)$, by $\mathbb{P}$, Z_N, and F_N, respectively, when there is no scope of confusion. Various thermodynamic properties of this model, which is alternatively referred to as the fully connected p-spin model or the ferromagnetic p-spin model, are studied in [2,29,33,36].

This paper studies the fluctuations of the (average) magnetization $\overline{\sigma}_N := \frac{1}{N}\sum_{i=1}^N \sigma_i$, given a sample $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_N) \sim \mathbb{P}_{\beta,h,p}$ from the p-spin Curie–Weiss model. While this has been extensively studied for the $p = 2$ case [14–17], to the best of our knowledge this is the first such result for the higher order ($p \geqslant 3$) Curie–Weiss model. In this paper we provide a complete description of the asymptotic distribution $\overline{\sigma}_N$ for the p-spin Curie–Weiss model, for $p \geqslant 3$. We provide a brief summary of the results below:

- We identify a region of 'regular' points in the parameter space where $\overline{\sigma}_N$ concentrates at a unique point and has fluctuations of order $N^{\frac{1}{2}}$ with a limiting Gaussian distribution centered around this point.

- More interestingly, there are certain 'critical' points, which form a 1-dimensional curve in the parameter space, where $\overline{\sigma}_N$ concentrates at either two or three points. In other words, $\overline{\sigma}_N$ converges to a discrete distribution with either two or three atoms for (β, h) on this critical curve. In particular, if $h \neq 0$ or p is odd, then $\overline{\sigma}_N$ concentrates at two points along this curve. On the other hand, when $h = 0$ and p is even, there is an addition (strongly) critical point, where $\overline{\sigma}_N$ concentrates at three points. Moreover, $\overline{\sigma}_N$ has Gaussian fluctuations centered around each of the atoms, when conditioned to lie in their respective neighborhoods.
- Finally, there are one or two 'special' points in the parameter space, depending on whether $p \geqslant 3$ is odd or even, respectively, where $\overline{\sigma}_N$ has fluctuations of the order $N^{-\frac{1}{4}}$ and a non-Gaussian limiting distribution.

The formal statement of the result is given in Sect. 2. The proofs require precise approximations of the partition function Z_N and a careful understanding of the maximizers of a certain mean-field variational problem at all points in the parameter space. One of the technical bottlenecks in dealing with p-spin models is the absence of the 'Gaussian transform', which allows one to relate the partition function with certain Gaussian integrals in models with quadratic Hamiltonians, as in the 2-spin Curie–Weiss model. This method, unfortunately, does not apply when $p \geqslant 3$, hence, to estimate the partition function we have to use a more bare-hands combinatorial approach. The details of the proof are given in Sect. 3. In Sect. 4 we discuss future directions. Various technical details are given in the "Appendix".

2. Statements of the Main Results

In this section we state our main results on the limiting properties of the magnetization in the p-spin Curie–Weiss model. The asymptotics of the magnetization are described in Sect. 2.1, and in Sect. 2.2 we summarize our results in a phase diagram.

2.1. Limiting distribution of the magnetization. The fundamental quantity of interest in understanding the asymptotic behavior of the p-spin Curie–Weiss model is the magnetization $\overline{\sigma}_N = \frac{1}{N}\sum_{i=1}^{N}\sigma_i$. As alluded to before, the limiting properties of $\overline{\sigma}_N$ has been carefully studied for the case $p = 2$ [7,14]. Here, we will consider the case $p \geqslant 3$, where, as discussed below, many surprises and interesting new phase transitions emerge.

In order to state the results we need a few definitions: For $p \geqslant 2$ and $(\beta, h) \in \Theta := [0, \infty) \times \mathbb{R}$, define the function $H = H_{\beta,h,p} : [-1, 1] \to \mathbb{R}$ as

$$H(x) := \beta x^p + hx - I(x), \tag{2.1}$$

where $I(x) := \frac{1}{2}\{(1+x)\log(1+x) + (1-x)\log(1-x)\}$, for $x \in [-1, 1]$, is the binary entropy function. The points of maxima of this function will determine the typical values of $\bar{\sigma}_N$ and, hence, play a crucial role in our results. A careful analysis of the function H (see "Appendix B.1") reveals that it can have one, two, or three global maximizers in the open interval $(-1, 1)$, which leads to the following definition:[1]

Definition 1. Fix $p \geqslant 2$ and $(\beta, h) \in \Theta$, and let H be as defined above in (2.1).

[1] For a smooth function $f : [-1, 1] \to \mathbb{R}$ and $x \in (-1, 1)$, the first and second derivatives of f at the point x will be denoted by $f'(x)$ and $f''(x)$, respectively. More generally, for $s \geqslant 3$, the s-th order derivative of f at the point x will be denoted by $f^{(s)}(x)$.

(1) The point (β, h) is said to be *p-regular*, if the function $H_{\beta,h,p}$ has a unique global maximizer $m_* = m_*(\beta, h, p) \in (-1, 1)$ and $H''_{\beta,h,p}(m_*) < 0$. (Note that a point $m \in (-1, 1)$ is said to be a global maximizer of H if $H(m) > H(x)$, for all $x \in [-1, 1]\backslash\{m\}$.) Denote the set of all p-regular points in Θ by $\mathcal{R}_p$.
(2) The point (β, h) is said to be *p-special*, if $H_{\beta,h,p}$ has a unique global maximizer $m_* = m_*(\beta, h, p) \in (-1, 1)$ and $H''_{\beta,h,p}(m_*) = 0$.
(3) The point (β, h) is said to be *p-critical*, if $H_{\beta,h,p}$ has more than one global maximizer.

Note that the three cases above form a disjoint partition of the parameter space Θ. Hereafter, we denote the set of p-critical points by $\mathscr{C}_p$, and the set of points (β, h) where $H_{\beta,h,p}$ has exactly two global maximizers by $\mathscr{C}_p{}^+$. We show in Lemma B.3 that the set of points in $\mathscr{C}_p$ form a continuous 1-dimensional curve in the parameter space Θ (see also Figs. 6 and 7). Next, we consider points with three global maximizers, that is $\mathscr{C}_p \backslash \mathscr{C}_p{}^+$. To this end, define

$$\tilde{\beta}_p := \sup \left\{ \beta \geqslant 0 : \sup_{x \in [-1,1]} H_{\beta,0,p}(x) = 0 \right\}. \tag{2.2}$$

Alternatively, Lemma B.3 shows that $\tilde{\beta}_p$ is the smallest $\beta \geqslant 0$ for which the point $(\beta, 0)$ is p-critical. Now, depending on whether p is odd or even we have the following two cases:

- $p \geqslant 3$ odd: In this case Lemma B.1 shows that, for all points $(\beta, h) \in \mathscr{C}_p$, the function $H_{\beta,h,p}$ has exactly two global maximizers, that is, $\mathscr{C}_p = \mathscr{C}_p{}^+$.
- $p \geqslant 4$ even: Here, Lemma B.1 shows that there is a unique point $\lambda_p := (\tilde{\beta}_p, 0) \in \mathscr{C}_p$, with $\tilde{\beta}_p$ as defined in (2.2), at which the function $H_{\tilde{\beta}_p,0,p}$ has exactly three global maximizers. For all other points in $(\beta, h) \in \mathscr{C}_p$, $H_{\beta,h,p}$ has exactly two global maximizers, that is, $\mathscr{C}_p = \mathscr{C}_p{}^+ \cup \{\lambda_p\}$. In this case we will refer to the point λ_p, or, equivalently, the point $\tilde{\beta}_p$, as the *p-strongly critical* point.[2] Hereafter, when the need will arise to distinguish strongly critical points from other critical points, we will refer to a point which is p-critical but not p-strongly critical, as *p-weakly critical*. Note that the collection of all p-weakly critical points is precisely the set $\mathscr{C}_p{}^+$.

It remains to describe the structure of p-special points. To this end, fix $p \geqslant 3$ and define the following quantities:

$$\check{\beta}_p := \frac{1}{2(p-1)} \left(\frac{p}{p-2} \right)^{\frac{p-2}{2}} \quad \text{and} \quad \check{h}_p := \tanh^{-1}\left(\sqrt{\frac{p-2}{p}} \right) - \check{\beta}_p p \left(\frac{p-2}{p} \right)^{\frac{p-1}{2}}. \tag{2.3}$$

Again, depending on whether p is even or odd there are two cases:

- $p \geqslant 3$ odd: In this case, Lemma B.2 shows that there is only one p-special point $\tau_p := (\check{\beta}_p, \check{h}_p)$.
- $p \geqslant 4$ even: Here, again from Lemma B.2 and the symmetry of the model about $h = 0$, there are two p-special points $\tau_p^+ := (\check{\beta}_p, \check{h}_p)$ and $\tau_p^- := (\check{\beta}_p, -\check{h}_p)$.

These points are especially interesting, because, as we will see in a moment, here the magnetization has fluctuations of order $N^{\frac{1}{4}}$ and a non-Gaussian limiting distribution.

[2] Note that the point $\tilde{\beta}_p$ is defined for all $p \geqslant 2$ (even or odd) as in (2.2). However, for $p \geqslant 3$ odd, this point is p-critical, but not p-strongly critical (that means it belongs to $\mathscr{C}_p^+$). On the other hand, for $p = 2$ this point is 2-special (see discussion in Remark 2.1).

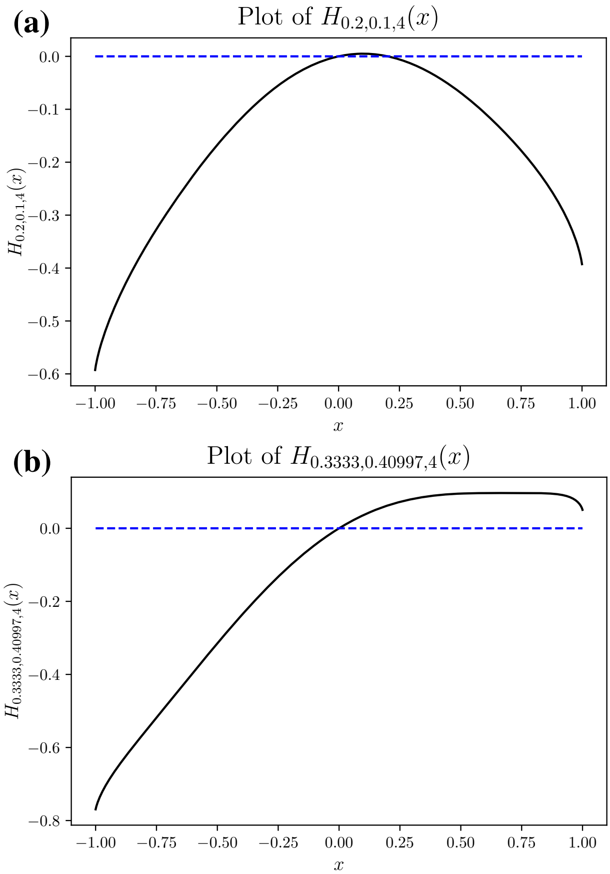

Fig. 1. (**a**) Plot of the function $H_{\beta,h,p}$ at the 4-regular point $(\beta, h) = (0.2, 0.1)$, where the function $H_{\beta,h,p}$ has a single global maximizer and the second derivative is negative at the maximizer; (**b**) Plot of the function $H_{\beta,h,p}$ at the 4-special point $(\beta, h) = (0.3333, 0.40997)$, where the function $H_{\beta,h,p}$ has a single global maximizer, but the second derivative is zero at the maximizer

The plots in Figs. 1 and 2 show instances of the different cases described above: Fig. 1a shows the plot of the function $H_{\beta,h,p}$ at the 4-regular point $(\beta, h) = (0.2, 0.1)$, and Fig. 1b shows the plot of the function $H_{\beta,h,p}$ at the 4-special point $(\beta, h) = (0.3333, 0.40997)$. On the other hand, Fig. 2a shows the plot of the function $H_{\beta,h,p}$ at the 3-critical point $(\beta, h) = (0.57, 0.12159)$, which has two global maximizers, and Fig. 2b shows the plot of the function at the 4-strongly critical point $(\beta, h) = (0.688, 0)$, where the function $H_{\beta,h,p}$ has three global maximizers. In fact, recalling that $\mathcal{R}_p$ denotes the set of all p-regular points and $\mathscr{C}_p^+$ the set of points (β, p) where $H_{\beta,h,p}$ has exactly two maximizers, the discussion above can be summarized as follows:

$$\Theta = \begin{cases} \mathcal{R}_p \bigcup \mathscr{C}_p^+ \bigcup \{\tau_p\} & \text{for } p \geqslant 3 \text{ odd}, \\ \mathcal{R}_p \bigcup \mathscr{C}_p^+ \bigcup \{\lambda_p, \tau_p^+, \tau_p^-\} & \text{for } p \geqslant 4 \text{ even}. \end{cases} \tag{2.4}$$

Figures 6 and 7 illustrates this decomposition of the parameter space for $p = 4$ and $p = 5$, respectively.

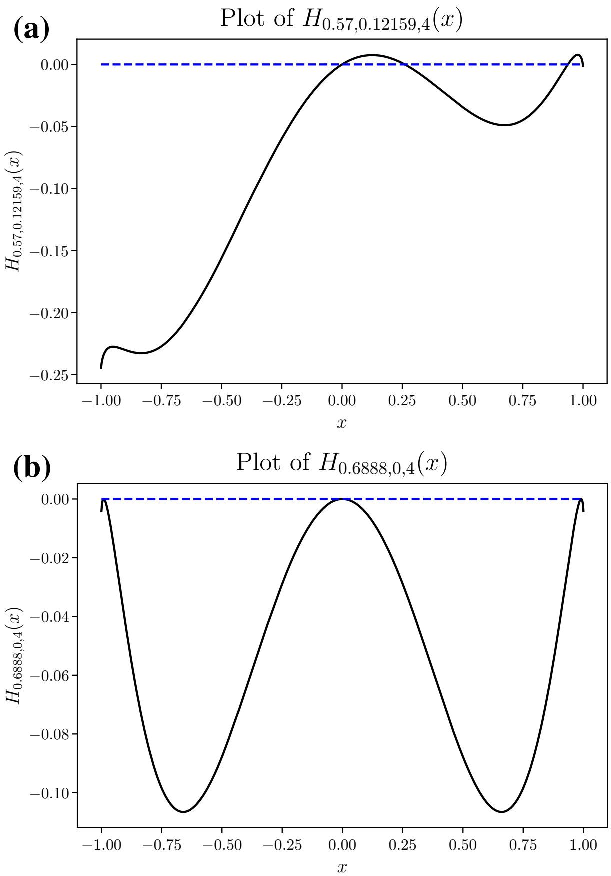

Fig. 2. Plots of the function $H_{\beta,h,p}$ at p-critical points. For the plot in (**a**) $p = 4$ and $(\beta, h) = (0.57, 0.12159)$ and the function $H_{\beta,h,p}$ has two global maximizers; and for (**b**) $p = 4$ and $(\beta, h) = (0.688, 0)$ and the function $H_{\beta,h,p}$ has three global maximizers, that is, the point $(0.688, 0)$ is 4-strongly critical

Remark 2.1. Note that (2.4) provides a complete characterization of the parameter space for $p \geqslant 3$. As mentioned before, in the well-studied case of $p = 2$, the situation is relatively simpler [11,14]. In this case, $H_{\beta,h,p}$ can have at most two global maximizers, that is, it has no strongly critical points, hence, $\mathcal{C}_2 = \mathcal{C}_2^+$. In fact, it follows from [14] that the set of points (β, h) with exactly two global maximizers $\mathcal{C}_2^+$ is the open half-line $(0.5, \infty) \times \{0\}$. Moreover, there is a single 2-special point $(0.5, 0)$ (where the function H has a unique maximum, but the double derivative is zero), and all the remaining points $\Theta \backslash [0.5, \infty)$ are 2-regular. This shows that for $p = 2$ there is no point in Θ with $h \neq 0$ that is critical. In contrast, for $p \geqslant 3$ odd, the set of critical points is a continuous curve in Θ which intersects the line $h = 0$ at a single point, and for $p \geqslant 4$ even, the set of critical points is a continuous curve in Θ which has two arms that intersect the line $h = 0$ in the half-line $[\tilde{\beta}_p, \infty)$ (see Lemma B.3 for the precise statement and Figs. 6 and 7 for an illustration.) Moreover, this curve has exactly one limit point (if $p \geqslant 3$ is odd) and exactly two limit points (if $p \geqslant 4$ is even) outside it, which is (are) precisely the p-special point(s).

Having described the behavior of the function $H_{\beta,h,p}$, we can now state the limiting distribution of $\overline{\sigma}_N$, which depends on whether the point (β, h) is regular, critical, or special.

Theorem 2.1 (Asymptotic distribution of the magnetization). *Fix $p \geqslant 3$ and $(\beta, h) \in \Theta$, and suppose $\boldsymbol{\sigma} \sim \mathbb{P}_{\beta,h,p}$. Then with $H = H_{\beta,p,h}$ as defined in (2.1), the following hold:*

(1) *Suppose (β, h) is p-regular and denote the unique maximizer of H by $m_* = m_*(\beta, h, p)$. Then, as $N \to \infty$,*

$$N^{\frac{1}{2}} (\overline{\sigma}_N - m_*) \xrightarrow{D} N\left(0, -\frac{1}{H''(m_*)}\right). \tag{2.5}$$

(2) *Suppose (β, h) is p-critical and denote the $K \in \{2, 3\}$ maximizers of H by $m_1 := m_1(\beta, h, p) < \ldots < m_K := m_K(\beta, h, p)$. Then, as $N \to \infty$,*

$$\overline{\sigma}_N \xrightarrow{D} \sum_{k=1}^{K} p_k \delta_{m_k}, \tag{2.6}$$

where for each $1 \leqslant k \leqslant K$,[3]

$$p_k := \frac{\left[(m_k^2 - 1)H''(m_k)\right]^{-1/2}}{\sum_{i=1}^{K} \left[(m_i^2 - 1)H''(m_i)\right]^{-1/2}}. \tag{2.7}$$

Moreover, if $A \subseteq [-1, 1]$ is an interval containing m_k in its interior for some $1 \leqslant k \leqslant K$, such that $H(m_k) > H(x)$ for all $x \in \mathrm{cl}(A)\backslash\{m_k\}$, then[4]

$$N^{\frac{1}{2}} (\overline{\sigma}_N - m_k) \Big| \{\overline{\sigma}_N \in A\} \xrightarrow{D} N\left(0, -\frac{1}{H''(m_k)}\right). \tag{2.8}$$

(3) *Suppose (β, h) is p-special and denote the unique maximizer of H by $m_* = m_*(\beta, h, p)$. Then, as $N \to \infty$,*

$$N^{\frac{1}{4}} (\overline{\sigma}_N - m_*) \xrightarrow{D} F,$$

where the density of F with respect to the Lebesgue measure is given by

$$\mathrm{d}F(x) = \frac{2}{\Gamma(\frac{1}{4})} \left(-\frac{H^{(4)}(m_*)}{24}\right)^{\frac{1}{4}} \exp\left(\frac{H^{(4)}(m_*)}{24} x^4\right) \mathrm{d}x, \tag{2.9}$$

with $H^{(4)}$ denoting the fourth derivative of the function H.

[3] Note that all the global maximizers of the function H belong to the open interval $(-1, 1)$, and if (β, p) is p-critical and $m_1, \ldots, m_K$ are the global maximizers of H, for some $K \in \{2, 3\}$, then $H''_{\beta,h,p}(m_i) < 0$, for all $1 \leqslant i \leqslant K$. These statements are proved in Lemmas B.1 and B.2, respectively. This implies that the probabilities $p_1, \ldots, p_K$ in (2.7) are well-defined. Moreover, when (β, h) is p-strongly critical, that is, $H_{\beta,h,p}$ has three global maximizers, the symmetry of the model about $h = 0$ (recall that $p \geqslant 4$ is even and $h = 0$ for a strongly critical point), implies that the three maximizers are $m_1, 0, -m_1$, for some $m_1 = m_1(\beta, h, p) < 0$.

[4] For any set $A \subseteq \mathbb{R}$, $\mathrm{int}(A)$ and $\mathrm{cl}(A)$ denote the topological interior and closure of A, respectively.

The proof of this result is given in Sect. 3. We describe below the key ideas involved in the proof of Theorem 2.1:

- In the p-regular case, the proof has three main steps: The first step is to prove a concentration inequality of $\overline{\sigma}_N$ in an asymptotically vanishing neighborhood m_* (Lemma 3.1). This not only shows that m_* is the typical value of $\overline{\sigma}_N$, but also implies that the partition function Z_N (which is the sum over all $\boldsymbol{\sigma} \in \mathcal{C}_N$ as in (1.2)), can be restricted over those $\boldsymbol{\sigma}$ for which $\overline{\sigma}_N$ lies within this concentration interval around m_*. The second step is to find an accurate asymptotic expansion of Z_N by first approximating this restricted sum by an integral over the concentration interval, and then applying saddle point techniques to get a further approximation to this integral (Lemma 3.2). The third and final step is to use this approximation of Z_N to compute the limit of the moment generating function of $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$, and show that the limit converges to that of the Gaussian distribution appearing in (2.5). Details are given in Sect. 3.1.
- The proof in the p-special case follows the same strategy as the p-regular case, with appropriate modifications to deal with the vanishing second derivative at the maximizer. As before, the first step is to prove the concentration of $\overline{\sigma}_N$ within a vanishing neighborhood of m_* which, in this case, requires a higher-order Taylor expansion, since $H''_{\beta,h,p}(m_*) = 0$ (Lemma 3.3). The second step, as before, is the approximation of the partition function (Lemma 3.4). The proof is completed by calculating the limit of the moment generating function of $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ using this approximation to the partition function. Details are given in Sect. 3.2.
- For the p-critical case, the basic proof strategy remains the same as above. However, to deal with the presence of multiple maximizers, we need to prove a conditional concentration result for the magnetization, that is, $\overline{\sigma}_N$ concentrates at one of the maximizers, given that $\overline{\sigma}_N$ lies in a small neighborhood of that maximizer (Lemma 3.7). Similarly, for the second step, we need to approximate a restricted partition function, where instead of taking a sum over all configurations $\boldsymbol{\sigma} \in \mathcal{C}_N$ as in (1.1), we sum over configurations $\boldsymbol{\sigma} \in \mathcal{C}_N$ such that $\overline{\sigma}_N$ lies in the neighborhood of one of the maximizers (Lemma 3.8). Details are given in Sect. 3.3.

To empirically validate the different results in Theorem 2.1, we fix $p \geqslant 3$, some $(\beta, h) \in \Theta$, and $N = 20{,}000$. Then we generate 10^6 replications from $\mathbb{P}_{\beta,h,p}$ and plot the histograms of the magnetizations. Figure 3a shows the histogram of $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ at the 4-regular point $(\beta, h) = (0.2, 0.1)$ where, as expected from (2.5), we see a limiting normal distribution. Next, Fig. 3b shows the histogram of $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ at the 4-special point $(\beta, h) = (0.3333, 0.40997)$, where a non-normal shape emerges, as predicted by (2.9). Figure 4 shows the histogram of $\overline{\sigma}_N$ at the 4-critical point $(\beta, h) = (0.57, 0.12159)$, where the function $H_{0.57,0.12159,4}$ has two global maximizers (see plot in Fig. 2a). Hence, the histogram of $\overline{\sigma}_N$ has two peaks located at two maximizers (as shown in (2.6)). Finally, in Fig. 5 we show the histogram of $\overline{\sigma}_N$ at a 4-strongly critical point $(\beta, h) = (0.688, 0)$. Here, the histogram has three peaks, since the function $H_{\beta,h,p}$ has three global maximizers (see plot in Fig. 2b). Note that the histograms of $\overline{\sigma}_N$ both in Figs. 4 and 5 look like a Gaussian distribution in a neighborhood of each of the maximizers, as predicted by (2.8) in the theorem above.

2.2. Summarizing the phase diagram. The results above can be compactly summarized and better visualized in a phase diagram, which shows the partition of the parameter space

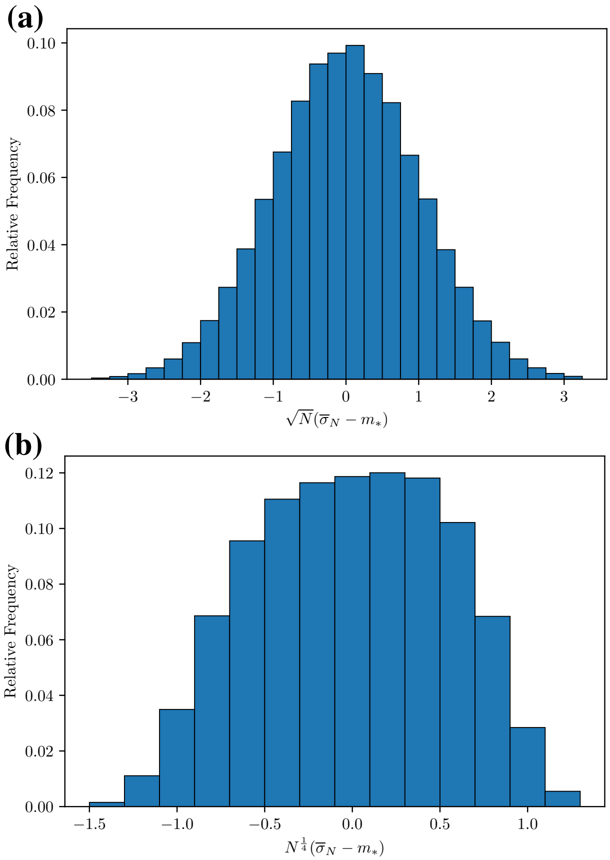

Fig. 3. (**a**) The histogram of $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ at the 4-regular point $(\beta, h) = (0.2, 0.1)$ and (**b**) the histogram of $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ at the 4-special point $(\beta, h) = (0.3333, 0.40997)$

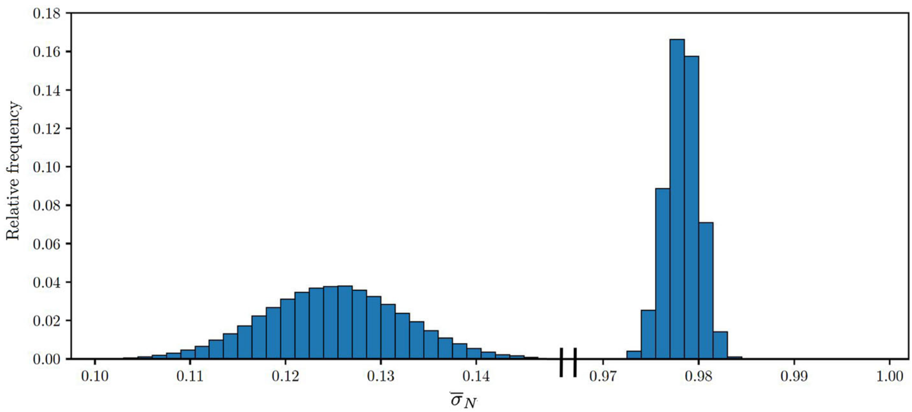

Fig. 4. Histogram of $\overline{\sigma}_N$ at the 4-critical point (0.57, 0.12159), where the function $H_{0.57,0.12159,4}$ has two global maximizers, around which $\overline{\sigma}_N$ concentrates

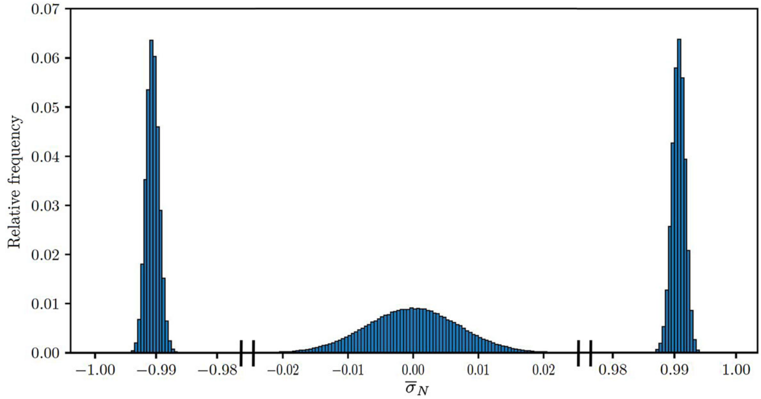

Fig. 5. Histogram of $\overline{\sigma}_N$ at the non 4-strongly critical point (0.6888, 0), where the function $H_{0.6888,0,4}$ has three global maximizers, around which $\overline{\sigma}_N$ concentrates

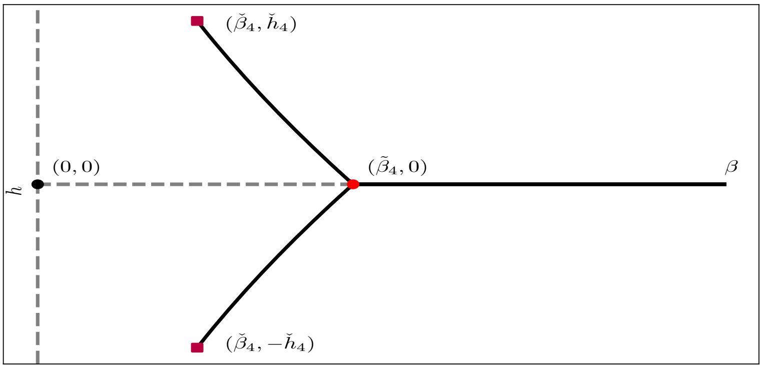

Fig. 6. The phase diagram for $p = 4$: The fluctuations of the magnetization in the different regions of the parameter space $\Theta = [0, \infty) \times \mathbb{R}$ are as follows: The (white) region: These are the p-regular points where H has a unique global maximizer $m_* \in [-1, 1]$ and $H''(m_*) < 0$. Hence, $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ is asymptotically normal by (2.5). The points: These are the p-special points. Here, H has a unique maximizer m_*, but $H''(m_*) = 0$. Hence, $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ converges to a non-Gaussian distribution as in (2.9). The curve: These are p-weakly critical points. Here, H has two global maximizers and, hence, $\overline{\sigma}_N$ is a 2-point mixture with Gaussian fluctuations centered around the maximizers (by (2.6) and (2.8)). The point: This is the p-strongly critical point. Here, H has three global maximizers and, hence, $\overline{\sigma}_N$ is a 3-point mixture with Gaussian fluctuations centered around the maximizers (by (2.6) and (2.8))

described in (2.4). The phase diagrams for $p = 4$ and $p = 5$, obtained by numerical optimization of the function H over a fine grid of parameter values, are shown in Figs. 6 and 7, respectively. The limiting distributions that arise in the different regions of the phase diagram are described in the figure legends.

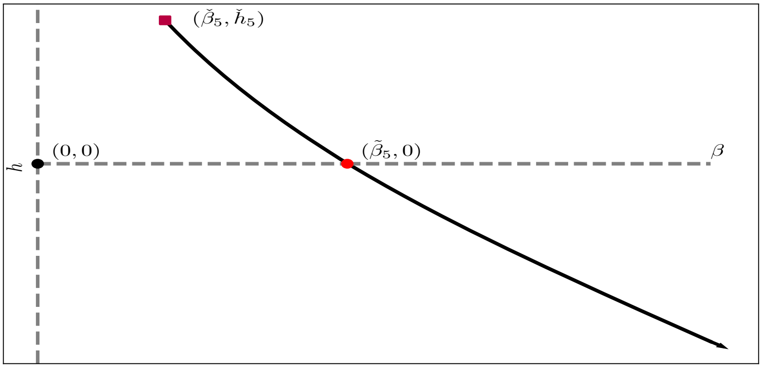

Fig. 7. The phase diagram for $p = 5$: The properties of the magnetization in the different regions of the parameter space $\Theta = [0, \infty) \times \mathbb{R}$ are as follows:The (white) region: These are the p-regular points where H has a unique global maximizer $m_* \in [-1, 1]$ and $H''(m_*) < 0$. Hence, $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ is asymptotically normal by (2.5). The ■ point: This is the only p-special point. Here, H has a unique maximizer m_*, but $H''(m_*) = 0$. Hence, $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ converges to a non-Gaussian distribution as in (2.9). The ▬ curve and the ● point: These are p-weakly critical points. Here, H has two global maximizers. Hence, $\overline{\sigma}_N$ is a 2-point mixture with Gaussian fluctuations centered around the maximizers (by (2.6) and (2.8))

3. Asymptotic Distribution of the Magnetization: Proof of Theorem 2.1

In this section, we will prove Theorem 2.1. To this end, note that the model (1.1) can be written more compactly as

$$\mathbb{P}_{\beta,h,p}(\boldsymbol{\sigma}) = \frac{1}{2^N Z_N(\beta, h, p)} \exp\left\{N\left(\beta\overline{\sigma}_N^p + h\overline{\sigma}_N\right)\right\},$$

where $\overline{\sigma}_N := \frac{1}{N}\sum_{i=1}^N \sigma_i$ is the magnetization. Therefore, the magnetization has the probability mass function,

$$\mathbb{P}_{\beta,h,p}(\overline{\sigma}_N = m) = \frac{1}{2^N Z_N(\beta, h, p)} \binom{N}{\frac{N(1+m)}{2}} e^{N(\beta m^p + hm)},$$
$$\text{for } m \in \left\{-1, -1 + \frac{2}{N}, \ldots, 1 - \frac{2}{N}, 1\right\}.$$

Observe that the probability mass function of $\overline{\sigma}_N$ involves the partition function $Z_N(\beta, h, p)$, which does not have a closed form. Therefore, obtaining limiting properties of $\overline{\sigma}_N$ requires accurate estimation of $Z_N(\beta, h, p)$.

We prove Theorem 2.1 in the p-regular case in Sect. 3.1 below. The proofs in the p-special and the p-critical cases are given in Sects. 3.2 and 3.3, respectively. For technical reasons, in the proofs, we will need to consider slightly perturbed parameter values (β, h_N), for some sequence $h_N \to h$ to be chosen later. Hereafter, we will denote $\mathbb{P}_{\beta,h_N,p}$, $Z_N(\beta, h_N, p)$, and $F_N(\beta, h_N, p)$, by $\bar{\mathbb{P}}$, $\bar{Z}_N$, and $\bar{F}_N$, respectively.

3.1. Proof of Theorem 2.1 when (β, h) is p-regular. Fix a p-regular point $(\beta, h) \in \Theta$ and consider a sequence $h_N \in \mathbb{R}$ (to be specified later) converging to h. It has been shown in Lemma B.4 that the function $H_N(x) := H_{\beta,h_N,p}(x)$ will have a unique global maximizer

$m_*(N)$, for all large N, and $m_*(N) \to m_*$ as $N \to \infty$. Choose this maximizer $m_*(N)$ and define, for $\alpha \in (0, 1)$,

$$A_{N,\alpha} := \left(m_*(N) - N^{-\frac{1}{2}+\alpha}, m_*(N) + N^{-\frac{1}{2}+\alpha}\right). \tag{3.1}$$

The first step in the proof of Theorem 2.1 when (β, h) is p-regular, is to show that under $\bar{\mathbb{P}}$, the magnetization $\overline{\sigma}_N$ concentrates around $m_*(N)$ at rate $N^{-\frac{1}{2}+\alpha}$, for any $\alpha > 0$.

Lemma 3.1. *Suppose $(\beta, h) \in \Theta$ is p-regular. Then for $\alpha \in \left(0, \frac{1}{6}\right]$ and $A_{N,\alpha}$ as defined above in (3.1),*[5]

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A^c_{N,\alpha}\right) = \exp\left\{\frac{1}{3}N^{2\alpha}H''(m_*)\right\} O(N^{\frac{3}{2}}).$$

Proof. Note that the support of the magnetization $\overline{\sigma}_N$ is the set

$$\mathcal{M}_N := \left\{-1, -1 + \frac{2}{N}, \ldots, 1 - \frac{2}{N}, 1\right\}.$$

It follows from [32], Equation (5.4), that for any $m \in \mathcal{M}_N$, the cardinality of the set

$$A_m := \{\boldsymbol{\sigma} \in \mathcal{C}_N : \overline{\sigma}_N = m\}$$

can be bounded by

$$\frac{2^N}{LN^{\frac{1}{2}}} \exp\{-NI(m)\} \leqslant |A_m| \leqslant 2^N \exp\{-NI(m)\} \tag{3.2}$$

for some universal constant L (recall that $I(\cdot)$ is the binary entropy function). Hence, we have from (3.2),

$$\begin{aligned}
\bar{\mathbb{P}}(\overline{\sigma}_N \in A^c_{N,\alpha}) &= \frac{\sum_{m\in\mathcal{M}_N \bigcap A^c_{N,\alpha}} |A_m| \exp\{N(\beta m^p + h_N m)\}}{\sum_{m\in\mathcal{M}_N} |A_m| \exp\{N(\beta m^p + h_N m)\}} \\
&\leqslant \frac{LN^{\frac{1}{2}}(N+1)\sup_{x\in A^c_{N,\alpha}} e^{NH_N(x)}}{\sup_{x\in[-1,1]} e^{NH_N(x)}} \\
&= \exp\left\{N\left(\sup_{x\in A^c_{N,\alpha}} H_N(x) - H_N(m_*(N))\right)\right\} O(N^{\frac{3}{2}}). \qquad (3.3)
\end{aligned}$$

By Lemma B.11, we know that for all large N, $\sup_{x\in A^c_{N,\alpha}} H_N(x)$ is either $H_N(m_*(N) - N^{-\frac{1}{2}+\alpha})$ or $H_N(m_*(N) + N^{-\frac{1}{2}+\alpha})$. Since $H'_N(m_*(N)) = 0$ and the functions $H^{(3)}_N$ are uniformly bounded on any closed interval contained in $(-1, 1)$, Taylor's theorem gives us:

$$\begin{aligned}
H_N\left(m_*(N) \pm N^{-\frac{1}{2}+\alpha}\right) - H_N(m_*(N)) &= \frac{1}{2}N^{-1+2\alpha}H''_N(m_*(N)) + O\left(N^{-\frac{3}{2}+3\alpha}\right) \qquad (3.4) \\
&\leqslant \frac{1}{3}N^{-1+2\alpha}H''(m_*) + O\left(N^{-\frac{3}{2}+3\alpha}\right). \qquad (3.5)
\end{aligned}$$

Note that (3.5) follows from (3.4) since $H''_N(m_*(N)) \to H''(m_*) < 0$. The proof of Lemma 3.1 is now complete, in view of (3.3). □

[5] For any set A, A^c denotes the complement of the set A.

Lemma 3.1 shows that almost all contribution to $\bar{Z}_N$ comes from configurations whose magnetization lies in a vanishing neighborhood of the maximizer $m_*(N)$ of H_N. This enables us to accurately approximate the partition function $\bar{Z}_N$. This involves a Riemann approximation of the sum of the mass function $\mathbb{P}_{\beta,h_N,p}(\boldsymbol{\sigma})$ over all $\boldsymbol{\sigma}$ whose mean lies in a vanishing neighborhood of m_*, followed by a further saddle-point approximation of the resulting integral.

Lemma 3.2. *Suppose $(\beta, h) \in \Theta$ is p-regular. Then for $\alpha > 0$ and N large enough, the partition function can be expanded as,*

$$\bar{Z}_N = \frac{e^{NH_N(m_*(N))}}{\sqrt{(m_*(N)^2 - 1)H_N''(m_*(N))}}\left(1 + O\left(N^{-\frac{1}{2}+\alpha}\right)\right), \tag{3.6}$$

where $m_*(N)$ is the unique maximizer of the function H_N. Moreover, for N large enough, the log-partition function can be expanded as,

$$\bar{F}_N = NH_N(m_*(N)) - \tfrac{1}{2}\log\left[(m_*(N)^2 - 1)H_N''(m_*(N))\right] + O\left(N^{-\frac{1}{2}+\alpha}\right). \tag{3.7}$$

Proof. Without loss of generality, let $\alpha \in \left(0, \frac{1}{6}\right]$ and note that

$$\bar{\mathbb{P}}(\overline{\sigma}_N \in A_{N,\alpha}) = \bar{Z}_N^{-1} \sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}} \binom{N}{N(1+m)/2} \exp\left\{N(\beta m^p + h_N m - \log 2)\right\}. \tag{3.8}$$

By Lemma 3.1, $\bar{\mathbb{P}}(\overline{\sigma}_N \in A_{N,\alpha}) = 1 - O\left(e^{-N^\alpha}\right)$ and hence (3.8) gives us

$$\begin{aligned}\bar{Z}_N &= \left(1 + O\left(e^{-N^\alpha}\right)\right) \sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}} \binom{N}{N(1+m)/2} \exp\left\{N(\beta m^p + h_N m - \log 2)\right\} \\ &= \left(1 + O\left(e^{-N^\alpha}\right)\right) \sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}} \zeta(m) \end{aligned} \tag{3.9}$$

where $\zeta : [-1, 1] \to \mathbb{R}$ is defined as

$$\zeta(x) := \binom{N}{N(1+x)/2} \exp\left\{N(\beta x^p + h_N x - \log 2)\right\}, \tag{3.10}$$

where $\binom{N}{N(1+x)/2}$ is interpreted as a continuous binomial coefficient (refer to "Appendix A.1" for the definition of continuous binomial coefficients). The next step is to approximate the sum in (3.9) by an integral, using Lemma A.2. Note that Lemma A.2 can be applied with $n = \Theta(N^{\frac{1}{2}+\alpha})$ to obtain (using Lemma B.7),

$$\begin{aligned}\left|\int_{A_{N,\alpha}} \zeta(x)\mathrm{d}x - \frac{2}{N}\sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}} \zeta(m)\right| &\leqslant \Theta(N^{-\frac{1}{2}+\alpha})N^{-1} \sup_{x \in A_{N,\alpha}} |\zeta'(x)| \\ &= O\left(N^{-\frac{1}{2}+\alpha} \cdot N^{-1} \cdot N^{\frac{1}{2}+\alpha}\right)\zeta(m_*(N)) \\ &= O\left(N^{-1+2\alpha}\right)\zeta(m_*(N)). \end{aligned} \tag{3.11}$$

It now follows from (3.11), Lemmas A.5, A.3 and B.6, that

$$\begin{aligned}
&\sum_{m\in\mathcal{M}_N\bigcap A_{N,\alpha}} \zeta(m) \\
&= \frac{N}{2}\int_{A_{N,\alpha}} \zeta(x)\mathrm{d}x + O(N^{2\alpha})\zeta(m_*(N)) \\
&= \frac{N^{\frac{1}{2}}}{2}\left(1+O(N^{-1})\right)\int_{A_{N,\alpha}} e^{NH_N(x)}\sqrt{\frac{2}{\pi(1-x^2)}}\mathrm{d}x \\
&\quad + O(N^{2\alpha})\zeta(m_*(N)) \\
&= \frac{N^{\frac{1}{2}}}{2}\sqrt{\frac{2\pi}{N|H_N''(m_*(N))|}}\sqrt{\frac{2}{\pi(1-m_*(N)^2)}}e^{NH_N(m_*(N))} \\
&\quad \left(1+O\left(N^{-\frac{1}{2}+3\alpha}\right)\right) + O(N^{2\alpha})\zeta(m_*(N)) \\
&= \frac{e^{NH_N(m_*(N))}}{\sqrt{(m_*(N)^2-1)H_N''(m_*(N))}}\left(1+O\left(N^{-\frac{1}{2}+3\alpha}\right)\right) \\
&\quad + \sqrt{\frac{2}{\pi N(1-m_*(N)^2)}}e^{NH_N(m_*(N))}\left(1+O(N^{-1})\right)O(N^{2\alpha}) \\
&= \frac{e^{NH_N(m_*(N))}}{\sqrt{(m_*(N)^2-1)H_N''(m_*(N))}}\left(1+O\left(N^{-\frac{1}{2}+3\alpha}\right)\right).
\end{aligned} \tag{3.12}$$

Combining (3.9) and (3.12), we have:

$$\begin{aligned}
\bar{Z}_N &= \left(1+O\left(e^{-N^{\alpha}}\right)\right)\left(1+O\left(N^{-\frac{1}{2}+3\alpha}\right)\right)\frac{e^{NH_N(m_*(N))}}{\sqrt{(m_*(N)^2-1)H_N''(m_*(N))}} \\
&= \left(1+O\left(N^{-\frac{1}{2}+3\alpha}\right)\right)\frac{e^{NH_N(m_*(N))}}{\sqrt{(m_*(N)^2-1)H_N''(m_*(N))}}.
\end{aligned} \tag{3.13}$$

This completes the proof of (3.6). If we take logarithm on all sides in (3.13) and use the fact that $\log(1+O(a_n)) = O(a_n)$ for any sequence $a_n = o(1)$, then we get (3.7), completing the proof. □

Completing the Proof of (2.5): We now have all the necessary ingredients in order to derive the CLT for $\overline{\sigma}_N$ when (β, h) is p-regular. Recall that $m_* = m_*(\beta, h, p)$ is the unique maximizer of H. To complete the proof we will show that the moment generating function of $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ under $\mathbb{P}_{\beta,h,p}$ converges pointwise to the moment generating function of the $N(0, -1/H''(m_*))$ distribution. Towards this, fix $t \in \mathbb{R}$ and note that the moment generating function of $N^{\frac{1}{2}}(\overline{\sigma}_N - m_*)$ at t can be expressed as

$$\mathbb{E}_{\beta,h,p}\left[e^{tN^{\frac{1}{2}}(\overline{\sigma}_N - m_*)}\right] = e^{-tN^{\frac{1}{2}}m_*}\frac{Z_N\left(\beta, h+N^{-\frac{1}{2}}t, p\right)}{Z_N(\beta,h,p)}. \tag{3.14}$$

Using Lemma 3.2 and the fact that $m_*(N) \to m_*$, the right side of (3.14) simplifies to

$$(1+o(1))e^{-tN^{\frac{1}{2}}m_*+N\left\{H_{\beta,h+N^{-\frac{1}{2}}t,p}\left(m_*\left(\beta,h+N^{-\frac{1}{2}}t,p\right)\right)-H_{\beta,h,p}(m_*(\beta,h,p))\right\}}. \quad (3.15)$$

Now, Lemma B.5 and a simple Taylor expansion gives us

$$\begin{aligned} m_*\left(\beta, h+N^{-\frac{1}{2}}t, p\right) - m_*(\beta, h, p) &= N^{-\frac{1}{2}}t\,\frac{\partial}{\partial \underline{h}} m_*(\beta, \underline{h}, p)\Big|_{\underline{h}=h} + O(N^{-1}) \\ &= -\frac{t}{N^{\frac{1}{2}}H''(m_*(\beta, h, p))} + O(N^{-1}). \quad (3.16) \end{aligned}$$

Using (3.16) and a further Taylor expansion, we have

$$\begin{aligned} &N\left\{H_{\beta,h,p}\left(m_*\left(\beta, h+N^{-\frac{1}{2}}t, p\right)\right) - H_{\beta,h,p}\left(m_*\left(\beta,h,p\right)\right)\right\} \\ &\quad = \frac{N}{2}\left\{m_*\left(\beta, h+N^{-\frac{1}{2}}t, p\right) - m_*\left(\beta,h,p\right)\right\}^2 H''\left(m_*\left(\beta,h,p\right)\right) + o(1) \\ &\quad = \frac{t^2}{2H''\left(m_*\left(\beta,h,p\right)\right)} + o(1) \\ &\quad = \frac{t^2}{2H''(m_*)} + o(1). \quad (3.17) \end{aligned}$$

Next, we have by Lemma B.5 and a Taylor expansion,

$$\begin{aligned} tN^{\frac{1}{2}}m_*\left(\beta, h+N^{-\frac{1}{2}}t, p\right) &= tN^{\frac{1}{2}}m_*(\beta,h,p) + t(t+\bar{h})\frac{\partial}{\partial \underline{h}} m_*(\beta,\underline{h},p)\Big|_{\underline{h}=h} + o(1) \\ &= tN^{\frac{1}{2}}m_* - \frac{t^2}{H''(m_*)} + o(1). \quad (3.18) \end{aligned}$$

Adding (3.17) and (3.18), and recalling the definition of the function H from (2.1), we have:

$$\begin{aligned} &N\left\{H_{\beta,h+N^{-\frac{1}{2}}t,p}\left(m_*\left(\beta, h+N^{-\frac{1}{2}}t, p\right)\right) - H_{\beta,h,p}\left(m_*\left(\beta,h,p\right)\right)\right\} \\ &\quad = tN^{\frac{1}{2}}m_* - \frac{t^2}{2H''(m_*)} + o(1). \quad (3.19) \end{aligned}$$

Using (3.19), the expression in (3.15) becomes

$$\exp\left\{-\frac{t^2}{2H''(m_*)}\right\} + o(1). \quad (3.20)$$

The constant in expression (3.20) is easily recognizable as the moment generating function of $N(0, \;-\frac{1}{H''(m_*)})$ evaluated at t. This completes the proof of Theorem 2.1 for p-regular points (β, h). □

3.2. Proof of Theorem 2.1 when (β, h) *is p-special.* When (β, h) is p-special, we consider local perturbations of the parameters $(\beta, h_N) := (\beta, h + \bar{h}N^{-\frac{3}{4}})$. Note that in this case the function $H_{\beta,h,p}$ still has a unique maximizer $m_* = m_*(\beta, h, p)$, but $H''_{\beta,h,p}(m_*) = 0$. The proof strategy here has the same broad roadmap as in the p-regular case, with relevant modifications while taking Taylor expansions, since $H''_{\beta,h,p}(m_*) = 0$. As before, the first step is to prove the concentration of $\overline{\sigma}_N$ within a vanishing neighborhood of m_* (Lemma 3.3). Here, the concentration window turns out to be a little more inflated, that is, its length is of order $N^{-\frac{1}{4}+\alpha}$, for $\alpha > 0$. Next, we approximate the partition function $\bar{Z}_N$, where, since the second derivative of H is zero at the maximizer, we need to consider derivatives up to order four to accurately approximate $\bar{Z}_N$ (Lemma 3.4). The details of the proof are presented below.

Throughout this section, as usual, we will denote $H_{\beta,h,p}$ by H, $H_{\beta,h_N,p}$ by H_N, the unique global maximizer of $H_{\beta,h_N,p}$ (for large N) by $m_*(N)$, $\mathbb{P}_{\beta,h_N,p}$ by $\bar{\mathbb{P}}$, $Z_N(\beta, h_N, p)$ by $\bar{Z}_N$ and $F_N(\beta, h_N, p)$ by $\bar{F}_N$. As outlined above, the first step in the proof of Theorem 2.1 when (β, h) is p-special, is to show the concentration of $\overline{\sigma}_N$ within a vanishing neighborhood of $m_*(N)$. In the p-special case, this is more delicate, because it requires Taylor expansions up to the fourth order term. Here, the concentration window turns out to be a bit more inflated as well, and is given by:

$$\mathcal{A}_{N,\alpha} := (m_*(N) - N^{-\frac{1}{4}+\alpha}, m_*(N) + N^{-\frac{1}{4}+\alpha}). \tag{3.21}$$

Lemma 3.3. *Suppose* $(\beta, h) \in \Theta$ *is p-special. Fix* $\alpha \in \left(0, \frac{1}{20}\right]$ *and let* $\mathcal{A}_{N,\alpha}$ *be as in* (3.21)*. Then,*

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in \mathcal{A}^c_{N,\alpha}\right) = \exp\left\{\frac{1}{24}N^{4\alpha}H^{(4)}(m_*)(1 + o(1))\right\} O(N^{\frac{3}{2}}).$$

Proof. It follows from (3.3) that

$$\begin{aligned}
&\bar{\mathbb{P}}(\overline{\sigma}_N \in \mathcal{A}^c_{N,\alpha}) \\
&\quad = \exp\left\{N\left(\sup_{x \in A^c_{N,\alpha}} H_N(x) - H_N\left(m_*(N)\right)\right)\right\} O(N^{\frac{3}{2}}) \\
&\quad \leqslant \exp\left\{N\left(H_N\left(m_*(N) \pm N^{-\frac{1}{4}+\alpha}\right) - H_N\left(m_*(N)\right)\right)\right\} O(N^{\frac{3}{2}}) \\
&\quad \text{(using } H''_N(m_*(N)) \leqslant 0 \text{ and Lemma } B.11\text{)} \\
&\quad \leqslant \exp\left\{\frac{1}{6}N^{\frac{1}{4}+3\alpha}H^{(3)}_N(m_*(N)) + \frac{1}{24}N^{4\alpha}H^{(4)}_N(m_*(N)) + O\left(N^{-\frac{1}{4}+5\alpha}\right)\right\} O(N^{\frac{3}{2}}).
\end{aligned} \tag{3.22}$$

Now, it follows from Lemma B.10, that $|H^{(3)}_N(m_*(N))| = O(N^{-1/4})$. Hence, $N^{(1/4)+3\alpha}$ $H^{(3)}_N(m_*(N)) + N^{4\alpha}H^{(4)}_N(m_*(N)) = N^{4\alpha}H^{(4)}(m_*)(1 + o(1))$, and Lemma 3.3 follows from (3.22). □

The next step in the proof of Theorem 2.1 when (β, h) is p-special is the approximation of the partition function.

Lemma 3.4. *Suppose $(\beta, h) \in \Theta$ is p-special, and let $h_N = h + N^{-\frac{3}{4}}t$, for $t \in \mathbb{R}$. Then for N large enough, the partition function $\bar{Z}_N$ can be expanded as*

$$\bar{Z}_N = \frac{N^{\frac{1}{4}} e^{N H_N(m_*(N))}}{\sqrt{2\pi(1-m_*(N)^2)}} \int_{-\infty}^{\infty} e^{\eta_{t,p}(y)} \mathrm{d}y \,(1 + o(1)),$$

where $\eta_{t,p}(y) = ay^2 + by^3 + cy^4$, with

$$a := \frac{(6t)^{\frac{2}{3}} \left(H^{(4)}(m_*)\right)^{\frac{1}{3}}}{4}, \; b := -\frac{(6t)^{\frac{1}{3}} \left(H^{(4)}(m_*)\right)^{\frac{2}{3}}}{6}, \; \textit{and } c := \frac{H^{(4)}(m_*)}{24}.$$

Proof. Once again, as in the proof of Lemma 3.2, it follows from Lemma 3.3, that for $\alpha \in \left(0, \frac{1}{20}\right]$,

$$\bar{Z}_N = \left(1 + O\left(e^{-N^{\alpha}}\right)\right) \sum_{m \in \mathcal{M}_N \bigcap \mathcal{A}_{N,\alpha}} \zeta(m), \tag{3.23}$$

where $\zeta : [-1, 1] \to \mathbb{R}$ is defined in (3.10) and $\mathcal{A}_{N,\alpha}$ is defined in (3.21). It also follows from Lemma A.2 and Lemma B.9, exactly as in the proof of Lemma 3.2, that

$$\left| \int_{\mathcal{A}_{N,\alpha}} \zeta(x) \mathrm{d}x - \frac{2}{N} \sum_{m \in \mathcal{M}_N \bigcap \mathcal{A}_{N,\alpha}} \zeta(m) \right| = O\left(N^{-1+4\alpha}\right) \zeta(m_*(N)). \tag{3.24}$$

Hence, we have from (3.24), Lemma B.6, Lemma A.4 and Lemma B.10,

$$\begin{aligned}
\sum_{m \in \mathcal{M}_N \bigcap \mathcal{A}_{N,\alpha}} \zeta(m) &= \frac{N}{2} \int_{\mathcal{A}_{N,\alpha}} \zeta(x) \mathrm{d}x + O(N^{4\alpha}) \zeta(m_*(N)) \\
&= \frac{N^{\frac{1}{2}}}{2} \left(1 + O(N^{-1})\right) \int_{\mathcal{A}_{N,\alpha}} e^{N H_N(x)} \sqrt{\frac{2}{\pi(1-x^2)}} \mathrm{d}x + O(N^{4\alpha}) \zeta(m_*(N)) \\
&= O(N^{4\alpha}) \zeta(m_*(N)) + \frac{N^{\frac{1}{4}}}{\sqrt{2\pi(1-m_*(N)^2)}} e^{N H_N(m_*(N))} \\
&\quad \int_{-N^{\alpha}}^{N^{\alpha}} e^{\eta_{t,p}(y)} \mathrm{d}y \left(1 + O\left(N^{-\frac{1}{4}+5\alpha}\right)\right) \\
&= \frac{N^{\frac{1}{4}} e^{N H_N(m_*(N))}}{\sqrt{2\pi(1-m_*(N)^2)}} \int_{-\infty}^{\infty} e^{\eta_{t,p}(y)} \mathrm{d}y (1 + o(1)) \left(1 + O\left(N^{-\frac{1}{4}+5\alpha}\right)\right) \\
&\quad + \sqrt{\frac{2}{\pi N(1-m_*(N)^2)}} e^{N H_N(m_*(N))} \left(1 + O(N^{-1})\right) O(N^{4\alpha}) \\
&= \frac{N^{\frac{1}{4}} e^{N H_N(m_*(N))}}{\sqrt{2\pi(1-m_*(N)^2)}} \int_{-\infty}^{\infty} e^{\eta_{t,p}(y)} \mathrm{d}y \,(1 + o(1)).
\end{aligned} \tag{3.25}$$

Combining (3.23) and (3.25), we have:

$$\begin{aligned}\bar{Z}_N &= \left(1+O\left(e^{-N^{\alpha}}\right)\right)(1+o(1))\frac{N^{\frac{1}{4}}e^{NH_N(m_*(N))}}{\sqrt{2\pi(1-m_*(N)^2)}}\int_{-\infty}^{\infty}e^{\eta_{t,p}(y)}\mathrm{d}y\\ &= (1+o(1))\frac{N^{\frac{1}{4}}e^{NH_N(m_*(N))}}{\sqrt{2\pi(1-m_*(N)^2)}}\int_{-\infty}^{\infty}e^{\eta_{t,p}(y)}\mathrm{d}y. \end{aligned} \tag{3.26}$$

This completes the proof of Lemma 3.4. □

Completing the Proof of **Theorem 2.1 when** (β, h) **is** p**-special**: As before, we start by computing the limiting moment generating function of

$$N^{\frac{1}{4}}\left(\overline{\sigma}_N - m_*(\beta, h, p)\right),$$

in the following lemma.

Lemma 3.5. *For every* p*-special point* $(\beta, h) \in \Theta$*, if* $\boldsymbol{\sigma} \sim \mathbb{P}_{\beta,h,p}$*, then*

$$\lim_{N\to\infty}\mathbb{E}_{\beta,h,p}\left[e^{tN^{\frac{1}{4}}(\overline{\sigma}_N - m_*(\beta,h,p))}\right] = C_p(t)\exp\left\{-\frac{3}{4}\left(\frac{6t^4}{H^{(4)}(m_*)}\right)^{\frac{1}{3}}\right\}, \tag{3.27}$$

where

$$C_p(t) := \frac{\int_{-\infty}^{\infty}e^{\eta_{t,p}(y)}\mathrm{d}y}{\int_{-\infty}^{\infty}e^{\frac{H^{(4)}(m_*)}{24}y^4}\mathrm{d}y}$$

and $\eta_{t,p}$ *is defined in the statement of Lemma 3.4.*

Proof. Once again, throughout this proof, we will denote $m_*(\beta, h, p)$ by m_*. Fix $t \in \mathbb{R}$ and note that the moment generating function of $N^{\frac{1}{4}}(\overline{\sigma}_N - m_*)$ at t can be expressed as

$$\mathbb{E}_{\beta,h,p}\left[e^{tN^{\frac{1}{4}}(\overline{\sigma}_N - m_*)}\right] = e^{-tN^{\frac{1}{4}}m_*}\frac{Z_N\left(\beta, h+N^{-\frac{3}{4}}t, p\right)}{Z_N(\beta,h,p)}. \tag{3.28}$$

Using Lemma 3.4 and the fact that $m_*(\beta, h+N^{-\frac{3}{4}}t, p) \to m_*$, the right side of (3.28) simplifies to

$$C_p(t)e^{-tN^{\frac{1}{4}}m_* + N\left\{H_{\beta,h+N^{-\frac{3}{4}}t,p}\left(m_*\left(\beta,h+N^{-\frac{3}{4}}t,p\right)\right)-H_{\beta,h,p}(m_*(\beta,h,p))\right\}}(1+o(1)). \tag{3.29}$$

By Lemma B.10, we have:

$$N^{\frac{1}{4}}\left(m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right) - m_*\right) = -R_p(t) + o(1), \tag{3.30}$$

where $R_p(t) := \left(6t/H^{(4)}(m_*)\right)^{\frac{1}{3}}$. By a further Taylor expansion and using (B.10), we have (denoting $H_N = H_{\beta,h_N,p}$),

$$N\left\{H_N\left(m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right)\right) - H\left(m_*\left(\beta,h,p\right)\right)\right\} = T_1 + T_2 + T_3 + T_4, \tag{3.31}$$

where

$$T_1 := \frac{N}{2}\left\{m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right) - m_*(\beta,h,p)\right\}^2 H''_{\beta,h,p}(m_*(\beta,h,p)) = o(1),$$
$$T_2 := \frac{N}{6}\left\{m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right) - m_*(\beta,h,p)\right\}^3 H^{(3)}_{\beta,h,p}(m_*(\beta,h,p)) = o(1),$$
$$T_3 := \frac{N}{24}\left\{m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right) - m_*(\beta,h,p)\right\}^4 H^{(4)}_{\beta,h,p}(m_*(\beta,h,p))$$
$$= \frac{1}{24}R_p(t)^4 H^{(4)}(m_*) + o(1),$$

and $T_4 := O(N\{m_*(\beta, h+N^{-\frac{3}{4}}t, p) - m_*(\beta,h,p)\}^5) = o(1)$.

Now, using both (3.30) and (3.31), we have

$$N\left[H_{\beta,h+N^{-\frac{3}{4}}t,p}\left(m_*\left(\beta, h+N^{-\frac{3}{4}}t, p\right)\right) - H_{\beta,h,p}(m_*(\beta,h,p))\right]$$
$$= tN^{\frac{1}{4}}m_* - tR_p(t) + \eta_{0,p}(-R_p(t)) + o(1).$$

Using the above with (3.28), (3.29), and noting that $\eta_{0,p}(y) = \frac{H^{(4)}(m_*)}{24}y^4$, the result in Lemma 3.5 follows. □

Although (3.27) is not readily recognizable as the moment generating function of any probability distribution, we will show below that it is indeed the moment generating function of the distribution F given by:

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} \propto \exp\left(\frac{H^{(4)}(m_*)}{24}x^4\right). \tag{3.32}$$

Lemma 3.6. *Let F be the distribution defined in (3.32). Then,*

$$\int e^{tx}\mathrm{d}F(x) = C_p(t)\exp\left\{-\frac{3}{4}\left(\frac{6t^4}{H^{(4)}(m_*)}\right)^{\frac{1}{3}}\right\}, \tag{3.33}$$

with notations as in Lemma 3.5.

Proof. Let us denote the right side of (3.33) by $M(t)$. Define

$$\Delta(t,y) := -\frac{3}{4}\left(\frac{6t^4}{H^{(4)}(m_*)}\right)^{\frac{1}{3}} + \eta_{t,p}(y).$$

Note that

$$M(t) = \frac{\int_{-\infty}^{\infty} e^{\Delta(t,y)}\mathrm{d}y}{\int_{-\infty}^{\infty} e^{\frac{H^{(4)}(m_*)}{24}y^4}\mathrm{d}y}. \tag{3.34}$$

Now, recall that $R_p(t) = \left(6t/H^{(4)}(m_*)\right)^{\frac{1}{3}}$. Using the change of variables $u = y - R_p(t)$ and a straightforward algebra, we have

$$\Delta(t,y) = \frac{H^{(4)}(m_*)}{24}u^4 + tu. \tag{3.35}$$

Lemma 3.6 now follows on substituting (3.35) in (3.34). □

It now follows from Lemmas 3.5 and 3.6, that for a p-special point (β, h),

$$N^{\frac{1}{4}}(\overline{\sigma}_N - m_*(\beta, h, p)) \xrightarrow{D} F \tag{3.36}$$

This completes the proof of Theorem 2.1 when (β, h) is p-special.

3.3. Proof of Theorem 2.1 when (β, h) is p-critical. Throughout this section we assume that $(\beta, h) \in \Theta$ is p-critical. This means, by definition and Lemma B.1, that the function $H = H_{\beta,h,p}$ has $K \in \{2, 3\}$ global maximizers, which we denote by $m_1 < \ldots < m_K$. It also follows from Lemma B.4, that for a sequence $h_N \to h$, the function $H_N := H_{\beta,h_N,p}$, for all large N, has local maximizers at $m_1(N), \ldots, m_K(N)$ such that $m_k(N) \to m_k$, as $N \to \infty$, for all $1 \leqslant k \leqslant K$. As before, $\bar{\mathbb{P}}$ and $\bar{Z}_N$ will denote $\mathbb{P}_{\beta,h_N,p}$ and $Z_N(\beta, h_N, p)$, respectively.

In presence of multiple global maximizers, the magnetization $\overline{\sigma}_N$ will concentrate around the set of all global maximizers. In fact, we can prove the following stronger result: Consider an open interval A around a local maximizer m such that m is the unique global maximizer of H over A. Then conditional on the event $\overline{\sigma}_N \in A$ (which is a rare event if m is not a global maximizer), $\overline{\sigma}_N$ concentrates around m. This is the first step in the proof of Theorem 2.1 when (β, h) is p-critical. To state the result precisely, assume that m is a local maximizer of H and let $m(N)$ be local maximizers of H_N converging to m, which exist by Lemma B.4. Define

$$A_{N,\alpha}(m(N)) = \left(m(N) - N^{-\frac{1}{2}+\alpha}, m(N) + N^{-\frac{1}{2}+\alpha}\right). \tag{3.37}$$

The following lemma gives the conditional and, hence, the unconditional, concentration result of $\overline{\sigma}_N$ around local maximizers.[6]

Lemma 3.7. *Suppose $(\beta, h) \in \Theta$ is p-critical. Then for $\alpha \in \left(0, \frac{1}{6}\right]$ fixed and $A_{N,\alpha}(m(N))$ as defined in (3.37),*

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m(N))^c \middle| \overline{\sigma}_N \in A\right) = \exp\left\{\frac{1}{3}N^{2\alpha}H''(m)\right\} O(N^{\frac{3}{2}}), \tag{3.38}$$

for any interval $A \subseteq [-1, 1]$ such that $m \in int(A)$ and $H(m) > H(x)$, for all $x \in$ $\mathrm{cl}(A)\backslash\{m\}$. As a consequence, for $A_{N,\alpha,K} := \bigcup_{k=1}^K A_{N,\alpha}(m_k(N))$,

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A^c_{N,\alpha,K}\right) = \exp\left\{\frac{1}{3}N^{2\alpha}\max_{1\leqslant k\leqslant K} H''(m_k)\right\} O(N^{\frac{3}{2}}). \tag{3.39}$$

Proof. It follows from Lemma B.4, that for all N sufficiently large, $H_N(m(N)) > H_N(x)$ for all $x \in \mathrm{cl}(A)\backslash\{m(N)\}$, whence we can apply Lemma B.11 to conclude that

$$\sup_{x\in A\backslash A_{N,\alpha}(m(N))} H_N(x) = H_N\left(m(N) \pm N^{-\frac{1}{2}+\alpha}\right),$$

[6] The unconditional concentration derived in (3.39) is not required in the proof of Theorem 2.1. Nevertheless, we include it for the sake of completeness.

for all large N such that $A_{N,\alpha}(m(N)) \subset A$, as well. Following the proof of Lemma 3.1, we have for all large N ,

$$\begin{aligned}
&\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m(N))^c \middle| \overline{\sigma}_N \in A\right) \\
&\leqslant \exp\left\{N\left(\sup_{x\in A\setminus A_{N,\alpha}(m(N))} H_N(x) - \sup_{x\in A} H_N(x)\right)\right\} O(N^{\frac{3}{2}}) \\
&= \exp\left\{N\left(H_N\left(m(N) \pm N^{-\frac{1}{2}+\alpha}\right) - H_N\left(m(N)\right)\right)\right\} O(N^{\frac{3}{2}}) \\
&\leqslant \exp\left\{\frac{N}{3}\left(N^{-1+2\alpha}H''(m) + O\left(N^{-\frac{3}{2}+3\alpha}\right)\right)\right\} O(N^{\frac{3}{2}}). \qquad (3.40)
\end{aligned}$$

The result (3.38) now follows from (3.40).

Next, we proceed to prove (3.39). Let $A_1 := [-1, (m_1 + m_2)/2)$, $A_K := [(m_{K-1} + m_K)/2, 1]$ and for $1 < k < K$, $A_k := [(m_{k-1} + m_k)/2, (m_k + m_{k+1})/2)$. Then, $A_1, A_2, \ldots, A_K$ are disjoint intervals uniting to $[-1, 1]$, $m_k \in \text{int}(A_k)$, and $H(m_k) > H(x)$ for all $x \in \text{cl}(A_k)\setminus\{m_k\}$ and all $1 \leqslant k \leqslant K$. Hence, by Lemma 3.7,

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m_k(N))^c \middle| \overline{\sigma}_N \in A_k\right) = \exp\left\{\frac{1}{3}N^{2\alpha}H''(m_k)\right\} O(N^{\frac{3}{2}}) \quad \text{for all } 1 \leqslant k \leqslant K.$$

Since $A_{N,\alpha}(m_k(N)) \subset A_k$ for all $1 \leqslant k \leqslant K$, for all large N, we have $A_{N,\alpha}(m_k(N))^c \bigcap A_k = A^c_{N,\alpha,K} \bigcap A_k$ for all $1 \leqslant k \leqslant K$, for all large N (recall the definition of $A_{N,\alpha,K}$ from the statement of Lemma 3.7). Hence, $\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m_k(N))^c \middle| \overline{\sigma}_N \in A_k\right) = \bar{\mathbb{P}}\left(\overline{\sigma}_N \in A^c_{N,\alpha,K} \middle| \overline{\sigma}_N \in A_k\right)$ for all $1 \leqslant k \leqslant K$, for all large N. Hence, for all large N, we have

$$\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A^c_{N,\alpha,K} \middle| \overline{\sigma}_N \in A_k\right) = \exp\left\{\frac{1}{3}N^{2\alpha}H''(m_k)\right\} O(N^{\frac{3}{2}}) \quad \text{for all } 1 \leqslant k \leqslant K. \tag{3.41}$$

It follows from (3.41) that for all large N,

$$\begin{aligned}
\bar{\mathbb{P}}(\overline{\sigma}_N \in A^c_{N,\alpha,K}) &= \sum_{k=1}^{K} \bar{\mathbb{P}}\left(\overline{\sigma}_N \in A^c_{N,\alpha,K} \middle| \overline{\sigma}_N \in A_k\right) \bar{\mathbb{P}}(\overline{\sigma}_N \in A_k) \\
&\leqslant \exp\left\{\frac{1}{3}N^{2\alpha} \max_{1\leqslant k\leqslant K} H''(m_k)\right\} O(N^{\frac{3}{2}}) \sum_{k=1}^{K} \bar{\mathbb{P}}(\overline{\sigma}_N \in A_k) \\
&= \exp\left\{\frac{1}{3}N^{2\alpha} \max_{1\leqslant k\leqslant K} H''(m_k)\right\} O(N^{\frac{3}{2}}). \qquad (3.42)
\end{aligned}$$

The result in (3.39) now follows from (3.42), completing the proof of Lemma 3.7. □

In order to derive a conditional CLT of $\overline{\sigma}_N$ around the local maximizer m, given that m is in A (where A is as in Lemma 3.7 above), we need precise estimates of the *restricted partition functions* defined as

$$\bar{Z}_N\big|_A := \frac{1}{2^N} \sum_{\boldsymbol{\sigma}\in\mathcal{C}_N:\overline{\sigma}_N\in A} \exp\left\{N(\beta\overline{\sigma}_N^p + h_N\overline{\sigma}_N)\right\}.$$

Note that $\bar{Z}_N\big|_A$ is the partition function of the conditional measure $\bar{\mathbb{P}}\left(\boldsymbol{\sigma} \in \cdot \middle| \overline{\sigma}_N \in A\right)$, in the sense that for any $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_N) \in \mathcal{C}_N$ such that $\bar{\boldsymbol{\tau}} \in A$, we have

$$\bar{\mathbb{P}}\left(\boldsymbol{\sigma} = \boldsymbol{\tau} \middle| \overline{\sigma}_N \in A\right) = \frac{1}{2^N \bar{Z}_N\big|_A} \exp\left\{N(\beta\overline{\sigma}_N^p + h_N\overline{\sigma}_N)\right\}.$$

The following lemma gives an approximation of the restricted and, hence, the unrestricted partition functions. To this end, recall that $m(N)$ is a local maximizer of H_N converging to m.

Lemma 3.8. *Suppose $(\beta, h) \in \Theta$ is p-critical. Then for $\alpha > 0$ and N large enough, the restricted partition function can be expanded as*

$$\bar{Z}_N\big|_A = \frac{e^{NH_N(m(N))}}{\sqrt{(m(N)^2 - 1)H_N''(m(N))}}\left(1 + O\left(N^{-\frac{1}{2}+\alpha}\right)\right), \tag{3.43}$$

where the set A is as in Lemma 3.7. This implies, for every $\alpha > 0$ and N large enough, the (unrestricted) partition function can be expanded as

$$\bar{Z}_N = \sum_{k=1}^{K} \frac{e^{NH_N(m_k(N))}}{\sqrt{(m_k(N)^2 - 1)H_N''(m_k(N))}}\left(1 + O\left(N^{-\frac{1}{2}+\alpha}\right)\right). \tag{3.44}$$

Proof. The arguments below are meant for all sufficiently large N. Without loss of generality, let $\alpha \in \left(0, \frac{1}{6}\right]$ and note that

$$\begin{aligned}
&\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m(N))\middle|\overline{\sigma}_N \in A\right) \\
&= \bar{Z}_N\big|_A^{-1} \sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}(m(N))} \binom{N}{N(1+m)/2} \exp\left\{N(\beta m^p + h_N m - \log 2)\right\}.
\end{aligned}\tag{3.45}$$

By Lemma 3.7, $\bar{\mathbb{P}}\left(\overline{\sigma}_N \in A_{N,\alpha}(m(N))\middle|\overline{\sigma}_N \in A\right) = 1 - O(e^{-N^\alpha})$ and hence (3.45) gives us

$$\begin{aligned}
\bar{Z}_N\big|_A = \left(1 + O(e^{-N^\alpha})\right) &\sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}(m(N))} \\
&\binom{N}{N(1+m)/2} \exp\left\{N(\beta m^p + h_N m - \log 2)\right\}.
\end{aligned}\tag{3.46}$$

Since $m(N)$ is the unique global maximizer of H_N over the interval $A_{N,\alpha}(m(N))$, by mimicking the proof of Lemma 3.2 on the interval $A_{N,\alpha}(m(N))$, it follows that

$$\begin{aligned}
&\sum_{m \in \mathcal{M}_N \bigcap A_{N,\alpha}(m(N))} \binom{N}{N(1+m)/2} \exp\left\{N(\beta m^p + h_N m - \log 2)\right\} \\
&= \frac{e^{NH_N(m(N))}}{\sqrt{(m(N)^2 - 1)H_N''(m(N))}}\left(1 + O\left(N^{-\frac{1}{2}+3\alpha}\right)\right).
\end{aligned}\tag{3.47}$$

The result in (3.43) now follows from (3.46) and (3.47).

For each $1 \leqslant k \leqslant K$, (3.43) immediately gives us

$$\bar{Z}_N\big|_{A_k} = \frac{e^{NH_N(m_k(N))}}{\sqrt{(m_k(N)^2-1)H_N''(m_k(N))}}\left(1+O\left(N^{-\frac{1}{2}+\alpha}\right)\right), \tag{3.48}$$

where the sets $A_1, \ldots, A_K$ are as defined in the proof of (3.39). The result in (3.44) now follows from (3.48) on observing that $\bar{Z}_N = \sum_{k=1}^{K} \bar{Z}_N\big|_{A_k}$. □

Lemmas 3.7 and 3.8 can now be used to complete the proof of Theorem 2.1 (2).

Completing the Proof of Theorem* 2.1 *when* (β, h) *is* p*-critical: For each $\varepsilon > 0$ and $1 \leqslant s \leqslant K$, define $B_{s,\varepsilon} = (m_s - \varepsilon, m_s + \varepsilon)$. Then for all $\varepsilon > 0$ small enough, $H(m_s) > H(x)$, for all $x \in B_{s,\varepsilon}\backslash\{m_s\}$. Now, for each $1 \leqslant s \leqslant K$, we have

$$\mathbb{P}_{\beta,h,p}(\overline{\sigma}_N \in B_{s,\varepsilon}) = \frac{Z_N(\beta,h,p)\big|_{B_{s,\varepsilon}}}{Z_N(\beta,h,p)}. \tag{3.49}$$

By Lemma 3.8 we have

$$Z_N(\beta,h,p)\big|_{B_{s,\varepsilon}} = \frac{e^{N\sup_{x\in[-1,1]} H(x)}}{\sqrt{(m_s^2-1)H''(m_s)}}(1+o(1)) \quad \text{for all } 1 \leqslant s \leqslant K, \tag{3.50}$$

and

$$Z_N(\beta,h,p) = e^{N\sup_{x\in[-1,1]} H(x)} \sum_{s=1}^{K} \frac{1}{\sqrt{(m_s^2-1)H''(m_s)}}(1+o(1)). \tag{3.51}$$

The result in (2.6) now follows from (3.49), (3.50) and (3.51).

Now, we proceed we prove (2.8). A direct calculation reveals that

$$\mathbb{E}_{\beta,h,p}\left[e^{tN^{\frac{1}{2}}(\overline{\sigma}_N - m)}\Big|\overline{\sigma}_N \in A\right] = e^{-tN^{\frac{1}{2}}m}\frac{Z_N(\beta,h+N^{-\frac{1}{2}}t,p)\big|_A}{Z_N(\beta,h,p)\big|_A}. \tag{3.52}$$

Using Lemma 3.8, the right side of (3.52) simplifies to

$$(1+o(1))e^{-tN^{\frac{1}{2}}m + N\left\{H_{\beta,h+N^{-\frac{1}{2}}t,p}\left(m\left(\beta,h+N^{-\frac{1}{2}}t,p\right)\right) - H_{\beta,h,p}(m(\beta,h,p))\right\}},$$

where $m(\beta,h,p)$ and $m(\beta,h+N^{-\frac{1}{2}}t,p)$ are the local maximizers of the functions $H_{\beta,h,p}$ and $H_{\beta,h+N^{-\frac{1}{2}}t,p}$ respectively, converging to m. We can mimic the proof of Theorem 2.1 (1) or (3) verbatim from this point onward, to conclude that as $N \to \infty$,

$$\mathbb{E}_{\beta,h,p}\left[e^{tN^{\frac{1}{2}}(\overline{\sigma}_N - m)}\Big|\overline{\sigma}_N \in A\right] \to \exp\left\{-\frac{t}{H''(m)} - \frac{t^2}{2H''(m)}\right\}. \tag{3.53}$$

The result in (2.8) now follows from (3.53). □

4. Discussion and Future Directions

In this paper we have derived the limiting distribution of the magnetization in the p-spin Curie–Weiss model (1.1) at all points in the parameter space. One natural way to generalize the model in (1.1) is to change the base measure from the Rademacher distribution (the uniform distribution on $\{-1, 1\}$) to a general probability measure μ supported on $[-1, 1]$. This gives rise to the following probability distribution on $[-1, 1]^N$:

$$\mathrm{d}\mathbb{P}_{\beta,h,p,\mu}(\boldsymbol{\sigma}) = \frac{\exp\left\{N\left(\beta\overline{\sigma}_N^p + h\overline{\sigma}_N\right)\right\}\prod_{i=1}^N \mathrm{d}\mu(\sigma_i)}{Z_N(\beta, h, p, \mu)}, \tag{4.1}$$

for $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_N) \in [-1, 1]^N$. Here, the normalizing constant is given by

$$Z_N(\beta, h, p, \mu) = \int_{[-1,1]^N} \exp\left\{N\left(\beta\overline{\sigma}_N^p + h\overline{\sigma}_N\right)\right\}\prod_{i=1}^N \mathrm{d}\mu(\sigma_i).$$

Clearly, (4.1) reduces to the model in (1.1) for $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ (the Rademacher distribution).

For the 2-spin case, fluctuations of the magnetization have been studied for general base measures [15,17]. In this direction, we expect results analogous to those obtained in Theorem 2.1 to hold for the p-spin model, for $p \geqslant 3$, with general base measures as well. Towards this, by an application of Cramér's theorem and Varadhan's lemma [12] we have,

$$\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta, h, p, \mu) = \sup_{x\in[-1,1]} H_{\beta,h,p,\mu}(x), \tag{4.2}$$

where $H_{\beta,h,p,\mu}(x) := \beta x^p + hx - I_\mu(x)$, with $I_\mu(x) := \sup_{\lambda\in\mathbb{R}}\{\lambda x - \phi_\mu(\lambda)\}$ and $\phi_\mu(\lambda) := \log \mathbb{E}_{X\sim\mu}[e^{\lambda X}]$. Note that $I_\mu(\cdot)$, which is the Legendre transform of the cumulant generating function $\phi_\mu(\cdot)$, is the large deviation rate function of the sample mean for the measure μ. When μ is the Rademacher distribution, $I_\mu(x) = I(x) = \frac{1}{2}\{(1 + x)\log(1 + x) + (1 - x)\log(1 - x)\}$ is the binary entropy function and we get back the function $H_{\beta,h,p}$ as defined in (2.1). The representation of the partition function in (4.2) suggests that the magnetization $\overline{\sigma}_N$, for $\boldsymbol{\sigma} \sim \mathbb{P}_{\beta,h,p,\mu}$ as in (4.1), concentrates around the global maximizers of the function $H_{\beta,h,p,\mu}$. Moreover, as in Theorem 2.1, we expect $\overline{\sigma}_N$ to have limiting distributions centered around the global maximizers (properly conditioned in case of multiple maximizers), where the order of the fluctuations and the nature of the asymptotic distribution will depend on the number of vanishing derivatives of the function $H_{\beta,h,p,\mu}$ at a particular maximizer. To establish this formally one would need precise estimates on the density of μ^{*n}, the n-fold convolution of the base measure μ. While such estimates are readily available for the Rademacher distribution, for general base measures this is more involved. Towards this, large deviation local limit type estimates for sums of i.i.d. random variables [5] can be useful. Computing the global maximizers of the function $H_{\beta,h,p,\mu}$, for a given measure μ, appears to be a rather delicate problem as well. Already in the Rademacher case, as summarized in Figs. 6 and 7 , many new phases emerge as one moves from the 2-spin model to the p-spin model. Understanding the landscape of the function $H_{\beta,h,p,\mu}$ for other natural base measures μ is an interesting problem for future research.

Acknowledgement. The authors thank the anonymous referee for many thoughtful comments which greatly improved the quality and the presentation of the paper. Bhaswar B. Bhattacharya was supported by NSF CAREER Grant DMS-2046393 and a Sloan Research Fellowship.

Appendix A. Properties of Special Functions and Approximation Lemmas

A.1. Special functions and their properties. In this section, we state few important properties of some special mathematical functions which arise in our analysis.

Definition 2. The gamma function $\Gamma : (0, \infty) \mapsto \mathbb{R}$ is defined as:

$$\Gamma(x) := \int_0^\infty u^{x-1} e^{-u}\, du.$$

Definition 3. The digamma function $\Gamma : (0, \infty) \mapsto \mathbb{R}$ is defined as:

$$\psi(x) := \frac{\mathrm{d}}{\mathrm{d}x} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

The following standard expansion of the digamma function will be very helpful in our analysis: As $x \to \infty$,

$$\psi(1 + x) = \log x + \frac{1}{2x} + O(x^{-2}). \tag{A.1}$$

Definition 4. For real numbers $x \geqslant y > 0$, the binomial coefficient x choose y is defined as

$$\binom{x}{y} := \frac{\Gamma(x + 1)}{\Gamma(y + 1)\Gamma(x - y + 1)}.$$

Lemma A.1. *Fix $u > 0$. Then, for every $x \in (0, u)$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}x}\binom{u}{x} = \binom{u}{x} \left[\psi(u - x + 1) - \psi(x + 1)\right].$$

Proof. Let $\iota(x) = \binom{u}{x}$. Then, $\log \iota(x) = \log \Gamma(u + 1) - \log \Gamma(x + 1) - \log \Gamma(u - x + 1)$ and hence,

$$\frac{\iota'(x)}{\iota(x)} = \frac{\mathrm{d}}{\mathrm{d}x} \log \iota(x) = -\psi(x + 1) + \psi(u - x + 1). \tag{A.2}$$

Lemma A.1 now follows from (A.2). □

A.2. Mathematical approximations. In this section, we give three different types of standard mathematical approximations, which play crucial roles in our analysis.

Lemma A.2 (Riemann Approximation). *Let $f : [a, b] \to \mathbb{R}$ be a differentiable function, and let $a = x_0 < x_1 < \ldots < x_n = b$. Let $x_s^* \in [x_{s-1}, x_s]$ for each $1 \leqslant k \leqslant n$. Then, we have:*

$$\left| \int_a^b f - \sum_{k=1}^n (x_s - x_{s-1}) f(x_s^*) \right| \leqslant \frac{1}{2}(b-a) \max_{1 \leqslant k \leqslant n} (x_s - x_{s-1}) \sup_{x \in [a,b]} |f'(x)|.$$

Proof. Lemma A.2 follows from the following string of inequalities:

$$\begin{aligned}
\left| \int_a^b f - \sum_{s=1}^n (x_s - x_{s-1}) f(x_s^*) \right| &= \left| \sum_{s=1}^n \int_{x_{s-1}}^{x_s} (f(x) - f(x_s^*)) \mathrm{d}x \right| \\
&\leqslant \sum_{s=1}^n \int_{x_{s-1}}^{x_s} \left| f(x) - f(x_s^*) \right| \mathrm{d}x \qquad \text{(A.3)} \\
&\leqslant \sup_{x \in [a,b]} |f'(x)| \sum_{s=1}^n \int_{x_{s-1}}^{x_s} \left| x - x_s^* \right| \mathrm{d}x \qquad \text{(A.4)} \\
&= \frac{1}{2} \sup_{x \in [a,b]} |f'(x)| \sum_{s=1}^n \left[(x_s^* - x_{s-1})^2 + (x_s - x_s^*)^2 \right] \\
&\leqslant \frac{1}{2} \sup_{x \in [a,b]} |f'(x)| \sum_{s=1}^n (x_s - x_{s-1})^2 \\
&\leqslant \frac{1}{2}(b-a) \max_{1 \leqslant s \leqslant n} (x_s - x_{s-1}) \sup_{x \in [a,b]} |f'(x)|.
\end{aligned}$$

Note that, in going from (A.3) to (A.4), we used the mean value theorem. □

The following lemma gives a Laplace-type approximation of an integral over a shrinking interval. For the classical Laplace approximation, which approximates integrals over fixed intervals, refer to [13,35]. Even though the proof of Lemma A.3 below is exactly similar to that of the classical Laplace approximation, we provide the proof here for the sake of completeness. To this end, for positive sequences $\{a_n\}_{n \geqslant 1}$ and $\{b_n\}_{n \geqslant 1}$, $a_n = O_\square(b_n)$ denotes $a_n \leqslant C_1(\square) b_n$ and $a_n = \Omega_\square(b_n)$ denotes $a_n \geqslant C_2(\square) b_n$, for all n large enough and positive constants $C_1(\square)$, $C_2(\square)$, which may depend on the subscripted parameters.

Lemma A.3 (Laplace-Type Approximation-I) . *Let $a < b$ be fixed real numbers, $g : [a, b] \mapsto \mathbb{R}$ be a differentiable function on (a, b), and $h_n : [a, b] \mapsto \mathbb{R}$ be a sequence of thrice differentiable functions on (a, b). Suppose that $\{x_n\}$ is a sequence in (a, b) that is bounded away from both a and b, satisfying $h_n'(x_n) = 0$ and $h_n''(x_n) < 0$ for all n. Suppose further, that for every $a < u < v < b$, $\sup_{x \in [u,v]} |g'(x)| = O_{u,v}(1)$, $\sup_{n \geqslant 1} \sup_{x \in [u,v]} |h_n^{(3)}(x)| = O_{u,v}(1)$ and $\inf_{x \in [u,v]} |g(x)| = \Omega_{u,v}(1)$. Also, suppose that $\inf_{n \geqslant 1} |h_n''(x_n)| > 0$. Then, for all $\alpha \in \left(0, \frac{1}{6}\right)$, we have as $n \to \infty$,*

$$\int_{x_n - n^{-\frac{1}{2}+\alpha}}^{x_n + n^{-\frac{1}{2}+\alpha}} g(x) e^{n h_n(x)} \mathrm{d}x = \sqrt{\frac{2\pi}{n \left| h_n''(x_n) \right|}} g(x_n) e^{n h_n(x_n)} \left(1 + O\left(n^{-\frac{1}{2}+3\alpha} \right) \right).$$

Proof. If we make the change of variables $y = \sqrt{n}(x - x_n)$, we have

$$\int_{x_n - n^{-\frac{1}{2}+\alpha}}^{x_n + n^{-\frac{1}{2}+\alpha}} g(x) e^{n h_n(x)} \mathrm{d}x = n^{-\frac{1}{2}} \int_{-n^\alpha}^{n^\alpha} g(y n^{-\frac{1}{2}} + x_n) e^{n h_n\left(y n^{-\frac{1}{2}} + x_n\right)} \mathrm{d}y. \quad \text{(A.5)}$$

By a Taylor expansion, we have for any sequence $y \in [-n^\alpha, n^\alpha]$,

$$e^{n h_n\left(y n^{-\frac{1}{2}} + x_n\right)} = \left(1 + O\left(n^{3\alpha - \frac{1}{2}}\right)\right) e^{n h_n(x_n) + \frac{y^2}{2} h_n''(x_n)}$$
$$\text{and } g(y n^{-\frac{1}{2}} + x_n) = \left(1 + O\left(n^{\alpha - \frac{1}{2}}\right)\right) g(x_n). \quad \text{(A.6)}$$

Using (A.6), the right side of (A.5) becomes

$$\begin{aligned}
& n^{-\frac{1}{2}} \left(1 + O\left(n^{3\alpha - \frac{1}{2}}\right)\right) g(x_n) e^{n h_n(x_n)} \int_{-n^\alpha}^{n^\alpha} e^{\frac{y^2}{2} h_n''(x_n)} \mathrm{d}y \\
&= \left(1 + O\left(n^{3\alpha - \frac{1}{2}}\right)\right) \sqrt{\frac{2\pi}{n \left|h_n''(x_n)\right|}} g(x_n) e^{n h_n(x_n)} \mathbb{P}\left(\left|N\left(0, \frac{1}{|h_n''(x_n)|}\right)\right| \leqslant n^\alpha\right) \\
&= \left(1 + O\left(n^{3\alpha - \frac{1}{2}}\right)\right) \sqrt{\frac{2\pi}{n \left|h_n''(x_n)\right|}} g(x_n) e^{n h_n(x_n)} \left(1 - O\left(e^{-n^\alpha}\right)\right) \\
&= \left(1 + O\left(n^{3\alpha - \frac{1}{2}}\right)\right) \sqrt{\frac{2\pi}{n \left|h_n''(x_n)\right|}} g(x_n) e^{n h_n(x_n)}.
\end{aligned}$$

The proof of Lemma A.3 is now complete. □

Lemma A.4 (Laplace-Type Approximation-II). *Let $a < b$ be fixed real numbers, $g : [a, b] \mapsto \mathbb{R}$ be a differentiable function on (a, b), and $h_n : [a, b] \mapsto \mathbb{R}$ be a sequence of 5-times differentiable functions on (a, b). Suppose that $\{x_n\}$ is a sequence in (a, b) that is bounded away from both a and b, satisfying $h_n'(x_n) = 0$ for all $n \geqslant 1$. Also, assume that $n^{\frac{1}{2}} h_n''(x_n) = C_1 + O(n^{-\frac{1}{4}})$, $n^{\frac{1}{4}} h_n^{(3)}(x_n) = C_2 + O(n^{-\frac{1}{4}})$, and $h_n^{(4)}(x_n) = C_3 + O(n^{-\frac{1}{4}})$, where C_1, C_2 and C_3 are real constants. Suppose further, that for every $a < u < v < b$, $\sup_{x \in [u,v]} |g'(x)| = O_{u,v}(1)$, $\sup_{n \geqslant 1} \sup_{x \in [u,v]} |h_n^{(5)}(x)| = O_{u,v}(1)$ and $\inf_{x \in [u,v]} |g(x)| = \Omega_{u,v}(1)$. Then, for all $\alpha \in \left(0, \frac{1}{20}\right)$, as $n \to \infty$,*

$$\int_{x_n - n^{-\frac{1}{4}+\alpha}}^{x_n + n^{-\frac{1}{4}+\alpha}} g(x) e^{n h_n(x)} \mathrm{d}x = n^{-\frac{1}{4}} g(x_n) e^{n h_n(x_n)} \int_{-n^\alpha}^{n^\alpha} e^{\frac{y^2}{2} C_1 + \frac{y^3}{6} C_2 + \frac{y^4}{24} C_3} \mathrm{d}y \left(1 + O\left(n^{5\alpha - \frac{1}{4}}\right)\right).$$

Proof. To begin with, by a change of variables $y = n^{\frac{1}{4}}(x - x_n)$, we have

$$\int_{x_n - n^{-\frac{1}{4}+\alpha}}^{x_n + n^{-\frac{1}{4}+\alpha}} g(x) e^{n h_n(x)} \mathrm{d}x = n^{-\frac{1}{4}} \int_{-n^\alpha}^{n^\alpha} g(y n^{-\frac{1}{4}} + x_n) e^{n h_n\left(y n^{-\frac{1}{4}} + x_n\right)} \mathrm{d}y \quad \text{(A.7)}$$

Now, by a Taylor expansion of $nh_n\left(yn^{-\frac{1}{4}}+x_n\right)$ around x_n, we have for any sequence $y \in [n^{-\alpha}, n^{\alpha}]$,

$$\begin{aligned} nh_n\left(yn^{-\frac{1}{4}}+x_n\right) &= nh_n(x_n) + \frac{n^{\frac{1}{2}}y^2}{2}h_n''(x_n) \\ &\quad + \frac{n^{\frac{1}{4}}y^3}{6}h_n^{(3)}(x_n) + \frac{y^4}{24}h_n^{(4)}(x_n) + O\left(n^{-\frac{1}{4}}y^5\right) \\ &= nh_n(x_n) + \frac{y^2}{2}C_1 + \frac{y^3}{6}C_2 + \frac{y^4}{24}C_3 + O\left(n^{5\alpha-\frac{1}{4}}\right). \end{aligned} \tag{A.8}$$

It follows from (A.8), that

$$e^{nh_n\left(yn^{-\frac{1}{4}}+x_n\right)} = \left(1+O\left(n^{5\alpha-\frac{1}{4}}\right)\right)e^{nh_n(x_n)+\frac{y^2}{2}C_1+\frac{y^3}{6}C_2+\frac{y^4}{24}C_3}. \tag{A.9}$$

Similarly, for any sequence $y \in [-n^{\alpha}, n^{\alpha}]$, we have

$$g(yn^{-\frac{1}{4}}+x_n) = \left(1+O\left(n^{\alpha-\frac{1}{4}}\right)\right)g(x_n). \tag{A.10}$$

Using (A.9) and (A.10), the right side of (A.7) becomes

$$n^{-\frac{1}{4}}g(x_n)e^{nh_n(x_n)}\int_{-n^{\alpha}}^{n^{\alpha}} e^{\frac{y^2}{2}C_1+\frac{y^3}{6}C_2+\frac{y^4}{24}C_3}\,\mathrm{d}y\left(1+O\left(n^{5\alpha-\frac{1}{4}}\right)\right).$$

The proof of Lemma A.4 is now complete. □

Lemma A.5 (Stirling's Approximation of the Binomial Coefficient). *Suppose that $x = x_N$ is a sequence in $(-1, 1)$ that is bounded away from both 1 and -1. Then, as $N \to \infty$,*

$$\binom{N}{N(1+x)/2} = 2^N\sqrt{\frac{2}{\pi N(1-x^2)}}\exp\left(-NI(x)\right)\left(1+O(N^{-1})\right).$$

Proof. First, note that by the usual Stirling approximation for the gamma function, we have the following as all of u, v and $u-v \to \infty$,

$$\begin{aligned} \binom{u}{v} &= \frac{\sqrt{2\pi u}\left(\frac{u}{e}\right)^u\left(1+O\left(\frac{1}{u}\right)\right)}{\sqrt{2\pi v}\left(\frac{v}{e}\right)^v\left(1+O\left(\frac{1}{v}\right)\right)\sqrt{2\pi(u-v)}\left(\frac{u-v}{e}\right)^{(u-v)}\left(1+O\left(\frac{1}{u-v}\right)\right)} \\ &= \sqrt{\frac{u}{2\pi v(u-v)}}\cdot\frac{u^u}{v^v(u-v)^{u-v}}\left(1+O\left(\frac{1}{u}\right)+O\left(\frac{1}{v}\right)+O\left(\frac{1}{u-v}\right)\right). \end{aligned}$$

Substituting $u = N$ and $v = N(1+x)/2$ (the hypothesis of the lemma indeed implies that u, v and $u - v \to \infty$), we have

$$\binom{N}{N(1+x)/2} = \sqrt{\frac{N}{2\pi \frac{N(1+x)}{2} \cdot \frac{N(1-x)}{2}}} \cdot \frac{N^N}{\left(\frac{N(1+x)}{2}\right)^{N(1+x)/2} \left(\frac{N(1-x)}{2}\right)^{N(1-x)/2}} \left(1 + O(N^{-1})\right)$$
$$= 2^N \sqrt{\frac{2}{\pi N(1-x^2)}} \exp\left(-\frac{N(1+x)}{2}\log(1+x) - \frac{N(1-x)}{2}\log(1-x)\right)\left(1 + O(N^{-1})\right)$$
$$= 2^N \sqrt{\frac{2}{\pi N(1-x^2)}} \exp\left(-NI(x)\right)\left(1 + O(N^{-1})\right).$$

This completes the proof of Lemma A.5. □

Appendix B. Properties of the Function H and Other Technical Lemmas

This section is devoted to proving several technical lemmas that are used throughout the proofs of our main results. In "Appendix B.1", we will prove several important properties of the function H. In "Appendix B.2" we collect the proofs of some other technical lemmas.

B.1. Properties of the function H. We start by showing that a p-strongly critical point arises if and only if $p \geqslant 4$ is even, and in that case, the only such point is $(\tilde{\beta}_p, 0)$ (recall (2.2)).

Lemma B.1 (Basic properties of the function H). *The function $H_{\beta,h,p}$ has the following properties.*

(1) $\sup_{x\in[-1,1]} H_{\beta,h,p}(x) \geqslant 0$ *and equality holds if and only if* $(\beta, h) \in [0, \tilde{\beta}_p] \times \{0\}$.
(2) *Every local maximizer of $H_{\beta,h,p}$ lies in* $(-1, 1)$.
(3) *$H_{\beta,h,p}$ can have at most two local maximizers for $p = 3$ and at most three local maximizers for $p \geqslant 4$. Further, it has three global maximizers if and only if $p \geqslant 4$ is even, $h = 0$ and $\beta = \tilde{\beta}_p$.*

Proof of (1). First note that $\sup_{x\in[-1,1]} H_{\beta,h,p}(x) \geqslant H_{\beta,h,p}(0) = 0$. Now, it follows from first principles, that $\lim_{\varepsilon\to 0} H_{\beta,h,p}(\varepsilon)/\varepsilon = H'_{\beta,h,p}(0) = h$. If $h > 0$, then there exists $0 < \varepsilon < 1$ such that $H_{\beta,h,p}(\varepsilon)/\varepsilon > h/2$, and if $h < 0$, then there exists $-1 < \varepsilon < 0$ such that $H_{\beta,h,p}(\varepsilon)/\varepsilon < h/2$. In either case, $\sup_{x\in[-1,1]} H_{\beta,h,p}(x) \geqslant H_{\beta,h,p}(\varepsilon) > \varepsilon h/2 > 0$. Therefore, equality in (1) implies that $h = 0$, and hence, by the definition in (2.2), we must have $\beta \leqslant \tilde{\beta}_p$. This proves the "only if" direction. For the "if" direction, suppose that $(\beta, h) \in [0, \tilde{\beta}_p] \times \{0\}$. Consider the case $\beta < \tilde{\beta}_p$ first, so that by the definition in (2.2), there exists $\beta' > \beta$ such that $\sup_{x\in[-1,1]} H_{\beta',0,p}(x) = 0$. Equality in (1) now follows from:

$$0 \leqslant \sup_{x\in[-1,1]} H_{\beta,0,p}(x) = \sup_{x\in[-1,1]} H_{\beta,0,p}(|x|)$$
$$\leqslant \sup_{x\in[-1,1]} H_{\beta',0,p}(|x|) = \sup_{x\in[-1,1]} H_{\beta',0,p}(x) = 0.$$

Finally, let $\beta = \tilde{\beta}_p$, and suppose towards a contradiction, that $H_{\beta,0,p}(x) > 0$ for some $x \in [-1, 1]$. Then, $H_{\beta,0,p}(|x|) \geqslant H_{\beta,0,p}(x) > 0$, and hence, there exists $\beta' < \beta$ such that

$$H_{\beta',0,p}(|x|) = H_{\beta,0,p}(|x|) + (\beta' - \beta)|x|^p > 0.$$

This contradicts our previous finding that $\sup_{x\in[-1,1]} H_{\underline{\beta},0,p}(x) = 0$ for all $\beta < \tilde{\beta}_p$. The proof of (1) is now complete. □

Proof of (2). Note that $\lim_{x\to-1^+} H'_{\beta,h,p}(x) = +\infty$ and $\lim_{x\to1^-} H'_{\beta,h,p}(x) = -\infty$. Hence, there exists $\varepsilon > 0$, such that $H_{\beta,h,p}$ is strictly increasing on $[-1, -1+\varepsilon]$ and strictly decreasing on $[1-\varepsilon, 1]$, showing that none of -1 and 1 can be a local maximizer of $H_{\beta,h,p}$. □

Proof of (3). Define

$$N_{\beta,h,p}(x) := (1 - x^2)H''_{\beta,h,p}(x) = \beta p(p - 1)x^{p-2}(1 - x^2) - 1,$$

for $x \in (-1, 1)$. Note that on $(-1, 1)$, $N'_{\beta,h,p}(x) = \beta p(p-1)x^{p-3}(p - 2 - px^2)$ has exactly two roots $\pm\sqrt{1 - 2/p}$, for $p = 3$, and an additional root 0 for $p \geqslant 4$. Define:

$$K_p := 2\mathbf{1}\{p = 3\} + 3\mathbf{1}\{p \geqslant 4\}.$$

Then, by Rolle's theorem, $N_{\beta,h,p}$, and hence, $H''_{\beta,h,p}$ can have at most $K_p + 1$ roots on $(-1, 1)$. This shows that $H'_{\beta,h,p}$ can have at most $K_p + 2$ roots on $(-1, 1)$, which by part (2), include all the local maximizers of $H_{\beta,h,p}$. We now claim that for any two local maximizers $a < b$ of $H_{\beta,h,p}$, there exists a root of $H'_{\beta,h,p}$ in (a, b). To see this, note that since a and b are local maximizers of $H_{\beta,h,p}$, by the mean value theorem, there must exist $a_1 < b_1 \in (a, b)$ such that $H'_{\beta,h,p}(a_1) \leqslant 0$ and $H'_{\beta,h,p}(b_1) \geqslant 0$. Now, by the intermediate value theorem applied on the continuous function $H'_{\beta,h,p}$, we conclude that there is a $\zeta \in (a_1, b_1)$ such that $H'_{\beta,h,p}(\zeta) = 0$. Hence, if there are ℓ local maximizers of $H_{\beta,h,p}$ on $(-1, 1)$, then there are at least $2\ell - 1$ roots of $H'_{\beta,h,p}$ on $(-1, 1)$. Thus,

$$2\ell - 1 \leqslant K_p + 2, \quad \text{i.e.} \quad \ell \leqslant (K_p + 3)/2,$$

which proves the first part of (3).

To prove the second part of (3), first suppose that $H_{\beta,h,p}$ has three global maximizers. By the first part, p must be at least 4. We will now show that p is even, by contradiction. If p is odd, then $H''_{\beta,h,p}(x) < 0$ for all $x \leqslant 0$, and hence, by Rolle's theorem, there can be at most one non-positive root of $H'_{\beta,h,p}$. Now, if $H'_{\beta,h,p}$ has at least four positive roots, then by repeated application of Rolle's theorem, $N'_{\beta,h,p}$ has at least two positive roots. This is a contradiction, since $\sqrt{1 - 2/p}$ is the only positive root of $N'_{\beta,h,p}$. Hence, $H'_{\beta,h,p}$ can have at most three positive roots. Thus, $H'_{\beta,h,p}$ can have at most four roots, and hence, $H_{\beta,h,p}$ can have at most two local maximizers, a contradiction. Hence, p must be even.

Next, we show that h must be 0. If $h > 0$, then $H_{\beta,h,p}(x) < H_{\beta,h,p}(-x)$ for all $x < 0$, and hence, all the three global maximizers of $H_{\beta,h,p}$ must be positive. Thus, $H'_{\beta,h,p}$ has at least 5 positive roots, which implies that $N'_{\beta,h,p}$ has at least three positive roots, a contradiction. Similarly, if $h < 0$, then all the three global maximizers of $H_{\beta,h,p}$ must be negative, and thus, $H'_{\beta,h,p}$ has at least 5 negative roots, which implies that $N'_{\beta,h,p}$ has at least three negative roots, once again a contradiction. This shows that $h = 0$.

Finally, we show that $\beta = \tilde{\beta}_p$. If $\beta > \tilde{\beta}_p$, then by the definition in (2.2), 0 is not a global maximizer of $H_{\beta,h,p}$ and hence, $H_{\beta,h,p}$ being an even function, must have an even number of global maximizers, a contradiction. Therefore, it suffices to assume that $\beta < \tilde{\beta}_p$. We will show that 0 is the only global maximizer of $H_{\beta,h,p}$, which is enough to complete the proof of the *only if* implication. Towards this, suppose that there is a non-zero global maximizer x^* of $H_{\beta,h,p}$. Since $\beta < \tilde{\beta}_p$, we must have $H_{\beta,h,p}(x^*) = 0$, and hence, for every $\beta' \in (\beta, \tilde{\beta}_p)$, we must have $H_{\beta',h,p}(x^*) > 0$, a contradiction to the definition in (2.2). This completes the proof of the *only if* implication.

For the *if* implication, let $\beta := \tilde{\beta}_p + \frac{1}{N}$, whence by part (1), $\sup_{x\in[-1,1]} H_{\beta,0,p}(x) > 0$ for all $N \geqslant 1$. Since $H_{\beta,0,p}(0) = 0$, for each N there exists $x_N \neq 0$ such that $H_{\beta,0,p}(x_N) > 0$. Let x_{N_k} be a convergent subsequence of x_N, converging to a point x^*. Then,

$$\lim_{k\to\infty} H_{\beta_{N_k},0,p}(x_{N_k}) = H_{\tilde{\beta}_p,0,p}(x^*),$$

and hence, $H_{\tilde{\beta}_p,0,p}(x^*) \geqslant 0$. However, by part (1), the reverse inequality is true, and hence, $H_{\tilde{\beta}_p,0,p}(x^*) = 0$, and hence, $0, x^*$ and $-x^*$ are all global maximizers of $H_{\tilde{\beta}_p,0,p}$. We will be done, if we can show that $x^* \neq 0$. Towards this, note that since $\lim_{\varepsilon\to 0} H_{\tilde{\beta}_p,0,p}(\varepsilon)/\varepsilon^2 = -\frac{1}{2}$, there exists $\delta > 0$ such that $H_{\tilde{\beta}_p,0,p}(\varepsilon) < -\varepsilon^2/4$ whenever $|\varepsilon| < \delta$. Suppose that $x^* = 0$, i.e. $x_{N_k} \to 0$ as $k \to \infty$. Then for all k large enough, we must have

$$H_{\beta_{N_k},0,p}(x_{N_k}) = H_{\tilde{\beta}_p,0,p}(x_{N_k}) + \frac{x_{N_k}^p}{N_k} < -\frac{x_{N_k}^2}{4} + \frac{x_{N_k}^p}{N_k} < 0,$$

a contradiction. This shows that $x^* \neq 0$. The proof of (3) and Lemma B.1 is now complete. □

Remark B.1. The argument in the last paragraph of the proof of Lemma B.1 can be adopted to show that for odd p, $H_{\tilde{\beta}_p,0,p}$ has exactly two global maximizers, one at 0 and the other one positive.

We now proceed to describe p-special points. To begin with, for convenience in the proof, we introduce the following notation.

Definition 5. A point $(\beta, h) \in [0, \infty) \times \mathbb{R}$ is said to be *p-locally special*, if the function $H_{\beta,h,p}$ has a local maximizer m satisfying $H''_{\beta,h,p}(m) = 0$.

We will see that every p-locally special point is actually p-special, and hence, the two notions are identical. In the following lemma, we give exact expressions for p-special points.

Lemma B.2 (Description of p-special points). *Define*

$$\check{\beta}_p := \frac{1}{2(p-1)}\left(\frac{p}{p-2}\right)^{\frac{p-2}{2}} \quad \textit{and} \quad \check{h}_p := \tanh^{-1}\left(\sqrt{\frac{p-2}{p}}\right) - p\check{\beta}_p\left(\frac{p-2}{p}\right)^{\frac{p-1}{2}}.$$

Then, we have the following:

(1) *If $p \geqslant 3$ is odd, then $\left(\check{\beta}_p, \check{h}_p\right)$ is the only p-locally special point in $[0, \infty) \times \mathbb{R}$. In this case, $m_* := \sqrt{1 - 2/p}$ is the only solution to the equation $H''_{\check{\beta}_p,\check{h}_p,p}(x) = 0$. In fact, m_* is a global maximizer of $H_{\check{\beta}_p,\check{h}_p,p}$ satisfying $H^{(3)}_{\check{\beta}_p,\check{h}_p,p}(m_*) = 0$ and $H^{(4)}_{\check{\beta}_p,\check{h}_p,p}(m_*) < 0$. Further, m_* is the unique stationary point of $H_{\check{\beta}_p,\check{h}_p,p}$.*

(2) *If $p \geqslant 4$ is even, then $\left(\check{\beta}_p, \check{h}_p\right)$ and $\left(\check{\beta}_p, -\check{h}_p\right)$ are the only p-locally special points in $[0, \infty) \times \mathbb{R}$. In this case, $m_*(1) := \sqrt{1 - 2/p}$ and $m_*(-1) := -m_*(1)$ are the only solutions to each of the equations $H''_{\check{\beta}_p,i\check{h}_p,p}(x) = 0$ for $i \in \{-1, 1\}$. In fact, $m_*(i)$ is a global maximizer of $H_{\check{\beta}_p,i\check{h}_p,p}$ for $i \in \{-1, 1\}$ satisfying*

$$H^{(3)}_{\check{\beta}_p,i\check{h}_p,p}(m_*(i)) = 0 \textit{ and } H^{(4)}_{\check{\beta}_p,i\check{h}_p,p}(m_*(i)) < 0, \quad \textit{for } i \in \{-1, 1\}.$$

Further, $m^*(i)$ is the unique global maximizer of $H_{\check{\beta}_p,i\check{h}_p,p}$ for $i \in \{-1, 1\}$.

Hence, a point (β, h) is p-locally special if and only if it is p-special.

Proof of Lemma B.2. We start with the following proposition: □

Proposition 1. *Let $\beta := \check{\beta}_p$, $h \in \mathbb{R}$, and let $y \in (0, 1)$ be a local maximum of $H_{\beta,h,p}$, satisfying $H''_{\beta,h,p}(y) = H^{(3)}_{\beta,h,p}(y) = 0$. Then $H^{(4)}_{\beta,h,p}(y) < 0$.*

Proof. For convenience, we will denote $N_{\beta,h,p} := (1 - x^2)H''_{\beta,h,p}(x)$ by N and $H_{\beta,h,p}$ by H. Note that

$$N''(x) = (1 - x^2)H^{(4)}(x) - 4xH^{(3)}(x) - 2H''(x).$$

By hypothesis, $N''(y) = (1 - y^2)H^{(4)}(y)$. Now,

$$N''(x) = \beta p(p-1)(p-2)(p-3)x^{p-4} - \beta p^2(p-1)^2 x^{p-2}$$

cannot have any root other than 0 and $\pm\sqrt{\frac{(p-2)(p-3)}{p(p-1)}}$. But we know from the proof of Lemma B.2 that $H''_{\beta,h,p}$ cannot have any root other than $\pm\sqrt{1 - 2/p}$ (note that Proposition 1 is not needed to reach this conclusion, and hence, there is no circularity in the argument), and for $p \geqslant 3$, we have $\frac{(p-2)(p-3)}{p(p-1)} < \frac{p-2}{p}$. Therefore, y is not a root of N'', and hence, not a root of $H^{(4)}$. Proposition 1 now follows from the standard higher derivative test. □

We are now proceed with the proof of Lemma B.2. We start by proving that the first coordinate of every p-locally special point in $[0, \infty) \times \mathbb{R}$ must be equal to $\check{\beta}_p$. Towards this, we first claim that $H''_{\beta,h,p}(x) < 0$, or equivalently, $N_{\beta,h,p}(x) < 0$ for all $x \in (-1, 1)$, if $\beta < \check{\beta}_p$. This will rule out the possibility of (β, h) being a candidate for a p-locally special point, for $\beta < \check{\beta}_p$. Towards proving this claim, we can assume that

$$\sup_{x\in(-1,1)} N_{\beta,h,p}(x) > -1,$$

since otherwise we would be done. Since $N_{\beta,h,p}(-1) = N_{\beta,h,p}(0) = N_{\beta,h,p}(1) = -1$, the function $N_{\beta,h,p}$ attains maximum at some $m \in (-1, 1)\backslash\{0\}$, and hence, m is a non-zero solution to the equation $N'_{\beta,h,p}(x) = 0$. Therefore, from the proof of (3) in Lemma B.1, that $m \in \{-q, q\}$, where $q := \sqrt{1 - 2/p}$. Since $N_{\beta,h,p}(q) \geqslant N_{\beta,h,p}(-q)$, we know for sure that q is a global maximizer of $N_{\beta,h,p}$. Our claim now follows from the observation that $\beta < \check{\beta}_p \implies N_{\beta,h,p}(q) < 0$.

Now, we are going to rule out the possibility $\beta > \check{\beta}_p$, as well. Suppose that $\beta > \check{\beta}_p$, and let m_* be a local maximizer of $H_{\beta,h,p}$ satisfying $H''_{\beta,h,p}(m_*) = 0$, i.e. $N_{\beta,h,p}(m_*) = 0$. Now, $N_{\beta,h,p}(0) = -1 \implies m_* \neq 0$. Next, since $\beta > \check{\beta}_p$, it follows that $N_{\beta,h,p}(q) > 0$, and hence, $m_* \neq q$. If p is even, then $N_{\beta,h,p}(-q) = N_{\beta,h,p}(q) > 0$, and if p is odd, then $N_{\beta,h,p}(x) < -1$ for all $x < 0$. Thus, in either case, $m_* \neq -q$. All these show that $N'_{\beta,h,p}(m_*) \neq 0$. Suppose that $N'_{\beta,h,p}(m_*) > 0$. Since $N_{\beta,h,p}(m_*) = 0$, there exists $\varepsilon > 0$ such that $N_{\beta,h,p}(x) > 0$ for all $x \in (m_*, m_* + \varepsilon)$ and $N_{\beta,h,p}(x) < 0$ for all $x \in (m_*, m_* - \varepsilon)$. Thus, $H''_{\beta,h,p}(x) > 0$ for all $x \in (m_*, m_* + \varepsilon)$ and $H''_{\beta,h,p}(x) < 0$ for all $x \in (m_* - \varepsilon, m_*)$. Since $H'_{\beta,h,p}(m_*) = 0$, we must have

$$H'_{\beta,h,p}(x) > 0 \quad \text{for all } x \in (m_* - \varepsilon, m_* + \varepsilon)\backslash\{m_*\}.$$

This implies that $H_{\beta,h,p}$ is strictly increasing on $[m_*, m_* + \varepsilon)$, contradicting that m_* is a local maximizer of $H_{\beta,h,p}$. Similarly, if $N'_{\beta,h,p}(m_*) < 0$, then there exists $\varepsilon > 0$ such that $H'_{\beta,h,p}(x) < 0$ for all $x \in (m_* - \varepsilon, m_* + \varepsilon)\backslash\{m_*\}$, and so, $H_{\beta,h,p}(x)$ is strictly decreasing on $(m_* - \varepsilon, m_*]$, contradicting once again, that m_* is a local maximizer of $H_{\beta,h,p}$. We have thus proved our claim, that the first coordinate of every p-special point in $[0, \infty) \times \mathbb{R}$ must be equal to $\check{\beta}_p$. In what follows, let $\beta := \check{\beta}_p$.

Proof of (1). Let $p \geqslant 3$ be odd and let m_* be any solution to the equation $H''_{\beta,h,p}(x) = 0$, or equivalently, to the equation $N_{\beta,h,p}(x) = 0$. Since $N_{\beta,h,p}(x) \leqslant -1$ for all $x \leqslant 0$, it follows that $m_* \in (0, 1)$. Now, we already know that the only positive root of $N'_{\beta,h,p}$ is $q := \sqrt{1 - 2/p}$, and since $N_{\beta,h,p}(q) = 0$, by Rolle's theorem, $N_{\beta,h,p}$ cannot have any positive root other than q. Thus, $m_* = q$ is the only root of $H''_{\beta,h,p}$. Since $N_{\beta,h,p}(m_*) = N'_{\beta,h,p}(m_*) = 0$, we have

$$H^{(3)}_{\beta,h,p}(m_*) = \frac{N'_{\beta,h,p}(m_*)(1 - m_*^2) + 2m_* N_{\beta,h,p}(m_*)}{(1 - m_*^2)^2} = 0.$$

Now, m_* is a stationary point of $H_{\beta,h,p}$, i.e. $H'_{\beta,h,p}(m_*) = 0$ if and only if $h = \check{h}_p$. Hence, $(\check{\beta}_p, \check{h}_p)$ is the only candidate for being a p-locally special point in $[0, \infty) \times \mathbb{R}$. Let $h := \check{h}_p$ throughout the rest of the proof of (a). Since $H'_{\beta,h,p}(m_*) = 0$ and m_* is the only root of $H''_{\beta,h,p}$, by Rolle's theorem, $H'_{\beta,h,p}$ cannot have any root other than m_*. This implies that the sign of $H'_{\beta,h,p}$ remains constant on each of the intervals $(-1, m^*)$ and $(m^*, 1)$. Since

$$\lim_{x\to -1^+} H'_{\beta,h,p}(x) = +\infty \quad \text{and} \quad \lim_{x\to 1^-} H'_{\beta,h,p}(x) = -\infty,$$

we conclude that $H'_{\beta,h,p} > 0$ on $(-1, m^*)$ and $H'_{\beta,h,p} < 0$ on $(m^*, 1)$, thereby showing that m^* is a global maximizer, and also the unique stationary point of $H_{\beta,h,p}$, and verifying that $(\check{\beta}_p, \check{h}_p)$ is actually a p-special point. The result in part (1) now follows from Proposition 1. □

Proof of (2). Let $p \geqslant 4$ be even. Since $m_*(1)$ and $m_*(-1)$ are the only non-zero roots of $N'_{\beta,h,p}$, and they are also roots of $N_{\beta,h,p}$, by Rolle's theorem, they are the only roots of $N_{\beta,h,p}$, as well. Hence, the only roots of $H''_{\beta,h,p}$ are $m_*(1)$ and $m_*(-1)$, and so, $H^{(3)}_{\beta,h,p}(m_*(1)) = H^{(3)}_{\beta,h,p}(m_*(-1)) = 0$.
For $i \in \{-1, 1\}$, note that $m_*(i)$ is a stationary point of $H_{\beta,h,p}$, i.e. $H'_{\beta,h,p}(m_*(i)) = 0$, if and only if $h = i\check{h}_p$. Hence, $(\check{\beta}_p, \check{h}_p)$ and $(\check{\beta}_p, -\check{h}_p)$ are the only candidates for being p-locally special points in $[0, \infty) \times \mathbb{R}$. Let $h := \check{h}_p$ throughout the rest of the proof of (2). Since $H'_{\beta,ih,p}(m_*(i)) = 0$ and $m_*(i)$ is the only root of $H''_{\beta,ih,p}$ with sign i, by Rolle's theorem, $H'_{\beta,ih,p}$ cannot have 0 or any point with sign i as a root, other than $m_*(i)$. This implies that the sign of $H'_{\beta,h,p}$ remains constant on each of the intervals $[0, m_*(1))$ and $(m_*(1), 1)$, and the sign of $H'_{\beta,-h,p}$ remains constant on each of the intervals $(-1, m_*(-1))$ and $(m_*(-1), 0]$. Since

$$\lim_{x\to -1^+} H'_{\beta,\pm h,p}(x) = +\infty \quad \text{and} \quad \lim_{x\to 1^-} H'_{\beta,\pm h,p}(x) = -\infty,$$

we conclude that $H'_{\beta,h,p} < 0$ on $(m_*(1), 1)$ and $H'_{\beta,-h,p} > 0$ on $(-1, m_*(-1))$. Now, note that

$$\begin{aligned}
h &= \tanh^{-1}\left(\sqrt{\frac{p-2}{p}}\right) - \check{\beta}_p p \left(\frac{p-2}{p}\right)^{\frac{(p-1)}{2}} \\
&= \left[\sum_{k=0}^{\infty} \frac{\left(\sqrt{\frac{p-2}{p}}\right)^{2k+1}}{2k+1}\right] - \frac{p}{2(p-1)}\sqrt{\frac{p-2}{p}} \\
&\geqslant \sqrt{\frac{p-2}{p}} - \frac{p}{2(p-1)}\sqrt{\frac{p-2}{p}} \\
&= \frac{p-2}{2(p-1)}\sqrt{\frac{p-2}{p}} > 0.
\end{aligned}$$

Hence, $H'_{\beta,h,p}(0) = h > 0$ and $H'_{\beta,-h,p}(0) = -h < 0$. Consequently, $H'_{\beta,h,p} > 0$ on $[0, m_*(1))$ and $H'_{\beta,-h,p} < 0$ on $(m_*(-1), 0]$. Thus, $m_*(i)$ is the unique global maximizer of $H_{\beta,ih,p}$ over the interval $\mathscr{I}_i := \{ix : x \in [0, 1]\}$. (Note that $\mathscr{I}_1 = [0, 1]$ and $\mathscr{I}_{-1} = [-1, 0]$.) Now, it is easy to see that $H_{\beta,ih,p}(x) < H_{\beta,ih,p}(-x)$, for all

$x \in [-1, 1]\backslash \mathscr{J}_i$. This shows that $m_*(i)$ is the unique global maximizer of $H_{\beta,ih,p}$ over $[-1, 1]$. Part (2) now follows from Proposition 1, and the proof of Lemma B.2 is now complete. □

Next, we give a description of p-weakly critical points that is, points (β, h) for which the function $H_{\beta,h,p}$ has exactly two global maximizers). Note that we already have a full characterization of p-strongly critical points (that is, points (β, h) for which the function $H_{\beta,h,p}$ has exactly three global maximizers) by part (3) of Lemma B.1. To elaborate, we know that there cannot be any p-strongly critical point if p is odd, and if $p \geqslant 4$ is even, then $(\tilde{\beta}_p, 0)$ is the only p-strongly critical point. In the following lemma, we show that the set of all p-critical points is a one-dimensional continuous curve in the plane $[0, \infty) \times \mathbb{R}$. We also prove some other interesting properties of this curve, for instance, the only limit point(s) of the curve which is (are) outside it, is (are) the p-special point(s).

Lemma B.3 (Description of p-weakly critical points). *For every $p \geqslant 3$, $\check{\beta}_p < \tilde{\beta}_p$, and the set ${\mathscr{C}_p}^+$ can be characterized as follows.*

(1) *For every even $p \geqslant 4$, there exists a continuous function $\varphi_p : (\check{\beta}_p, \infty) \mapsto [0, \infty)$ which is strictly decreasing on $(\check{\beta}_p, \tilde{\beta}_p)$ and vanishing on $[\tilde{\beta}_p, \infty)$, such that*

$$\mathscr{C}_p^+ = \left\{(\beta, \pm\varphi_p(\beta)) : \beta \in (\check{\beta}_p, \infty)\backslash\{\tilde{\beta}_p\}\right\}.$$

(2) *For every odd $p \geqslant 3$, there exists a strictly decreasing, continuous function φ_p : $(\check{\beta}_p, \infty) \mapsto \mathbb{R}$ satisfying $\varphi_p(\tilde{\beta}_p) = 0$ and $\lim_{\beta\to\infty} \varphi_p(\beta) = -\infty$, such that*

$$\mathscr{C}_p^+ = \left\{(\beta, \varphi_p(\beta)) : \beta \in (\check{\beta}_p, \infty)\right\}.$$

In both cases, $\lim_{\beta\to\check{\beta}_p^+} \varphi_p(\beta) = \tanh^{-1}(m_*) - p\check{\beta}_p m_*^{p-1}$, *where* $m_* := \sqrt{\frac{p-2}{p}}$.

Proof. First, we prove that $\check{\beta}_p < \tilde{\beta}_p$ for all $p \geqslant 3$. Since

$$\sup_{x\in[-1,1]} H_{\beta,0,p+1}(x) = \sup_{x\in[0,1]} H_{\beta,0,p+1}(x) \leqslant \sup_{x\in[0,1]} H_{\beta,0,p}(x) = \sup_{x\in[-1,1]} H_{\beta,0,p}(x),$$

it follows that $\tilde{\beta}_{p+1} \geqslant \tilde{\beta}_p$, i.e. $\tilde{\beta}_p$ is increasing in p. Therefore, $\tilde{\beta}_p \geqslant \tilde{\beta}_2 = \frac{1}{2}$ for all $p \geqslant 3$. First note that $\check{\beta}_3 = \frac{\sqrt{3}}{4} < \frac{1}{2}$. Next, note that for $p \geqslant 4$,

$$\check{\beta}_p = \frac{1}{2(p-1)}\left(1 + \frac{2}{p-2}\right)^{\frac{p-2}{2}} \leqslant \frac{e}{2(p-1)} \leqslant \frac{e}{6} < \frac{1}{2}.$$

Hence, $\check{\beta}_p < \frac{1}{2} \leqslant \tilde{\beta}_p$ for all $p \geqslant 3$.

Next, we show that $\mathscr{C}_p^+ \subseteq (\check{\beta}_p, \infty) \times \mathbb{R}$. Towards this, first let $\beta < \check{\beta}_p$ and $h \in \mathbb{R}$. It follows from the proof of Lemma B.2, that $H''_{\beta,h,p} < 0$ on $[-1, 1]$, so $H_{\beta,h,p}$ is strictly concave on $[-1, 1]$, and hence, can have at most one global maximum. Therefore, $(\beta, h) \notin \mathscr{C}_p^+$. Now, let $\beta = \check{\beta}_p$ and $h \in \mathbb{R}$. From the proof of Lemma B.2, we know that $H''_{\beta,h,p}$ cannot have any root on $[-1, 1]$ other than possibly $\pm\sqrt{1 - 2/p}$. Since $H''_{\beta,h,p}(-1) = H''_{\beta,h,p}(1) = -\infty$, $H''_{\beta,h,p}(0) = -1$ and $H''_{\beta,h,p}$ is continuous, $H''_{\beta,h,p}(x) < 0$ for all $x \in [-1, 1]\backslash\{\pm\sqrt{1 - 2/p}\}$. This shows that $H'_{\beta,h,p}$ is strictly decreasing on $[-1, 1]$, and hence, $H_{\beta,h,p}$ can have at most one stationary point. Consequently, $(\beta, h) \notin \mathscr{C}_p^+$, proving our claim that $\mathscr{C}_p^+ \subseteq (\check{\beta}_p, \infty) \times \mathbb{R}$. We now consider the cases of even and odd p separately.

Proof of (1). Let $p \geqslant 4$ be even. Since $x \mapsto \beta x^p - I(x)$ is an even function, the set $\mathscr{C}_p^+$ is symmetric about the line $h = 0$, i.e. $(\beta, h) \in \mathscr{C}_p{}^+ \implies (\beta, -h) \in \mathscr{C}_p{}^+$. Next, we show that for every $\beta > \check{\beta}_p$, there exists at most one $h \geqslant 0$ such that $(\beta, h) \in \mathscr{C}_p{}^+$. Suppose towards a contradiction, that there exists $\beta > \check{\beta}_p$ and $h_2 > h_1 \geqslant 0$, such that both (β, h_1) and $(\beta, h_2) \in \mathscr{C}_p{}^+$. Letting $m_* := \sqrt{1 - 2/p}$, it follows that $H''_{\beta,h,p}(m_*) > 0$ for all $h \in \mathbb{R}$. Recalling that $H''_{\beta,h,p}$ can have at most two roots in $[0, 1]$, and using the facts

$$H''_{\beta,h,p}(0) = -1,\ H''_{\beta,h,p}(1) = -\infty,$$

it follows that there exist $0 < a_1 < m_* < a_2 < 1$, such that $H''_{\beta,h,p} < 0$ on $[0, a_1)$, $H''_{\beta,h,p}(a_1) = 0$, $H''_{\beta,h,p} > 0$ on (a_1, a_2), $H''_{\beta,h,p}(a_2) = 0$ and $H''_{\beta,h,p} < 0$ on $(a_2, 1]$. This shows that $H'_{\beta,h,p}$ is strictly decreasing on $[0, a_1]$, strictly increasing on $[a_1, a_2]$ and strictly decreasing on $[a_2, 1]$.
First assume that $h_1 > 0$, whence the two global maximizers $m_1(h_i) < m_2(h_i)$ of $H_{\beta,h_i,p}$ must be positive roots of $H'_{\beta,h_i,p}$ for $i \in \{1, 2\}$. Note that the monotonicity pattern of the function $H'_{\beta,h_i,p}$ implies that $m_1(h_i) \in (0, a_1)$ and $m_2(h_i) \in (a_2, 1)$. Hence, $H'_{\beta,h_i,p}(a_1) < 0$ and $H'_{\beta,h_i,p}(a_2) > 0$, and by the intermediate value theorem, there exists $m(h_i) \in (a_1, a_2)$ such that

$$H'_{\beta,h_i,p}(m(h_i)) = 0.$$

Observe that $H'_{\beta,h_i,p}$ is positive on $[0, m_1(h_i))$, negative on $(m_1(h_i), m(h_i))$, positive on $(m(h_i), m_2(h_i))$ and negative on $(m_2(h_i), 1]$. Since $h_2 > h_1$, it follows that $H'_{\beta,h_2,p} > 0$ on $[0, m_1(h_1)]$ and on $[m(h_1), m_2(h_1)]$. However, since $m_1(h_2)$, $m(h_2)$ and $m_2(h_2)$ are roots of $H'_{\beta,h_2,p}$ on $(0, a_1)$, (a_1, a_2) and $(a_2, 1)$ respectively, it follows that $m_1(h_1) < m_1(h_2)$, $m(h_2) < m(h_1)$ and $m_2(h_1) < m_2(h_2)$. Combining all these, gives

$$\int_{m_1(h_1)}^{m(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m(h_2)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t \tag{B.1}$$

and

$$\int_{m(h_1)}^{m_2(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m(h_1)}^{m_2(h_1)} H'_{\beta,h_2,p}(t)\mathrm{d}t < \int_{m(h_2)}^{m_2(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t \tag{B.2}$$

Adding (B.1) and (B.2), we have

$$\int_{m_1(h_1)}^{m_2(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m_2(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t. \tag{B.3}$$

This is a contradiction, since both sides of (B.3) are 0.
Therefore, it must be that $h_1 = 0$. In this case, the global maximizers $m_1(h_1) < m_2(h_1)$ of $H_{\beta,h_1,p}$ satisfy $m_1(h_1) = -m_2(h_1)$. Since $H'_{\beta,h_1,p}$ vanishes at 0, it must be negative on $(0, a_1]$. Hence, $m_2(h_1) \in (a_2, 1)$. This shows that $H'_{\beta,h_1,p}(a_2) > 0$, and hence,

there exists $m(h_1) \in (a_1, a_2)$ such that $H'_{\beta,h_1,p}(m(h_1)) = 0$. Observe that $H'_{\beta,h_1,p}$ is negative on $(0, m(h_1))$, positive on $(m(h_1), m_2(h_1))$ and negative on $(m_2(h_1), 1)$. Therefore, since $h_2 > h_1$, $H'_{\beta,h_2,p} > 0$ on $[m(h_1), m_2(h_1)]$. Since $m(h_2)$ and $m_2(h_2)$ are roots of $H'_{\beta,h_2,p}$ on (a_1, a_2) and $(a_2, 1)$ respectively, we must have $m(h_2) < m(h_1)$ and $m_2(h_1) < m_2(h_2)$. Hence, we have

$$\int_0^{m(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m(h_2)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t \tag{B.4}$$

and

$$\int_{m(h_1)}^{m_2(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m(h_1)}^{m_2(h_1)} H'_{\beta,h_2,p}(t)\mathrm{d}t < \int_{m(h_2)}^{m_2(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t \tag{B.5}$$

Adding (B.4) and (B.5), gives

$$\int_0^{m_2(h_1)} H'_{\beta,h_1,p}(t)\mathrm{d}t < \int_{m_1(h_2)}^{m_2(h_2)} H'_{\beta,h_2,p}(t)\mathrm{d}t. \tag{B.6}$$

Once again, this is a contradiction, since the right side of (B.6) is 0, whereas the left side of (B.6) is non-negative. This completes the proof of our claim that for every $\beta > \check{\beta}_p$, there exists at most one $h \geqslant 0$ such that $(\beta, h) \in \mathscr{C}_p{}^+$.

We now show that for all $\beta \in (\check{\beta}_p, \infty)\backslash\{\tilde{\beta}_p\}$, there exists at least one $h \geqslant 0$ such that $(\beta, h) \in \mathscr{C}_p{}^+$. First, suppose that $\beta > \tilde{\beta}_p$. In this case, $\sup_{x\in[-1,1]} H_{\beta,0,p}(x) > 0$ by the definition in (2.2), and hence, $H_{\beta,0,p}$ has a non-zero global maximizer m_*. Since $H_{\beta,0,p}$ is an even function, $-m_*$ is also a global maximizer. It now follows from part (3) of Lemma B.1, that $H_{\beta,0,p}$ has exactly two global maximizers, and hence, $(\beta, 0) \in \mathscr{C}_p{}^+$. Next, let $\beta \in (\check{\beta}_p, \tilde{\beta}_p)$. Recall that the function $H'_{\beta,0,p}$ is continuous and strictly decreasing on each of the intervals $[0, a_1]$ and $[a_2, 1)$. Hence, the functions

$$\psi_1 := H'_{\beta,0,p}\Big|_{[0,a_1]} \quad \text{and} \quad \psi_2 := H'_{\beta,0,p}\Big|_{[a_2,1)}$$

are invertible, and by Proposition 2.1 in [22], the functions ψ_1^{-1} and ψ_2^{-1} are continuous. Hence, the function $\Lambda : [H'_{\beta,0,p}(a_1), \min\{0, H'_{\beta,0,p}(a_2)\}] \to \mathbb{R}$ defined as:

$$\Lambda(h) := \int_{\psi_1^{-1}(h)}^{\psi_2^{-1}(h)} H'_{\beta,-h,p}(t)\mathrm{d}t = \int_{\psi_1^{-1}(h)}^{\psi_2^{-1}(h)} H'_{\beta,0,p}(t)\mathrm{d}t + h\left(\psi_1^{-1}(h) - \psi_2^{-1}(h)\right)$$

is continuous. Since the function $t \mapsto H'_{\beta,0,p}(t) - H'_{\beta,0,p}(a_1)$ is strictly positive on the interval $(a_1, \psi_2^{-1}(H'_{\beta,0,p}(a_1)))$ (because it is strictly increasing on $[a_1, a_2]$, strictly decreasing on $[a_2, 1)$, and vanishes at the endpoints a_1 and $\psi_2^{-1}(H'_{\beta,0,p}(a_1))$ of the interval),

$$\Lambda(H'_{\beta,0,p}(a_1)) = \int_{a_1}^{\psi_2^{-1}(H'_{\beta,0,p}(a_1))} \left(H'_{\beta,0,p}(t) - H'_{\beta,0,p}(a_1)\right)\mathrm{d}t > 0. \tag{B.7}$$

Next, suppose that $H'_{\beta,0,p}(a_2) \leqslant 0$. Since the function $t \mapsto H'_{\beta,0,p}(t) - H'_{\beta,0,p}(a_2)$ is strictly negative on the interval $(\psi_1^{-1}(H'_{\beta,0,p}(a_2)), a_2)$ (because it is strictly decreasing on $[0, a_1]$, strictly increasing on $[a_1, a_2]$, and vanishes at the endpoints $\psi_1^{-1}(H'_{\beta,0,p}(a_2))$ and a_2 of the interval),

$$\Lambda(H'_{\beta,0,p}(a_2)) = \int_{\psi_1^{-1}(H'_{\beta,0,p}(a_2))}^{a_2} \left(H'_{\beta,0,p}(t) - H'_{\beta,0,p}(a_2)\right) \mathrm{d}t < 0. \tag{B.8}$$

Finally, suppose that $H'_{\beta,0,p}(a_2) > 0$. Then we have

$$\Lambda(0) = \int_0^{\psi_2^{-1}(0)} H'_{\beta,0,p}(t)\mathrm{d}t = H_{\beta,0,p}(\psi_2^{-1}(0)) < 0. \tag{B.9}$$

The last inequality in (B.9) follows from the facts that $\psi_2^{-1}(0) > 0$ and $\beta < \tilde{\beta}_p$. Using (B.7), (B.8), (B.9) and the intermediate value theorem, we conclude that there exists $h(\beta) \in (H'_{\beta,0,p}(a_1), \min\{0, H'_{\beta,0,p}(a_2)\})$ such that $\Lambda(h(\beta)) = 0$, i.e.

$$H_{\beta,-h(\beta),p}(\psi_1^{-1}(h(\beta))) = H_{\beta,-h(\beta),p}(\psi_2^{-1}(h(\beta))). \tag{B.10}$$

Now, $\psi_1^{-1}(h(\beta)) \in (0, a_1)$ and $\psi_2^{-1}(h(\beta)) \in (a_2, 1)$, and hence, $H'_{\beta,-h(\beta),p}$ is strictly decreasing on some open neighborhoods of $\psi_1^{-1}(h(\beta))$ and $\psi_2^{-1}(h(\beta))$. Since $H'_{\beta,-h(\beta),p}(\psi_1^{-1}(h(\beta))) = H'_{\beta,-h(\beta),p}(\psi_2^{-1}(h(\beta))) = 0$, the points $\psi_1^{-1}(h(\beta))$ and $\psi_2^{-1}(h(\beta))$ are local maximizers of $H_{\beta,-h(\beta),p}$. Since $-h(\beta) > 0$, any global maximizer of $H_{\beta,-h(\beta),p}$ must be a positive root of $H'_{\beta,-h(\beta),p}$, and further, it cannot lie on the interval $[a_1, a_2]$, since $H'_{\beta,-h(\beta),p}$ is strictly increasing on this interval. Hence, one of $\psi_1^{-1}(h(\beta))$ and $\psi_2^{-1}(h(\beta))$ must be a global maximizer of $H_{\beta,-h(\beta),p}$, and by (B.10), both must be global maximizers of $H_{\beta,-h(\beta),p}$. By part (3) of Lemma B.1, these are the only global maximizers of $H_{\beta,-h(\beta),p}$, and hence, $(\beta, -h(\beta)) \in \mathscr{C}_p{}^+$.
Next, if $\beta = \tilde{\beta}_p$, then $H_{\beta,0,p}$ has three global maximizers, so $(\beta, 0) \notin \mathscr{C}_p{}^+$. One of these global maximizers is 0 and the other two are negative of one another. It follows from the argument used in proving the uniqueness of h under the case $h_1 = 0$, that

$$\int_{m_1(h)}^{m_2(h)} H'_{\beta,h,p}(t)\mathrm{d}t > 0,$$

for every $h > 0$, where $m_2(h) > m_1(h) > 0$ are possible global maximizers of $H_{\beta,h,p}$ (see inequality (B.6)), which is a contradiction. Hence,

$$\mathscr{C}_p{}^+ \subseteq \left(\{\tilde{\beta}_p\} \times \mathbb{R}\right)^c.$$

At this point, we completed proving that for every $\beta \in (\check{\beta}_p, \infty)\backslash\{\tilde{\beta}_p\}$, there exists unique $h \geqslant 0$ such that $(\beta, h) \in \mathscr{C}_p{}^+$, and further, there exists no such h for $\beta = \tilde{\beta}_p$. Denote by $\varphi_p(\beta)$, this unique h corresponding to $\beta \in (\check{\beta}_p, \infty)\backslash\{\tilde{\beta}_p\}$. Our proof so far, also reveals that $\varphi_p(\beta) = 0$ for $\beta > \tilde{\beta}_p$ and $\varphi_p(\beta) > 0$ for $\beta \in (\check{\beta}_p, \tilde{\beta}_p)$. Define $\varphi_p(\tilde{\beta}_p) = 0$ for the sake of completing its definition on the whole of $(\check{\beta}_p, \infty)$.
We now show that φ_p is strictly decreasing on $(\check{\beta}_p, \tilde{\beta}_p)$. Towards this, take $\check{\beta}_p < \beta_1 < \beta_2 < \tilde{\beta}_p$. Let $h_1 := \varphi_p(\beta_1)$ and $h_2 := \varphi_p(\beta_2)$ (we already know from the proof of the existence part, that h_1 and h_2 are positive), and suppose towards a contradiction, that $h_1 \leqslant h_2$. Then, $H'_{\beta_1,h_1,p} < H'_{\beta_2,h_2,p}$ on $(0, 1]$. Let $m_{11} < m_{13}$ be the global

maximizers of $H_{\beta_1,h_1,p}$ and $m_{21} < m_{23}$ be the global maximizers of $H_{\beta_2,h_2,p}$. Also, let $m_{12} \in (m_{11}, m_{13})$ and $m_{22} \in (m_{21}, m_{23})$ be local minimizers of $H_{\beta_1,h_1,p}$ and $H_{\beta_2,h_2,p}$, respectively. We have already shown that for $i \in \{1, 2\}$, the function $H'_{\beta_i,h_i,p}$ is positive on $[0, m_{i1})$, negative on (m_{i1}, m_{i2}), positive on (m_{i2}, m_{i3}) and negative on $(m_{i3}, 1)$. Since $H'_{\beta_2,h_2,p} > 0$ on $[0, m_{11}]$, we must have $m_{21} > m_{11}$. On the other hand, we have $m_{21} < m_* := \sqrt{1 - 2/p} < m_{13}$. This, combined with the fact that $H'_{\beta_2,h_2,p} > 0$ on $[m_{12}, m_{13}]$, implies that $m_{21} < m_{12}$. Next, since $H'_{\beta_1,h_1,p} < 0$ on $[m_{21}, m_{22}]$ and $H'_{\beta_1,h_1,p}(m_{12}) = 0$, it follows that $m_{22} < m_{12}$. Finally, since $H'_{\beta_1,h_1,p} < 0$ on $[m_{23}, 1)$, we must have $m_{13} < m_{23}$. Hence, we have

$$m_{11} < m_{21} < m_{22} < m_{12} < m_{13} < m_{23}.$$

Using this and proceeding exactly as in the proof of the uniqueness of h, we have

$$\int_{m_{11}}^{m_{12}} H'_{\beta_1,h_1,p}(t)\mathrm{d}t < \int_{m_{21}}^{m_{22}} H'_{\beta_2,h_2,p}(t)\mathrm{d}t$$
$$\text{and} \quad \int_{m_{12}}^{m_{13}} H'_{\beta_1,h_1,p}(t)\mathrm{d}t < \int_{m_{22}}^{m_{23}} H'_{\beta_2,h_2,p}(t)\mathrm{d}t.$$

Adding the above two inequalities, we have

$$\int_{m_{11}}^{m_{13}} H'_{\beta_1,h_1,p}(t)\mathrm{d}t < \int_{m_{21}}^{m_{23}} H'_{\beta_2,h_2,p}(t)\mathrm{d}t,$$

which is a contradiction once again, since both sides of the above inequality are 0. Hence, we must have $h_1 > h_2$, showing that φ_p is strictly decreasing on $(\check{\beta}_p, \tilde{\beta}_p)$.

Next, we show that φ_p is continuous on $(\check{\beta}_p, \tilde{\beta}_p]$. Towards this, first take $\beta \in (\check{\beta}_p, \tilde{\beta}_p)$, and let $\{\beta_n\}_{n \geqslant 1}$ be a monotonic sequence in $(\check{\beta}_p, \tilde{\beta}_p)$ converging to β. Since φ_p is decreasing on $(\check{\beta}_p, \tilde{\beta}_p)$, it follows that $\varphi_p(\beta_n)$ is monotonic as well (the direction of monotonicity being opposite to that of β_n). Moreover, $\varphi_p(\beta_n)$ is bounded between $\varphi_p(\beta_1)$ and $\varphi_p(\beta)$. Hence, $\lim_{n\to\infty} \varphi_p(\beta_n)$ exists, which we call h. Let $m_1(n) < m_2(n)$ denote the global maximizers of $H_{\beta_n,\varphi_p(\beta_n),p}$. Choose a subsequence n_k such that $m_1(n_k) \to m_1$ and $m_2(n_k) \to m_2$ for some $m_1, m_2 \in [-1, 1]$. Since

$$H_{\beta_{n_k},\varphi_p(\beta_{n_k}),p}(m_i(n_k)) \geqslant H_{\beta_{n_k},\varphi_p(\beta_{n_k}),p}(x) \quad \text{for all } x \in [-1, 1] \text{ and } i \in \{1, 2\},$$

taking limit as $k \to \infty$ on both sides, we have $H_{\beta,h,p}(m_i) \geqslant H_{\beta,h,p}(x)$ for all $x \in [-1, 1]$ and $i \in \{1, 2\}$, showing that m_1 and m_2 are global maximizers of $H_{\beta,h,p}$. We now show that $m_1 < m_2$. Since $\beta_n \to \beta > \check{\beta}_p$, there exists $\underline{\beta} > \check{\beta}_p$ such that $\beta_n > \underline{\beta}$ for all large n. If $a_1(\underline{\beta}) < a_2(\underline{\beta})$ are the positive roots of $H''_{\underline{\beta},0,p}$, then $H''_{\beta_n,0,p} > 0$ on $[a_1(\underline{\beta}), a_2(\underline{\beta})]$ for all large n, and hence, $m_1(n) < a_1(\underline{\beta})$ and $m_2(n) > a_2(\underline{\beta})$ for all large n. This shows that

$$m_1 \leqslant a_1(\underline{\beta}) < a_2(\underline{\beta}) \leqslant m_2$$

and hence, $m_1 < m_2$. Thus $H_{\beta,h,p}$ has at least two global maximizers. But $\beta \neq \tilde{\beta}_p$, and $H_{\beta,h,p}$ must therefore have exactly two global maximizers, showing that $(\beta, h) \in \mathscr{C}_p{}^+$. Since $h \geqslant 0$, by the uniqueness property, we must have $h = \varphi_p(\beta)$. Hence, $\lim_{n\to\infty} \varphi_p(\beta_n) = \varphi_p(\beta)$, showing that φ_p is continuous on $(\check{\beta}_p, \tilde{\beta}_p)$.

To show that $\lim_{\beta \to (\tilde{\beta}_p)^-} \varphi_p(\beta) = 0$, take a sequence $\beta_n \in (\check{\beta}_p, \tilde{\beta}_p)$ increasing to $\tilde{\beta}_p$, whence $\varphi_p(\beta_n)$ decreases to some $h \geqslant 0$. By the same arguments as before, it follows that $H_{\tilde{\beta}_p,h,p}$ has at least two global maximizers. If $h > 0$, then $H_{\tilde{\beta}_p,h,p}$ will have exactly two global maximizers. Therefore $(\tilde{\beta}_p, h) \in \mathscr{C}_p{}^+$, contradicting our finding that $\mathscr{C}_p{}^+ \subseteq (\{\tilde{\beta}_p\} \times \mathbb{R})^c$. This shows that $h = 0$, completing the proof of (1). □

Proof of (2). Let $p \geqslant 3$ be odd. In this case, $H''_{\beta,0,p} < 0$ on $[-1, 0]$ for all $\beta \geqslant 0$. Let $\beta > \check{\beta}_p$. Once again, $H''_{\beta,0,p}$ can have at most two positive roots, which, together with the facts $H''_{\beta,0,p}(m_*) > 0$ and $H''_{\beta,0,p}(1) = -\infty$, imply the existence of $0 < a_1 < m_* < a_2 < 1$, such that $H''_{\beta,0,p} < 0$ on $[-1, a_1) \bigcup (a_2, 1]$ and $H''_{\beta,0,p} > 0$ on (a_1, a_2). One can now follow the proof of (a) modulo obvious modifications, to show that there exists at most one $h \in \mathbb{R}$ such that $(\beta, h) \in \mathscr{C}_p{}^+$.

To show the existence of at least one such $h \in \mathbb{R}$, one can once again essentially follow the proof of (a) modulo a couple of minor modifications. To be specific, if we modify the definition of ψ_1 to $H'_{\beta,0,p}\big|_{(-1,a_1]}$, and change the domain of Λ to $[H'_{\beta,0,p}(a_1), H'_{\beta,0,p}(a_2)]$, then by following the proof of (a), we can show the existence of $h(\beta) \in (H'_{\beta,0,p}(a_1), H'_{\beta,0,p}(a_2))$ such that $(\beta, -h(\beta)) \in \mathscr{C}_p{}^+$. If we denote the unique h corresponding to each $\beta > \check{\beta}_p$ such that $(\beta, h) \in \mathscr{C}_p{}^+$ by $\varphi_p(\beta)$, then continuity and the strict decreasing nature of φ_p once again follow from the proof of (a).

Next, it follows from Remark B.1, that $\varphi_p(\tilde{\beta}_p) = 0$. We now show that $\lim_{\beta \to \infty} \varphi_p(\beta) = -\infty$. Towards this, note that the monotonicity pattern of $H'_{\beta,\varphi_p(\beta),p}$ for $\beta > \check{\beta}_p$ implies that $H_{\beta,\varphi_p(\beta),p}$ has exactly two local maximizers $m_1(\beta) \in (-1, a_1(\beta))$ and $m_2(\beta) \in (a_2(\beta), 1)$, where $a_1(\beta)$ and $a_2(\beta)$ are the inflection points of $H_{\beta,\varphi_p(\beta),p}$, satisfying $0 < a_1(\beta) < m_* < a_2(\beta) < 1$ for all $\beta > \check{\beta}_p$. Hence, $m_1(\beta)$ and $m_2(\beta)$ are global maximizers of $H_{\beta,\varphi_p(\beta),p}$. Let $\beta > \tilde{\beta}_p$, whence the strictly decreasing nature of φ_p implies that $\varphi_p(\beta) < 0$. Since $H'_{\beta,\varphi_p(\beta),p}(-1) = \infty$ and $H'_{\beta,\varphi_p(\beta),p}(0) = \varphi_p(\beta) < 0$, the intermediate value theorem implies that $m_1(\beta) < 0$. Hence,

$$\beta(m_1(\beta))^p - I(m_1(\beta)) < 0, \quad \text{that is,} \quad H_{\beta,\varphi_p(\beta),p}(m_1(\beta)) < \varphi_p(\beta) m_1(\beta).$$

Now, since

$$H_{\beta,\varphi_p(\beta),p}(m_1(\beta)) = H_{\beta,\varphi_p(\beta),p}(m_2(\beta)) = \beta(m_2(\beta))^p + \varphi_p(\beta) m_2(\beta) - I(m_2(\beta)),$$

we have $\beta(m_2(\beta))^p + \varphi_p(\beta) m_2(\beta) - I(m_2(\beta)) < \varphi_p(\beta) m_1(\beta)$. This implies,

$$-2\varphi_p(\beta) > \varphi_p(\beta)(m_1(\beta) - m_2(\beta)) > \beta(m_2(\beta))^p - I(m_2(\beta)) \geqslant \beta m_*^p - I(m_2(\beta)). \tag{B.11}$$

The proof of our claim now follows from (B.11) since $\lim_{\beta \to \infty} \beta m_*^p - I(m_2(\beta)) = \infty$. This completes the proof of part (2).

Finally, we prove that $\lim_{\beta \to \check{\beta}_p^+} \varphi_p(\beta) = \tanh^{-1}(m_*) - p\check{\beta}_p m_*^{p-1}$, where $m_* := \sqrt{1 - 2/p}$. Towards this, let $0 < \varepsilon < \tilde{\beta}_p - \check{\beta}_p$ be given, and take any

$$\beta \in \left(\check{\beta}_p, \check{\beta}_p + \frac{\varepsilon}{2p(p-1)}\right).$$

As before, let $0 < a_1 < a_2 < 1$ be the points such that $H''_{\beta,0,p} < 0$ on $[0, a_1) \bigcup (a_2, 1]$ and $H''_{\beta,0,p} > 0$ on (a_1, a_2). Since $H''_{\check{\beta}_p,0,p} \leqslant 0$ on $[0, 1]$, it follows that $H''_{\beta,0,p} \leqslant (\beta - \check{\beta}_p)p(p-1) < \varepsilon/2$ on $[0, 1]$. Hence, for every $h \in \mathbb{R}$, we have

$$H'_{\beta,h,p}(a_2) - H'_{\beta,h,p}(a_1) = \int_{a_1}^{a_2} H''_{\beta,0,p}(t)\mathrm{d}t \leqslant \varepsilon(a_2 - a_1)/2 < \varepsilon/2. \quad \text{(B.12)}$$

Since $H''_{\beta,0,p}(m_*) > 0$, we must have $m_* \in (a_1, a_2)$. If $m_1 < m_2$ are the two global maximizers of $H_{\beta,\varphi_p(\beta),p}$, then $m_1 \in (0, a_1)$ and $m_2 \in (a_2, 1)$. Since $H'_{\beta,\varphi_p(\beta),p}$ is strictly decreasing on each of the intervals $[0, a_1]$ and $[a_2, 1)$, we must have $H'_{\beta,\varphi_p(\beta),p}(a_1) < 0$ and $H'_{\beta,\varphi_p(\beta),p}(a_2) > 0$. Hence, there exists $a_3 \in (a_1, a_2)$ such that $H'_{\beta,\varphi_p(\beta),p}(a_3) = 0$. Now, since $H'_{\beta,\varphi_p(\beta),p}$ is increasing on $[a_1, a_2]$, we have from (B.12),

$$\left|H'_{\beta,\varphi_p(\beta),p}(a_3) - H'_{\beta,\varphi_p(\beta),p}(m_*)\right| \leqslant H'_{\beta,\varphi_p(\beta),p}(a_2) - H'_{\beta,\varphi_p(\beta),p}(a_1) < \varepsilon/2,$$

and hence, $\left|H'_{\beta,\varphi_p(\beta),p}(m_*)\right| = \left|\tanh^{-1}(m_*) - p\beta m_*^{p-1} - \varphi_p(\beta)\right| < \varepsilon/2$. Now, $\left|p\beta m_*^{p-1} - p\check{\beta}_p m_*^{p-1}\right| \leqslant p(\beta - \check{\beta}_p) < \varepsilon/2$. By triangle inequality, we thus have

$$\begin{aligned}\left|\tanh^{-1}(m_*) - p\check{\beta}_p m_*^{p-1} - \varphi_p(\beta)\right| &\leqslant \left|\tanh^{-1}(m_*) - p\beta m_*^{p-1} - \varphi_p(\beta)\right| \\ &\quad + \left|p\beta m_*^{p-1} - p\check{\beta}_p m_*^{p-1}\right| \\ &< \varepsilon. \end{aligned} \quad \text{(B.13)}$$

Our claim now follows from (B.13). The proof of (2) and Lemma B.3 is now complete. □

Now, we will prove some properties of the function H, when the underlying parameter (β, h) is perturbed to (β, h_N), where $(\beta, h_N) \to (\beta, h)$, as $N \to \infty$. Investigating the properties of the function $H_{\beta,h_N,p}$ is especially important, since our analysis hinges more upon these perturbed functions, rather than the original function $H_{\beta,h,p}$.

Lemma B.4. *Suppose that $(\beta, h_N) \in [0, \infty) \times \mathbb{R}$ is a sequence converging to a point $(\beta, h) \in [0, \infty) \times \mathbb{R}$. Then, we have the following:*

(1) *Suppose that (β, h) is a p-regular point, and let m_* be the global maximizer of $H_{\beta,h,p}$. Then, for any sequence $(\beta, h_N) \in [0, \infty) \times \mathbb{R}$ converging to (β, h), the function $H_{\beta,h_N,p}$ will have unique global maximizer $m_*(N)$ for all large N, and $m_*(N) \to m_*$ as $N \to \infty$.*
(2) *Let m be a local maximizer of the function $H_{\beta,h,p}$, where the point (β, h) is not p-special. Suppose that $(\beta, h_N) \in [0, \infty) \times \mathbb{R}$ is a sequence converging to (β, h). Then for all large N, the function $H_{\beta,h_N,p}$ will have a local maximizer $m(N)$, such that $m(N) \to m$ as $N \to \infty$. Further, if $A \subseteq [-1, 1]$ is a closed interval such that $m \in int(A)$ and $H_{\beta,h,p}(m) > H_{\beta,h,p}(x)$ for all $x \in A\backslash\{m\}$, then there exists $N_0 \geqslant 1$, such that for all $N \geqslant N_0$, we have $H_N(m(N)) > H_N(x)$ for all $x \in A\backslash\{m(N)\}$.*

Proof of (1). The set $\mathcal{R}_p$ of all p-regular points is an open subset of $[0, \infty) \times \mathbb{R}$. To see this, note that $\mathcal{R}_p^c$ is given by $\mathscr{C}_p \bigcup \{(\check{\beta}_p, \check{h}_p)\}$ if p is odd, and by $\mathscr{C}_p \bigcup \{(\check{\beta}_p, \check{h}_p), (\check{\beta}_p, -\check{h}_p)\}$ if p is even. By Lemma B.3, $\mathcal{R}_p^c$ is a closed set in either case. Hence, the function $H_{\beta,h_N,p}$ will have unique global maximizer $m_*(N)$ for all large N.

To show that $m_*(N) \to m_*$, let $\{N_k\}_{k \geqslant 1}$ be a subsequence of the natural numbers. Then, $\{N_k\}_{k \geqslant 1}$ will have a further subsequence $\{N_{k_\ell}\}_{\ell \geqslant 1}$, such that $m_*(N_{k_\ell})$ converges to some $m' \in [-1, 1]$. Since $H_{\beta_{N_{k_\ell}}, h_{N_{k_\ell}}, p}\left(m_*(N_{k_\ell})\right) \geqslant H_{\beta_{N_{k_\ell}}, h_{N_{k_\ell}}, p}(x)$ for all $x \in [-1, 1]$, by taking limit as $\ell \to \infty$ on both sides, we have $H_{\beta,h,p}(m') \geqslant H_{\beta,h,p}(x)$ for all $x \in [-1, 1]$, showing that m' is a global maximizer of $H_{\beta,h,p}$. Since m_* is the unique global maximizer of $H_{\beta,h,p}$, it follows that $m' = m_*$, completing the proof of (1). □

Proof of (2). Let us denote $H_{\beta,h,p}$ by H and $H_{\beta,h_N,p}$ by H_N. It is easy to show that there exists $M \geqslant 1$ odd, and points $-1 = a_0 < a_1 < \ldots < a_M = 1$, such that H' is strictly decreasing on $[a_{2i}, a_{2i+1}]$ and strictly increasing on $[a_{2i+1}, a_{2i+2}]$ for all $0 \leqslant i \leqslant \frac{M-1}{2}$. Hence, the local maximizer m of H lies in (a_{2i}, a_{2i+1}) for some $0 \leqslant i \leqslant \frac{M-1}{2}$. Since $H'(a_{2i}) > 0$ and $H'(a_{2i+1}) < 0$, we also have $H_N'(a_{2i}) > 0$ and $H_N'(a_{2i+1}) < 0$ for all large N, and hence H_N' has a root $m(N) \in (a_{2i}, a_{2i+1})$ for all large N.
Let us now show that $m(N) \to m$. Towards this, let $\{N_k\}_{k \geqslant 1}$ be a subsequence of the natural numbers, whence there is a further subsequence $\{N_{k_\ell}\}_{\ell \geqslant 1}$ of $\{N_k\}_{k \geqslant 1}$, such that $m(N_{k_\ell}) \to m'$ for some $m' \in [a_{2i}, a_{2i+1}]$. Since $H'_{N_{k_\ell}}(m(N_{k_\ell})) = 0$ for all $\ell \geqslant 1$, we have $H'(m') = 0$. But the strict decreasing nature of H' on $[a_{2i}, a_{2i+1}]$ implies that m is the only root of H' on this interval, and hence, $m' = m$. This shows that $m(N) \to m$.
Next, we show that $m(N)$ is a local maximizer of H_N for all N sufficiently large. For this, we prove something stronger than needed, because this will be useful in proving the last statement of (2). Since $H''(m) < 0$, there exists $\varepsilon > 0$ such that $[m-\varepsilon, m+\varepsilon] \subset (a_{2i}, a_{2i+1})$ and $H'' < 0$ on $[m - \varepsilon, m + \varepsilon]$. If $m_0 \in [m - \varepsilon, m + \varepsilon]$ is such that $H''(m_0) = \sup_{x \in [m-\varepsilon, m+\varepsilon]} H''(x) < 0$, then since H_N'' converges to H'' uniformly on $(-1, 1)$,

$$\sup_{x \in [m-\varepsilon, m+\varepsilon]} H_N''(x) < H''(m_0)/2 \quad \text{for all large } N.$$

In particular, since $m(N) \in [m - \varepsilon, m + \varepsilon]$ for all large N, we have $H_N''(m(N)) < 0$ for all large N, showing that $m(N)$ is a local maximizer of H_N for all large N. Also, since $H_N'(m(N)) = 0$ and $\sup_{x \in [m-\varepsilon, m+\varepsilon]} H_N''(x) < 0$ for all large N, we must have

$$H_N(m(N)) > H_N(x) \quad \text{for all } x \in [m - \varepsilon, m + \varepsilon] \backslash \{m(N)\}, \quad \text{for all large} N.$$

Finally, suppose that $A \subseteq [-1, 1]$ is a closed interval such that $m \in \text{int}(A)$ and $H(m) > H(x)$ for all $x \in A \backslash \{m\}$. By Lemma B.11, there exists $\varepsilon' > 0$ such that for all $0 < \delta \leqslant \varepsilon'$, $\sup_{x \in A \backslash (m-\delta, m+\delta)} H(x) = H(m \pm \delta)$. Let $\alpha = \min\{\varepsilon, \varepsilon'\}$. Then,

$$H_N(m(N)) > H_N(x) \quad \text{for all } x \in [m - \alpha, m + \alpha] \backslash \{m(N)\}, \quad \text{for all large } N, \tag{B.14}$$

and $\sup_{x \in A \backslash (m-\alpha, m+\alpha)} H(x) = H(m \pm \alpha) < H(m)$ (since $H'(m) = 0$ and $H'' < 0$ on $[m - \alpha, m + \alpha]$). Hence,

$$\sup_{x \in A \backslash (m-\alpha, m+\alpha)} H_N(x) < H_N(m(N)) \quad \text{for all large } N. \tag{B.15}$$

The proof of (2) now follows from (B.14) and (B.15), and the proof of Lemma B.4 is now complete. □

B.2. Other technical lemmas. In this section, we collect the proofs of the remaining technical lemmas, which are used in the proofs of the main results in various places. We start with a result that gives implicit expressions for the partial derivatives of any stationary point of $H_{\beta,h,p}$ with respect to β and h.

Lemma B.5. *Let $m = m(\beta, h, p)$ satisfy the implicit relation $H'_{\beta,h,p}(m) = 0$, and suppose that $H''_{\beta,h,p}(m) \neq 0$. Then, the partial derivatives of m with respect to β and h are given by:*

$$\frac{\partial m}{\partial \beta} = -\frac{pm^{p-1}}{H''_{\beta,h,p}(m)} \quad \text{and} \quad \frac{\partial m}{\partial h} = -\frac{1}{H''_{\beta,h,p}(m)}. \tag{B.16}$$

Moreover, $\left|\frac{\partial^2 m}{\partial \beta^2}\right| < \infty$ *and* $\left|\frac{\partial^2 m}{\partial h^2}\right| < \infty$, *if* $H''_{\beta,h,p}(m) \neq 0$.

Proof. Differentiating both sides of the identity $\beta p m^{p-1} + h - \tanh^{-1}(m) = 0$ with respect to β and h separately, we get the following two first order partial differential equations, respectively:

$$pm^{p-1} + \beta p(p-1)m^{p-2}\frac{\partial m}{\partial \beta} - \frac{1}{1-m^2}\frac{\partial m}{\partial \beta} = 0,$$

$$\text{that is,} \quad pm^{p-1} + H''_{\beta,h,p}(m)\frac{\partial m}{\partial \beta} = 0\,; \tag{B.17}$$

$$\beta p(p-1)m^{p-2}\frac{\partial m}{\partial h} + 1 - \frac{1}{1-m^2}$$

$$\frac{\partial m}{\partial h} = 0, \quad \text{that is,} \quad 1 + H''_{\beta,h,p}(m)\frac{\partial m}{\partial h} = 0\,; \tag{B.18}$$

The expressions in (B.16) follow from (B.17) and (B.18). Another implicit differentiation of (B.17) with respect to β and (B.18) with respect to h yields the following two second order partial differential equations, respectively:

$$2p(p-1)m^{p-2}\frac{\partial m}{\partial \beta} + H^{(3)}_{\beta,h,p}(m)\left(\frac{\partial m}{\partial \beta}\right)^2 + H''_{\beta,h,p}(m)\frac{\partial^2 m}{\partial \beta^2} = 0; \tag{B.19}$$

$$H^{(3)}_{\beta,h,p}(m)\left(\frac{\partial m}{\partial h}\right)^2 + H''_{\beta,h,p}(m)\frac{\partial^2 m}{\partial h^2} = 0; \tag{B.20}$$

The finiteness of the second order partial derivatives of m as long as $H''_{\beta,h,p}(m) \neq 0$, now follow from the fact that $H''_{\beta,h,p}(m)$ is the coefficient of $\frac{\partial^2 m}{\partial \beta^2}$ and $\frac{\partial^2 m}{\partial h^2}$ in the differential equations (B.19) and (B.20). □

We now derive some important properties of the function ζ defined in (3.10). The following lemma is used in the proof of Lemma 3.2.

Lemma B.6. *For any sequence $x \in (-1, 1)$ that is bounded away from both 1 and -1, we have*

$$\zeta(x) = \sqrt{\frac{2}{\pi N(1-x^2)}}\, e^{N H_N(x)}\left(1 + O(N^{-1})\right).$$

Proof. The proof of Lemma B.6 follows immediately from Lemma A.5. □

Now, we bound the derivative of the function ζ in a neighborhood of the point $m_*(N)$. This result appears in the proof of Lemma 3.2.

Lemma B.7. *For every $\alpha \geqslant 0$ and p-regular point (β, h), we have the following bound:*

$$\sup_{x \in A_{N,\alpha}} |\zeta'(x)| = \zeta(m_*(N)) O\left(N^{\frac{1}{2}+\alpha}\right),$$

where $m_*(N)$ is the global maximizer of H_N and $A_{N,\alpha} := \left(m_*(N) - N^{-\frac{1}{2}+\alpha}, m_*(N) + N^{-\frac{1}{2}+\alpha}\right)$.

Proof of Lemma B.7. We begin with the following lemma: □

Lemma B.8. *For any sequence $x \in (-1, 1)$ that is bounded away from both 1 and -1, we have*

$$\zeta'(x) = \zeta(x)\left(NH_N'(x) + \frac{x}{1-x^2} + O(N^{-1})\right).$$

Proof. By Lemma A.1 and (A.1), we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\binom{N}{N(1+x)/2} \\
&= \frac{N}{2}\binom{N}{N(1+x)/2}\left[\psi\left(1 - \frac{Nx}{2} + \frac{N}{2}\right) - \psi\left(1 + \frac{Nx}{2} + \frac{N}{2}\right)\right] \\
&= \frac{N}{2}\binom{N}{N(1+x)/2}\left(\log\left(\frac{N}{2}(1-x)\right) - \log\left(\frac{N}{2}(1+x)\right)\right. \\
&\quad \left. + \frac{1}{N(1-x)} - \frac{1}{N(1+x)} + O(N^{-2})\right) \\
&= \binom{N}{N(1+x)/2}\left[-N\tanh^{-1}(x) + \frac{x}{1-x^2} + O(N^{-1})\right].
\end{aligned} \tag{B.21}$$

We thus have by the product rule of differential calculus and (B.21),

$$\begin{aligned}
\zeta'(x) &= \zeta(x)(N\beta p x^{p-1} + Nh_N) + \exp\left\{N(\beta x^p + h_N x - \log 2)\right\}\frac{\mathrm{d}}{\mathrm{d}x}\binom{N}{N(1+x)/2} \\
&= \zeta(x)(N\beta p x^{p-1} + Nh_N) + \zeta(x)\left[-N\tanh^{-1}(x) + \frac{x}{1-x^2} + O(N^{-1})\right] \\
&= \zeta(x)\left(NH_N'(x) + \frac{x}{1-x^2} + O(N^{-1})\right),
\end{aligned}$$

completing the proof of Lemma B.8. □

Now, we proceed with the proof of Lemma B.7. First note that, since $H_N'(m_*(N)) = 0$, we have by the mean value theorem,

$$\sup_{x \in A_{N,\alpha}} \left|H_N'(x)\right| \leqslant \sup_{x \in A_{N,\alpha}} \left|x - m_*(N)\right| \sup_{x \in A_{N,\alpha}} |H_N''(x)| = O\left(N^{-\frac{1}{2}+\alpha}\right). \tag{B.22}$$

It follows from (B.22) and Lemma B.8 that

$$\sup_{x \in A_{N,\alpha}} |\zeta'(x)| \leqslant O\left(N^{\frac{1}{2}+\alpha}\right) \sup_{x \in A_{N,\alpha}} \zeta(x). \tag{B.23}$$

Now, Lemma B.6 implies that

$$\sup_{x \in A_{N,\alpha}} \zeta(x) \leqslant \left(1 + O(N^{-1})\right) \zeta(m_*(N)) \sup_{x \in A_{N,\alpha}} \sqrt{\frac{1 - m_*(N)^2}{1 - x^2}} = \zeta(m_*(N)) O(1). \tag{B.24}$$

Lemma B.7 now follows from (B.23) and (B.24). □

Lemma B.7 has an analogous version for p-special points (β, h), which is stated below. In this case, the bound on ζ' is better, and holds on a slightly larger region, too.

Lemma B.9. *Let $m_*(N)$ be the unique global maximizer of $H_N := H_{\beta,h_N,p}$, where $h_N := h + \bar{h} N^{-3/4}$ for some $\bar{h} \in \mathbb{R}$, and (β, h) is a p-special point. Then, for all $\alpha \geqslant 0$,*

$$\sup_{x \in \mathcal{A}_{N,\alpha}} |\zeta'(x)| = \zeta(m_*(N)) O\left(N^{\frac{1}{4}+3\alpha}\right)$$

where $\mathcal{A}_{N,\alpha} := \left(m_*(N) - N^{-\frac{1}{4}+\alpha}, m_*(N) + N^{-\frac{1}{4}+\alpha}\right)$.

Proof. The proof of Lemma B.9 is similar to that of Lemma B.7, the only difference being a change in the estimate of $\sup_{x \in \mathcal{A}_{N,\alpha}} |H_N'(x)|$ from the estimate in (B.22). Note that

$$\sup_{x \in \mathcal{A}_{N,\alpha}} |H_N''(x)| \leqslant \sup_{x \in \mathcal{A}_{N,\alpha}} \tfrac{1}{2}(x - m_*)^2 \sup_{x \in \mathcal{I}(\mathcal{A}_{N,\alpha} \cup \{m_*\})} H^{(4)}(x) = O\left(N^{-\frac{1}{2}+2\alpha}\right),$$

where m_* denotes the global maximizer of $H_{\beta,h,p}$ and for a set $A \subseteq \mathbb{R}$, $\mathcal{I}(A)$ denotes the smallest interval containing A. The last equality follows from the observation

$$\sup_{x \in \mathcal{A}_{N,\alpha}} |x - m_*| \leqslant \sup_{x \in \mathcal{A}_{N,\alpha}} |x - m_*(N)| + |m_*(N) - m_*| \leqslant N^{-\frac{1}{4}+\alpha} + O\left(N^{-\frac{1}{4}}\right)$$
$$= O\left(N^{-\frac{1}{4}+\alpha}\right),$$

by Lemma B.10. Following (B.22), we have

$$\sup_{x \in \mathcal{A}_{N,\alpha}} |H_N'(x)| = O\left(N^{-\frac{3}{4}+3\alpha}\right).$$

The rest of the proof is exactly same as that of Lemma B.7. □

The following lemma provides estimates of the first four derivatives of the function H at the maximizer $m_*(N)$ for a perturbation of a p-special point. This key result is used in the proof of Lemma 3.4.

Lemma B.10. *Let (β, h) be a p-special point and $h_N := h + \bar{h}N^{-\frac{3}{4}}$ for some $\bar{h} \in \mathbb{R}$. If m_* and $m_*(N)$ denote the unique global maximizers of $H := H_{\beta,h,p}$ and $H_N := H_{\beta,h_N,p}$ respectively, then we have the following:*

$$N^{\frac{1}{4}}(m_*(N) - m_*) = -\left(\frac{6\bar{h}}{H^{(4)}(m_*)}\right)^{\frac{1}{3}} + O\left(N^{-\frac{1}{4}}\right), \tag{B.25}$$

$$N^{\frac{1}{2}}H''(m_*(N)) = \frac{1}{2}\left(6\bar{h}\right)^{\frac{2}{3}}\left(H^{(4)}(m_*)\right)^{\frac{1}{3}} + O\left(N^{-\frac{1}{4}}\right), \tag{B.26}$$

$$N^{\frac{1}{4}}H^{(3)}(m_*(N)) = -\left(6\bar{h}\right)^{\frac{1}{3}}\left(H^{(4)}(m_*)\right)^{\frac{2}{3}} + O\left(N^{-\frac{1}{4}}\right), \tag{B.27}$$

$$H^{(4)}(m_*(N)) = H^{(4)}(m_*) + O\left(N^{-\frac{1}{4}}\right). \tag{B.28}$$

Proof. Let us start by noting that

$$H'(m_*(N)) = H'_N(m_*(N)) - \bar{h}N^{-\frac{3}{4}} = -\bar{h}N^{-\frac{3}{4}}.$$

On the other hand, by a Taylor expansion of H' around m_* and using the fact $H'(m_*) = H''(m_*) = H^{(3)}(m_*) = 0$ (see Lemma B.2), we have

$$H'(m_*(N)) = \tfrac{1}{6}(m_*(N) - m_*)^3 H^{(4)}(\zeta_N),$$

where ζ_N lies between $m_*(N)$ and m_*. Hence,

$$N^{\frac{3}{4}}(m_*(N) - m_*)^3 = -\frac{6\bar{h}}{H^{(4)}(\zeta_N)}.$$

Now, it follows from the proof of Lemma B.4, part (1), that $m_*(N) \to m_*$, and hence, $\zeta_N \to m_*$. This implies that

$$\lim_{N\to\infty} N^{\frac{1}{4}}(m_*(N) - m_*) = -\left(\frac{6\bar{h}}{H^{(4)}(m_*)}\right)^{\frac{1}{3}}. \tag{B.29}$$

By a 5-term Taylor expansion of $H'(m_*(N))$ around m_*, one obtains

$$\tfrac{1}{6}(m_*(N) - m_*)^3 H^{(4)}(m_*) + \tfrac{1}{24}(m_*(N) - m_*)^4 H^{(5)}(\zeta'_N) = -\bar{h}N^{-\frac{3}{4}}. \tag{B.30}$$

for some sequence ζ'_N lying between $m_*(N)$ and m_*. From (B.30) and (B.29), we have

$$\begin{aligned} N^{\frac{3}{4}}(m_*(N) - m_*)^3 &= -\frac{6\bar{h}}{H^{(4)}(m_*)} - \frac{N^{\frac{3}{4}}(m_*(N) - m_*)^4 H^{(5)}(\zeta'_N)}{4H^{(4)}(m_*)} \\ &= -\frac{6\bar{h}}{H^{(4)}(m_*)} + O\left(N^{-\frac{1}{4}}\right). \end{aligned} \tag{B.31}$$

(B.25) now follows from (B.31), and (B.26), (B.27), (B.28) follow by substituting (B.25) into the following expansions

$$H''(m_*(N)) = \tfrac{1}{2}\left(m_*(N) - m_*\right)^2 H^{(4)}(m_*) + O\left((m_*(N) - m_*)^3\right),$$

$$H^{(3)}(m_*(N)) = \left(m_*(N) - m_*\right) H^{(4)}(m_*) + O\left((m_*(N) - m_*)^2\right),$$

and $H^{(4)}(m_*(N)) = H^{(4)}(m_*) + O(m_*(N) - m_*)$. □

The final lemma shows that if a function has non-vanishing curvature at a unique point of maxima, then for every sufficiently small open interval I around that point of maxima, it attains its maximum on I^c at either of the endpoints of I. This fact is used in the proofs of Lemmas 3.1 and 3.7 .

Lemma B.11. *Let $A \subseteq [-1, 1]$ be a closed interval. Suppose that $f : A \mapsto \mathbb{R}$ is continuous on A and twice continuously differentiable on* $\text{int}(A)$. *Suppose that there exists $x_* \in \text{int}(A)$ such that $f(x_*) > f(x)$ for all $x \in A \backslash \{x_*\}$, and $f''(x_*) < 0$. Then, there exists $\eta > 0$ such that for all $0 < \varepsilon \leqslant \eta$, f attains maximum on the set $A \backslash (x_* - \varepsilon, x_* + \varepsilon)$ at either $x_* - \varepsilon$ or $x_* + \varepsilon$.*

Proof. Since f'' is continuous on $\text{int}(A)$ and negative at x_*, there exists $\delta > 0$ such that $f''(x) < 0$ for all $x \in (x_* - \delta, x_* + \delta)$. Hence, f' is strictly decreasing on $(x_* - \delta, x_* + \delta)$. Since $f'(x_*) = 0$, we have $f'(x) > 0$ for all $x \in (x_* - \delta, x_*)$ and $f'(x) < 0$ for all $x \in (x_*, x_* + \delta)$. Hence, f is strictly increasing on $(x_* - \delta, x_*]$ and strictly decreasing on $[x_*, x_* + \delta)$.
Suppose now, towards a contradiction, that the lemma is not true. Then, there is a sequence $\varepsilon_n \to 0$ such that neither $x_* - \varepsilon_n$ nor $x_* + \varepsilon_n$ is a point of maximum of f on $A \backslash (x_* - \varepsilon_n, x_* + \varepsilon_n)$. Let $x_n \in A \backslash [x_* - \varepsilon_n, x_* + \varepsilon_n]$ be such that $f(x_n) = \sup_{x \in A \backslash (x_* - \varepsilon_n, x_* + \varepsilon_n)} f(x)$, which exists by the continuity of f and compactness of the set $A \backslash (x_* - \varepsilon_n, x_* + \varepsilon_n)$. Since $f(x_* - \varepsilon_n) \leqslant f(x_n) \leqslant f(x_*)$ for all n, and f is continuous, it follows that $f(x_n) \to f(x_*)$. If x_{n_k} is a convergent subsequence of x_n converging to some $y \in A$, then by continuity of f, we have $f(y) = f(x_*)$. This implies that $y = x_*$. Therefore, there exists k such that $x_{n_k} \in (x_* - \delta, x_* + \delta) \backslash \{x_*\}$ and $\varepsilon_{n_k} < \delta$. For this k, we have $f(x_{n_k}) < \max\{f(x_* - \varepsilon_{n_k}), f(x_* + \varepsilon_{n_k})\}$. This contradicts the fact that x_{n_k} maximizes f on the set $A \backslash (x_* - \varepsilon_{n_k}, x_* + \varepsilon_{n_k})$, completing the proof of Lemma B.11. □

References

1. Banerjee, S., Carlin, B.P., Gelfand, A.E.: Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC, Boca Raton (2014)
2. Barra, A.: Notes on ferromagnetic p-spin and REM. Math. Methods Appl. Sci. **32**(7), 783–797 (2009)
3. Bovier, A., Gayrard, V.: The thermodynamics of the Curie–Weiss model with random couplings. J. Stat. Phys. **72**(3–4), 643–664 (1993)
4. Van Can, H.: Annealed limit theorems for the Ising model on random regular graphs. Ann. Appl. Probab. **29**(3), 1398–1445 (2019)
5. Chaganty, N.R., Sethuraman, J.: Large deviation local limit theorems for arbitrary sequences of random variables. Ann. Probab. **13**(1), 97–114 (1985)
6. Chatterjee, S., Shao, Q.M.: Non-normal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. Ann. Appl. Probab. **21**(2), 464–483 (2011)
7. Comets, F., Gidas, B.: Asymptotics of maximum likelihood estimators for the Curie–Weiss model. Ann. Stat. **19**(2), 557–578 (1991)
8. Daskalakis, C., Dikkala, N., Panageas, I.: Logistic regression with peer-group effects via inference in higher-order Ising models. arXiv:2003.08259 (2020)
9. Daskalakis, C., Mossel, E., Roch, S.: Evolutionary trees and the Ising model on the Bethe lattice: a proof of Steel's conjecture. Probab. Theory Relat. Fields **149**(1), 149–189 (2011)
10. Deb, N., Mukherjee, S.: Fluctuations in mean-field Ising models. arXiv:2005.00710 (2020)
11. Dembo, A., Montanari, A.: Gibbs measures and phase transitions on sparse random graphs. Braz. J. Probab. Stat. **24**(2), 137–211 (2010)
12. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Springer, Berlin (2010)
13. de Bruijn, N.G.: Asymptotic Methods in Analysis. Interscience Publishers Inc., New York (1957)
14. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Springer, Berlin (2007)
15. Ellis, R.S., Newman, C.M.: Limit theorems for sums of dependent random variables occurring in statistical mechanics. Z. Wahrscheinlichkeitstheorie verw. Geb. **44**, 117–139 (1978)

16. Ellis, R.S., Newman, C.M.: The statistics of Curie–Weiss models. J. Stat. Phys. **19**, 149–161 (1978)
17. Ellis, R.S., Newman, C.M., Rosen, J.S.: Limit theorems for sums of dependent random variables occurring in statistical mechanics II. Conditioning, multiple phases, and metastability. Z. Wahrscheinlichkeitstheorie verw. Geb. **51**, 153–169 (1980)
18. Geman, S., Graffigne, C.: Markov random field image models and their applications to computer vision. In: Proceedings of the International Congress of Mathematicians, pp. 1496–1517 (1986)
19. Giardiná, C., Giberti, C., van der Hofstad, R., Prioriello, M.L.: Ising critical behavior of inhomogeneous Curie–Weiss models and annealed random graphs. Commun. Math. Phys. **348**, 221–263 (2016)
20. Green, P.J., Richardson, S.: Hidden Markov models and disease mapping. J. Am. Stat. Assoc. **97**, 1055–1070 (2002)
21. Heringa, J.R., Blöte, H.W.J., Hoogland, A.: Phase transitions in self-dual Ising models with multispin interactions and a field. Phys. Rev. Lett. **63**, 1546–1549 (1989)
22. Hoffman, H.: On the continuity of the inverses of strictly monotonic functions. Ir. Math. Soc. Bull. **75**, 45–57 (2015)
23. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U.S.A. **79**, 2554–2558 (1982)
24. Ising, E.: Beitrag zur theorie der ferromagnetismus. Z. Phys. **31**, 253–258 (1925)
25. Kabluchko, Z., Löwe, M., Schubert, K.: Fluctuations of the magnetization for Ising models on dense Erdős–Rényi random graphs. J. Stat. Phys. **177**, 78–94 (2019)
26. Kabluchko, Z., Löwe, M., Schubert, K.: Fluctuations of the magnetization for Ising models on Erdős–Rény random graphs—the regimes of low temperature and external magnetic field. arXiv:2012.08204 (2021)
27. Levin, D.A., Luczak, M.J., Peres, Y.: Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. Probab. Theory Relat. Fields **146**(1–2), 223–265 (2010)
28. Montanari, A., Saberi, A.: The spread of innovations in social networks. Proc. Natl. Acad. Sci. **107**(47), 20196–20201 (2010)
29. Ohkuwa, M., Nishimori, H., Lidar, D.A.: Reverse annealing for the fully connected p-spin model. Phys. Rev. A **98**, 022314 (2018)
30. Suzuki, M.: Solution and critical behavior of some "three-dimensional" Ising models with a four-spin interaction. Phys. Rev. Lett. **28**, 507–510 (1972)
31. Suzuki, M., Fisher, M.E.: Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models. J. Math. Phys. **12**(2), 235–246 (1971)
32. Talagrand, M.: Spin Glasses: A Challenge for Mathematicians-Cavity and Mean Field Models. Springer, Berlin (2003)
33. Jörg, T., Krzakala, F., Kurchan, J., Maggs, A.C., Pujos, J.: Energy gaps in quantum first-order mean-field–like transitions: the problems that quantum annealing cannot solve. EPL (Europhys. Lett.) **89**(4), 40004 (2010)
34. Turban, L.: One-dimensional Ising model with multispin interactions. J. Phys. A Math. Theor. **49**(35), 355002 (2016)
35. Wong, R.: Asymptotic Approximations of Integrals. Society for Industrial and Applied Mathematic, Philadelphia (2001)
36. Yamashiro, Yu., Ohkuwa, M., Nishimori, H., Lidar, D.A.: Dynamics of reverse annealing for the fully-connected p-spin model. Phys. Rev. A **100**, 052321 (2019)

Communicated by S. Chatterjee