mathematics

MDPI

Article

Limit Theorems for the Non-Convex Multispecies Curie–Weiss Model

Francesco Camilli [1], Emanuele Mingione [2] and Godwin Osabutey [2,*,†]

[1] Quantitative Life Sciences, International Centre for Theoretical Physics, 34151 Trieste, Italy; fcamilli@ictp.it
[2] Department of Matematics, Alma Mater Studiorum, University of Bologna, 40126 Bologna, Italy; emanuele.mingione2@unibo.it
* Correspondence: gosabutey@unimore.it
† Current address: Department of Physics, Computer Science and Mathematics, University of Modena and Reggio Emilia, 41121 Modena, Italy.

Abstract: We study the thermodynamic properties of the generalized non-convex multi-species Curie–Weiss model, where interactions among different types of particles (forming the species) are encoded in a generic matrix. For spins with a generic prior distribution, we compute the thermodynamic limit of the generating functional for the moments of the Boltzmann–Gibbs measure using simple interpolation techniques. For Ising spins, we further analyze the fluctuations of the magnetization in the thermodynamic limit under the Boltzmann–Gibbs measure. It is shown that a central limit theorem (CLT) holds for a rescaled and centered vector of species magnetizations, which converges to either a centered or non-centered multivariate normal distribution, depending on the rate of convergence of the relative sizes of the species.

Keywords: non-convex Curie–Weiss model; central limit theorem; Ising model; multispecies mean-field model; arbitrary spin distribution

MSC: 82D40; 60F05; 82B26

check for updates

Academic Editor: Davide Valenti

Received: 6 March 2025
Revised: 17 April 2025
Accepted: 18 April 2025
Published: 19 April 2025

Citation: Camilli, F.; Mingione, E.; Osabutey, G. Limit Theorems for the Non-Convex Multispecies Curie–Weiss Model. *Mathematics* **2025**, *13*, 1343. https://doi.org/10.3390/math13081343

1. Introduction

The Curie–Weiss model, also known as the mean-field Ising model, is one of the simplest models of magnetism that exhibits phase transitions [1]. In this model, the spins assume discrete binary values (± 1) and interact uniformly with one another. Due to its simplicity and analytical tractability, it has been applied in a variety of fields, including voting dynamics [2–4] and social collective behavior [4–7]. A multispecies extension of the Curie–Weiss model [8–10] has been proposed to capture the large-scale behavior of interacting systems involving multiple types of interacting particles, whose strength depends on which species each particle belongs to. These extensions were originally introduced in statistical physics as approximations of lattice models [11,12] and meta-magnets [13,14], which exhibit both ferromagnetic and antiferromagnetic interactions. From a mathematical physics perspective, the free energy of a two-species ferromagnet was rigorously derived in [15] and further investigated in [16–18] with the Hamilton–Jacobi formalism. Beyond the computation of the free energy, the fluctuations of the order parameter of the multispecies Curie–Weiss model were initially studied in [19], under a convexity assumption on the Hamiltonian, using the same approach as [20]. More recently, stronger results have been proved using different approaches. Notably, refs. [21,22] showed the validity of central limit theorems (CLTs) and provided their convergence rate via Stein's method, while in [23–25] a moment generating functional approach was used. We stress that in all the aforementioned literature, the convexity assumption on the interaction matrix always plays a crucial role. A similar convexity condition appears in the context of disordered multispecies models, where a formula for the free energy is only known in the convex case [26,27], for spherical models [28,29], and on the Nishimori line [30,31]. The non-convex case remains an open problem (see [32,33] for partial results on the subject). Regarding models with random interactions, few fluctuation theorems for the order parameter are known: for the Sherrington–Kirkpatrick model, the validity of CLTs is limited to high-temperature regions [34–36], unless the model is on the Nishimori line [37], while the multispecies case was studied here [38].

In this paper, we introduce a multispecies Curie–Weiss model with non-convex energy arising from coupling constants, thereby allowing for arbitrary interaction matrices. This extension is motivated by the need to model systems in which interspecies interactions may be competitive—a scenario relevant not only in statistical physics but also in areas such as game theory [39], neural networks [40], and social science [41,42], where it may find applications. Furthermore, we allow for generic spin values within $[-1,1]$, extending the model beyond the classical Ising framework.

The aim of this work is to study the multispecies Curie–Weiss model in full generality, in particular, without any convexity assumption for the interaction matrix. The work is divided into two main parts. The first part examines the limiting free energy of the model for arbitrary spin distributions supported on $[-1,1]$. Using a combination of interpolation methods and decoupling techniques, we derive a variational formula for the free energy. The second part focuses on the asymptotic behavior of the vector of species magnetization in the case of Ising spins. The methods used in the latter are inspired by [43,44]. By generalizing these methods, we demonstrate that the rescaled vector of species magnetization follows a standard CLT in the region of the phase space, where the order parameter concentrates on a single value. On the other hand, when concentration occurs at multiple points, a conditional CLT still applies.

This paper is structured as follows: Section 2 gives a description for the generalized multispecies Curie–Weiss model with a generic coupling matrix. The main results are presented in Section 3, followed by detailed proofs in Section 4. Section 5 concludes the paper and discusses future directions. The Appendix A contains some technical results used in the proofs.

2. Model Description and Definitions

Let ρ be a probability measure supported on $[-1,1]$, and K an integer representing the number of different species. Consider a Hamiltonian system consisting of N interacting spins, each labeled by an integer, σ_i, that lies in the set of indices $\Lambda = \{1,2,\dots,N\}$. To introduce a multispecies structure, we divide the set of indices Λ into K disjoint subsets Λ_p of cardinality N_p for $p = 1,\dots,K$, namely

$$\Lambda_p \cap \Lambda_l = \emptyset \;\; \forall\, p \neq l, \quad \sum_{p=1}^{K} |\Lambda_p| = N_1 + \cdots + N_K = N\,. \tag{1}$$

For future convenience, we also introduce the *form factors*, or relative sizes ratios $(\alpha_p)_{p\le K} = (N_p/N)_{p\le K}$, that shall be collected into a diagonal $K\times K$ matrix $\boldsymbol{\alpha} = \mathrm{diag}(\alpha_p)_{p\le K}$. Naturally, one has $\sum_{p\le K}\alpha_p = 1$. In the following, we also allow these ratios to depend on N, $\alpha_{N,p}$, and we shall call their limit α_p:

$$\lim_{N\to\infty}\alpha_{N,p} = \lim_{N\to\infty}\frac{N_p}{N} = \alpha_p, \quad p\le K. \tag{2}$$

However, for simplicity, we adopt the notation α_p instead of $\alpha_{N,p}$.

Let us turn back to our Hamiltonian system. The configuration space is $[-1,1]^N = \Omega_N$. Now that the multispecies structure has been introduced, let the interaction be governed by an indefinite symmetric matrix $\mathbf{J} \in \mathbb{R}^{K\times K}$ and the species specific external field $\mathbf{h} \in \mathbb{R}^K$. A schematic representation of the interaction network is displayed in Figure 1.

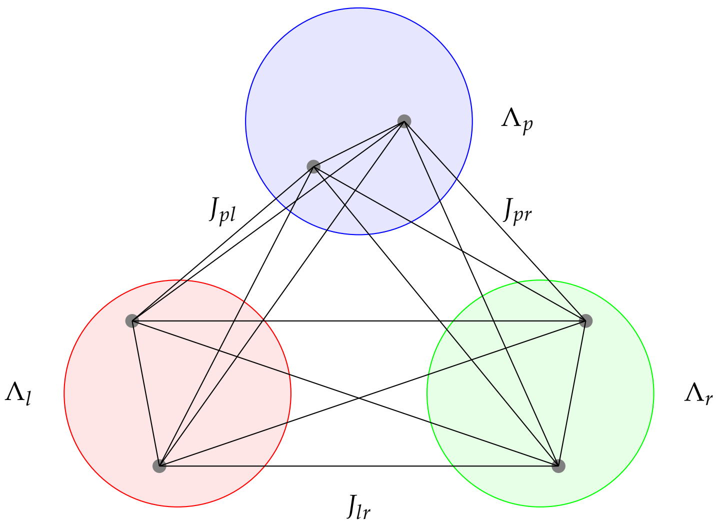

Figure 1. Interaction scheme for the multispecies model.

The model under study is thus defined as follows: Let $\sigma = (\sigma_i)_{i\in\Lambda} \in \Omega_N$, then for a given $(\mathbf{J}, \mathbf{h}) \in \mathbb{R}^{K\times K} \times \mathbb{R}^K$, the multispecies model is defined by the Hamiltonian:

$$H_N(\sigma) = -\frac{1}{2N}\sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j - \sum_{i=1}^{N} h_i\sigma_i = -\frac{N}{2}\sum_{p,l=1}^{K} m_p\alpha_p J_{pl}\alpha_l m_l - N\sum_{p=1}^{K}\alpha_p h_p m_p \tag{3}$$

where for each species p the magnetization density is

$$m_p(\sigma) = \frac{1}{N_p}\sum_{i\in\Lambda_p}\sigma_i \tag{4}$$

and we denote by $\mathbf{m}_N = (m_p)_{p\le K} \in [-1,1]^K$ the magnetization vector. Here, J_{pl} denotes the coupling strength between all species of types p and l, while h_p is the external field acting on all species of type p. The Hamiltonian (3) can now be rewritten as

$$H_N = -\frac{N}{2}(\mathbf{m}_N, \Delta\mathbf{m}_N) - N(\tilde{\mathbf{h}}, \mathbf{m}_N) \tag{5}$$

where $(\cdot,\cdot)$ denotes the scalar product in $\mathbb{R}^K$ and

$$\Delta = \boldsymbol{\alpha}\mathbf{J}\boldsymbol{\alpha} \qquad \text{and} \qquad \tilde{\mathbf{h}} = \boldsymbol{\alpha}\mathbf{h}. \tag{6}$$

The joint distribution of σ is governed by a Boltzmann–Gibbs measure

$$\mathcal{G}_N(\sigma) := \frac{e^{-H_N(\sigma)}}{Z_N}\prod_{i=1}^{N} d\rho(\sigma_i), \tag{7}$$

where $Z_N = \int_{[-1,1]^N} e^{-H_N(\sigma)}\prod_{i=1}^{N} d\rho(\sigma_i)$ is the partition function. Note that, in Equation (7), the usual inverse temperature $\tilde{\beta}$ is absorbed into the model parameters $\mathbf{J}$ and $\mathbf{h}$. Averages with respect to $\mathcal{G}_N$ will be denoted by $\omega_N(\cdot)$. The generating function for the moments of the measure in (7) is given by

$$p_N = \frac{1}{N} \log Z_N. \tag{8}$$

Note that p_N coincides, up to a multiplicative factor, with the free energy of the model.

3. Main Results

In this section, we provide the variational formula for the large N limit of the generating functional (8) of the generalized multispecies Curie–Weiss model (5). Additionally, in the case of Ising spins, we present central limit theorems for the magnetization vector $\mathbf{m}_N$. We reiterate that the coupling matrix $\mathbf{J}$ used throughout this work is an arbitrary symmetric real matrix.

3.1. Thermodynamic Limit of the Generating Functional

Our first result expresses the large N limit of p_N (8) as a variational problem in $\mathbb{R}^K$. Let us define the following variational functional:

$$p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \equiv p_{var}(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}, \Delta\mathbf{x}) + \sum_{p=1}^{K} \alpha_p \log \int_{\mathbb{R}} d\rho(\sigma) \exp\left[\sigma\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right]. \tag{9}$$

Now, let O be the orthogonal matrix diagonalizing Δ. Then, the following theorem holds:

Theorem 1. *For any $(\mathbf{J}, \mathbf{h}) \in \mathbb{R}^{K\times K} \times \mathbb{R}^K$, we have that*

$$\lim_{N\to\infty} p_N = \inf_{z_1,\dots,z_a} \sup_{z_{a+1},\dots,z_K} p_{var}(O\mathbf{z}). \tag{10}$$

It is worth noting that the large N limit of p_N can also be derived using large deviation techniques [45,46]. However, in this work, we adopt interpolation bounds. For further discussion, we refer interested readers to [47], where the phase diagram of the Curie–Weiss model with ferromagnetic interaction and generic compact spin distribution is analyzed.

3.2. Fluctuations of the Magnetization

Here, we state fluctuation results for the species magnetization in the binary spin case, i.e., $\sigma = (\sigma_i)_{i\le N} \in \{-1,1\}^N$. Let us define the function

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, \Delta\mathbf{x}) + (\tilde{\mathbf{h}}, \mathbf{x}) - (\hat{\boldsymbol{\alpha}}, I(\mathbf{x})), \quad \mathbf{x} \in [-1,1]^K \tag{11}$$

where $I(\mathbf{x}) = (I(x_p))_{p\le K} \in \mathbb{R}^K$ and

$$I(x) = \frac{1-x}{2}\log\left(\frac{1-x}{2}\right) + \frac{1+x}{2}\log\left(\frac{1+x}{2}\right), \quad x \in [-1,1] \tag{12}$$

and $\hat{\boldsymbol{\alpha}}$ is the vector associated with the diagonal matrix $\boldsymbol{\alpha}$.

In what follows, we assume that the convergence in (2) is sufficiently fast. More precisely, setting $\boldsymbol{\alpha}_N = \mathrm{diag}(\alpha_{N,p})_{p\le K}$, we assume that

$$\boldsymbol{\alpha}_N = \boldsymbol{\alpha} + N^{-\theta}\mathrm{diag}(\boldsymbol{\beta}) \tag{13}$$

where $\boldsymbol{\beta} = (\beta_p)_{p\le K}$ with $0 < \beta_p < \infty$ and $\theta \in \left[\frac{1}{2}, \infty\right)$. The fluctuations of the magnetization vector $\mathbf{m}_N$ depend on the global maximum point(s) of the function f in (11).

A detailed analysis of the critical points and the maximizers can be found in [15,21,48] for $K = 2$. We also note that for general K and zero external fields, a high temperature condition [22,25] implies that the zero vector is the unique global maximum of f, meaning the system exhibits no spontaneous magnetization.

The following theorems present central limit theorems for the vector of global species magnetization $\mathbf{m}_N = (m_1, ..., m_K)$ under the measure $\mathcal{G}_N$ (7). We begin with the case of a unique, non-degenerate global maximum.

Theorem 2. *Assume that* $(\mathbf{J}, \mathbf{h}) \in \mathbb{R}^{K\times K} \times \mathbb{R}^K$ *are such that* f *in* (11) *has a unique global maximizer* $\boldsymbol{\mu} = (\mu_r)_{r\le K}$ *with Hessian* $\mathcal{H}_f(\boldsymbol{\mu}) \prec 0$. *Then, under the measure* $\mathcal{G}_N$, *the following convergence in distribution holds:*

$$\left(\sqrt{N}\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N - \boldsymbol{\mu})\right) \xrightarrow[N\to\infty]{\mathcal{D}} \begin{cases} \mathcal{N}(\boldsymbol{\nu}, \sqrt{\boldsymbol{\alpha}}\mathcal{H}_f^{-1}(\boldsymbol{\mu})\sqrt{\boldsymbol{\alpha}}) & \text{if } \theta = \frac{1}{2} \\ \mathcal{N}(\mathbf{0}, \sqrt{\boldsymbol{\alpha}}\mathcal{H}_f^{-1}(\boldsymbol{\mu})\sqrt{\boldsymbol{\alpha}}) & \text{if } \theta > \frac{1}{2} \end{cases} \tag{14}$$

where $\boldsymbol{\nu} = -\sqrt{\boldsymbol{\alpha}}\, \mathcal{H}_f^{-1}(\boldsymbol{\mu})\, \boldsymbol{\alpha}\, \mathbf{J}\, diag(\boldsymbol{\mu})\boldsymbol{\beta}$.

When multiple global maxima coexist, Gaussianity can be restored under suitable conditions:

Theorem 3. *Assume that* $(\mathbf{J}, \mathbf{h}) \in \mathbb{R}^{K\times K} \times \mathbb{R}^K$ *are such that* f *in* (11) *has* n *global maximizers* $\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^n$, *each with Hessian* $\mathcal{H}_f(\boldsymbol{\mu}^i) \prec 0$ *for all* $i = 1, \dots, n$. *For any collection* $(A_i)_{i\le n}$ *of subsets of* $[-1,1]^K$ *such that* $\boldsymbol{\mu}^i \in int(A_i)$ *and* $f(\boldsymbol{\mu}^i) > f(\mathbf{x})$ *for all* $\mathbf{x} \in cl(A_i)\backslash\{\boldsymbol{\mu}^i\}$, *then*

$$\left(\sqrt{N}\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N - \boldsymbol{\mu}^i)\right)\Big|\{\mathbf{m}_N \in A_i\} \xrightarrow[N\to\infty]{\mathcal{D}} \begin{cases} \mathcal{N}(\boldsymbol{\nu}^i, \sqrt{\boldsymbol{\alpha}}\mathcal{H}_f^{-1}(\boldsymbol{\mu}^i)\sqrt{\boldsymbol{\alpha}}) & \text{if } \theta = \frac{1}{2} \\ \mathcal{N}(\mathbf{0}, \sqrt{\boldsymbol{\alpha}}\mathcal{H}_f^{-1}(\boldsymbol{\mu}^i)\sqrt{\boldsymbol{\alpha}}) & \text{if } \theta > \frac{1}{2} \end{cases} \tag{15}$$

and $\boldsymbol{\nu}^i = -\sqrt{\boldsymbol{\alpha}}\, \mathcal{H}_f^{-1}(\boldsymbol{\mu}^i)\, \boldsymbol{\alpha}\, \mathbf{J}\, diag(\boldsymbol{\mu}^i)\boldsymbol{\beta}$ *for all* $i = 1, \dots, n$.

4. Proofs

4.1. Proof of Theorem 1

We begin by noting that any indefinite matrix can be decomposed as the sum of two semi-definite matrices:

$$\Delta = \Delta_+ + \Delta_- \quad \text{where } \Delta_+ \ge 0,\ \Delta_- \le 0. \tag{16}$$

Thus, the scalar products seen so far split into two terms with definite signs. For instance,

$$(\mathbf{m}_N, \Delta\mathbf{m}_N) = (\mathbf{m}_N, \Delta_+\mathbf{m}_N) + (\mathbf{m}_N, \Delta_-\mathbf{m}_N). \tag{17}$$

Next, we order the species and eigenvectors so that the diagonal components satisfy

$$\Delta^D = \text{diag}(\lambda_1, \dots, \lambda_a, \lambda_{a+1}, \dots, \lambda_K), \quad \text{where } \lambda_1, \dots, \lambda_a \le 0, \quad \lambda_{a+1}, \dots, \lambda_K \ge 0, \tag{18}$$

$$\Delta^D = \Delta^D_+ + \Delta^D_- = \text{diag}(\lambda_1, \dots, \lambda_a, 0, \dots, 0) + \text{diag}(0, \dots, 0, \lambda_{a+1}, \dots, \lambda_K). \tag{19}$$

Here, the superscript D indicates a diagonal matrix. Additionally, note that the orthogonal matrix O simultaneously diagonalizes Δ, Δ_+, and Δ_-. We now establish the following key result:

Proposition 1 (Hybrid Sum Rule)**.** *For every N and $\mathbf{x} \in \mathbb{R}^K$, the following holds:*

$$p_N = -\frac{1}{2}(\mathbf{x}, \Delta_- \mathbf{x}) + \frac{1}{N} \log \int_{\mathbb{R}^K} d\rho(\sigma) e^{\frac{\beta N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) + \beta N(\tilde{\mathbf{h}} + \Delta_- \mathbf{x}, \mathbf{m}_N)} + \frac{1}{2}\int_0^1 dt\, \omega_{N,t}\Big[(\mathbf{m}_N - \mathbf{x}, \Delta_-(\mathbf{m}_N - \mathbf{x}))\Big], \quad (20)$$

where $\omega_{N,t}$ is the Gibbs state induced by a suitable interpolating Hamiltonian.

Proof. Consider the following interpolating Hamiltonian:

$$H_N(t) = -\frac{N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) - \frac{Nt}{2}(\mathbf{m}_N, \Delta_- \mathbf{m}_N) - (1-t)(\Delta_- \mathbf{x}, \mathbf{m}_N) - N(\tilde{\mathbf{h}}, \mathbf{m}_N). \quad (21)$$

The associated Boltzmann–Gibbs average is $\omega_{N,t}$, and the interpolating generating functional $p_N(t)$ satisfies

$$p_N(1) = p_N \quad (22)$$

$$p_N(0) = \frac{1}{N} \log \int_{\mathbb{R}^K} d\rho(\sigma) e^{\frac{\beta N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) + \beta N(\tilde{\mathbf{h}} + \Delta_- \mathbf{x}, \mathbf{m}_N)}. \quad (23)$$

Differentiating and completing the square yields

$$p_N'(t) = -\frac{1}{N}\omega_{N,t}(H_N(t)) = -\frac{1}{2}(\mathbf{x}, \Delta_- \mathbf{x}) + \frac{1}{2}\omega_{N,t}\Big[(\mathbf{m}_N - \mathbf{x}, \Delta_-(\mathbf{m}_N - \mathbf{x}))\Big].$$

The result follows by an application of the fundamental theorem of calculus. □

Lemma 1. *Assume the eigenvalue ordering in (19). Then, for any $(\mathbf{J}, \boldsymbol{h}) \in \mathbb{R}^{K\times K} \times \mathbb{R}^K$, one has*

$$p_N(\Delta, \mathbf{h}) \le \mathcal{O}\left(\frac{\log N}{N}\right) + \inf_{z_1 \dots z_a} \sup_{z_{a+1} \dots z_K} p_{var}(\Delta, \mathbf{h}; O\mathbf{z}). \quad (24)$$

Proof. We define the grid of hypercubes as $A_k^p = \left[y_k^p, y_k^p + \frac{2}{N_p}\right]$, where $y_k^p = -1 + 2(k-1)/N$ for $k \in \Lambda$, and p is the species label. The vertices of this grid are identified by a multi-index $\gamma = (\gamma_1, \dots, \gamma_K) \in \Lambda^K$, and and we denote them as $\mathbf{y}_\gamma$ in the following.

Next, observe that

$$\begin{aligned}
&\frac{1}{N}\log \int_{\mathbb{R}^N} d\rho(\sigma) e^{\frac{N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) + N(\tilde{\mathbf{h}} + \Delta_- \mathbf{x}, \mathbf{m}_N)} \\
&\quad = \frac{1}{N}\log \int_{\mathbb{R}^N} d\rho(\sigma) \prod_{p=1}^K \sum_{k=1}^{N_p} \mathbb{1}(m_p \in A_k^p) e^{\frac{N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) + N(\tilde{\mathbf{h}} + \Delta_- \mathbf{x}, \mathbf{m}_N)} \\
&\quad = \frac{1}{N}\log \sum_{\gamma_1, \dots, \gamma_K = 1}^{N_1, \dots, N_K} \int_{\mathbb{R}^N} d\rho(\sigma) \prod_{p=1}^K \mathbb{1}(m_p \in A_{\gamma_p}^p) e^{\frac{N}{2}(\mathbf{m}_N, \Delta_+ \mathbf{m}_N) + N(\tilde{\mathbf{h}} + \Delta_- \mathbf{x}, \mathbf{m}_N)}.
\end{aligned} \quad (25)$$

Enforcing the constraint $\prod_{p=1}^K \mathbb{1}(m_p \in A_{\gamma_p}^p)$, we obtain

$$0 \le (\mathbf{m}_N - \mathbf{y}_\gamma, \Delta_+(\mathbf{m}_N - \mathbf{y}_\gamma)) \le \frac{C}{N^2} \quad (26)$$

where $C > 0$, and the multi-index $\gamma = (\gamma_1, \dots, \gamma_K)$. Using this bound, we obtain

$$
\begin{aligned}
&\frac{1}{N}\log\int_{\mathbb{R}^N} d\rho(\sigma) e^{\frac{N}{2}(\mathbf{m}_N,\Delta_+\mathbf{m}_N)+N(\tilde{\mathbf{h}}+\Delta_-\mathbf{x},\mathbf{m}_N)} \leq \frac{C}{N^2} \\
&\quad+\frac{1}{N}\log\sum_{\gamma_1,\dots,\gamma_K=1}^{N_1,\dots,N_K}\int_{\mathbb{R}^N} d\rho(\sigma)\prod_{p=1}^{K}\mathbb{1}(m_p\in A^p_{\gamma_p})e^{-\frac{N}{2}(\mathbf{y}_\gamma,\Delta_+\mathbf{y}_\gamma)+N(\mathbf{m}_N,\Delta_+\mathbf{y}_\gamma)+N(\tilde{\mathbf{h}}+\Delta_-\mathbf{x},\mathbf{m}_N)} \\
&\quad\leq \frac{C}{N^2}+\frac{1}{N}\log\sum_{\gamma_1,\dots,\gamma_K=1}^{N_1,\dots,N_K}\int_{\mathbb{R}^N} d\rho(\sigma)e^{-\frac{N}{2}(\mathbf{y}_\gamma,\Delta_+\mathbf{y}_\gamma)+N(\mathbf{m}_N,\Delta_+\mathbf{y}_\gamma)+N(\tilde{\mathbf{h}}+\Delta_-\mathbf{x},\mathbf{m}_N)} \\
&\quad=\mathcal{O}\left(\frac{\log N}{N}\right)+\frac{1}{N}\log\sup_{\mathbf{y}}\int_{\mathbb{R}^N} d\rho(\sigma)e^{-\frac{N}{2}(\mathbf{y},\Delta_+\mathbf{y})+N(\tilde{\mathbf{h}}+\Delta_-\mathbf{x}+\Delta_+\mathbf{y},\mathbf{m}_N)} \\
&\quad=\mathcal{O}\left(\frac{\log N}{N}\right)+\frac{1}{N}\log\sup_{\mathbf{y}} e^{-\frac{N}{2}(\mathbf{y},\Delta_+\mathbf{y})}\int_{\mathbb{R}^N} d\rho(\sigma)e^{N\sum_{p=1}^K\alpha_p\frac{1}{N_p}\sum_{i\in\Lambda_p}\sigma_i(h_p+\hat{\boldsymbol{\alpha}}^{-1}(\Delta_-\mathbf{x}+\Delta_+\mathbf{y})_p)} \\
&\quad=\sup_{\mathbf{y}}\left\{-\frac{(\mathbf{y},\Delta_+\mathbf{y})}{2}+\sum_{p=1}^{K}\alpha_p\log\int_{\mathbb{R}} d\rho(\sigma)\exp\{\sigma[\mathbf{h}+\hat{\boldsymbol{\alpha}}^{-1}(\Delta_-\mathbf{x}+\Delta_+\mathbf{y})]_p\}\right\}+\mathcal{O}\left(\frac{\log N}{N}\right).
\end{aligned} \tag{27}
$$

Due to the decomposition of Δ, certain combinations of the components of $\mathbf{x}$ and $\mathbf{y}$ do not appear, i.e., those corresponding to the zero eigenvalues of Δ_+ and Δ_-. Therefore, we have

$$\Delta_+\mathbf{y}+\Delta_-\mathbf{x} = O\Delta_+^D\mathbf{y}^D+O\Delta_-^D\mathbf{x}^D = O\Delta^D\mathbf{z}^D = \Delta\mathbf{z} \tag{28}$$

$$\text{where}\quad \mathbf{z}^D = (x_1^D,\dots,x_a^D,y_{a+1}^D,\dots,y_K^D) \tag{29}$$

and D, when accompanying vectors, indicates vectors read in the basis of the eigenvectors of Δ (i.e., they denote vectors expressed in the eigenbasis of Δ). Hence, only K degrees of freedom remain. Now, by Equation (23), we have

$$
\begin{aligned}
p_N(0) &\leq \mathcal{O}\left(\frac{\log N}{N}\right)+\sup_{y^D_{a+1},\dots,y^D_K} p_{var}\left(\Delta_+,\hat{\boldsymbol{\alpha}}^{-1}\Delta_-\mathbf{x}+\mathbf{h};O\mathbf{y}^D\right) \\
&= \mathcal{O}\left(\frac{\log N}{N}\right)+\sup_{y^D_{a+1},\dots,y^D_K}\left\{-\frac{1}{2}\left(O\mathbf{y}^D,\Delta_+O\mathbf{y}^D\right)+\sum_{p=1}^{K}\alpha_p\log\int_{\mathbb{R}} d\rho(\sigma)e^{\sigma(\hat{\boldsymbol{\alpha}}^{-1}\Delta O\mathbf{z}^D+\mathbf{h})_p}\right\}.
\end{aligned} \tag{30}
$$

Note that this holds for any $x_1^D,\dots,x_a^D$ corresponding to the negative definite matrix. Adding the missing quadratic term (Δ_-) from the derivative and optimize over the remaining degrees of freedom, we obtain

$$p_N \leq \mathcal{O}\left(\frac{\log N}{N}\right)+\inf_{z_1,\dots,z_a}\sup_{z_{a+1},\dots,z_K} p_{var}(O\mathbf{z}).$$

We omitted D since $\mathbf{z}$ is now a dummy variable. □

Now we need to find the other bound.

Lemma 2. *For any* $(\mathbf{J},\boldsymbol{h})\in\mathbb{R}^{K\times K}\times\mathbb{R}^K$*, we have*

$$p_N \geq \inf_{z_1,\dots,z_a}\sup_{z_{a+1},\dots,z_K} p_{var}(O\mathbf{z})+\mathcal{O}\left(\frac{1}{N}\right). \tag{31}$$

Proof. For any $\mathbf{z} \in \mathbb{R}^K$, consider the interpolating Hamiltonian:

$$H_N(t) = -\frac{N}{2}(\mathbf{m}_N, \Delta \mathbf{m}_N) + \frac{Nt}{2}(\mathbf{m}_N - \mathbf{z}, \Delta(\mathbf{m}_N - \mathbf{z})) - N(\tilde{\mathbf{h}}, \mathbf{m}_N)\,. \tag{32}$$

The corresponding interpolating generating functional $p_N(t)$ satisfies

$$p_N(0) = p_N \tag{33}$$
$$p_N(1) = p_{var}(\Delta, \mathbf{h}; \mathbf{z}) \tag{34}$$
$$p'_N(t) = -\frac{1}{2}\omega_{N,t}[(\mathbf{m}_N - \mathbf{z}, \Delta(\mathbf{m}_N - \mathbf{z}))]\,. \tag{35}$$

By the convexity of the interpolating generating functional $p_N(t)$, $p_N(1) \leq p_N(0) + p'_N(0)$. Thus, for every $\mathbf{z} \in \mathbb{R}^K$, we obtain

$$p_N \geq \frac{1}{2}\omega_{N,0}[(\mathbf{m}_N - \mathbf{z}, \Delta(\mathbf{m}_N - \mathbf{z}))] + p_{var}(\Delta, \mathbf{h}; \mathbf{z}). \tag{36}$$

Now suppose that $\mathbf{z}$ equals a critical point $\bar{\mathbf{z}}^D = O^{-1}\bar{\mathbf{z}}$ of p_{var}, then

$$\frac{\partial}{\partial \mathbf{z}^D} p_{var}(O\bar{\mathbf{z}}^D) = O\frac{\partial}{\partial \mathbf{z}} p_{var}(\bar{\mathbf{z}}) = O\left[-\Delta\bar{\mathbf{z}} + \frac{\Delta \int_{\mathbb{R}} \sigma \exp \sigma(\boldsymbol{\alpha}^{-1}\Delta\bar{\mathbf{z}} + \mathbf{h})d\rho(\sigma)}{\int_{\mathbb{R}} \exp \tau(\boldsymbol{\alpha}^{-1}\Delta\bar{\mathbf{z}} + \mathbf{h})d\rho(\tau)}\right] = 0 \tag{37}$$

where the above exponentials and division are applied component-wise. Criticality implies from (37) that

$$\bar{\mathbf{z}} - \frac{\int_{\mathbb{R}} \sigma \exp \sigma(\boldsymbol{\alpha}^{-1}\Delta\bar{\mathbf{z}} + \mathbf{h})d\rho(\sigma)}{\int_{\mathbb{R}} \exp \tau(\boldsymbol{\alpha}^{-1}\Delta\bar{\mathbf{z}} + \mathbf{h})d\rho(\tau)} \in \mathrm{Ker}(\Delta). \tag{38}$$

Observe that the above can also be recast as

$$\bar{\mathbf{z}} - \omega_{N,0}(\mathbf{m}_N) \in \mathrm{Ker}(\Delta).$$

Expanding the quadratic form and separating diagonal and off-diagonal terms, we obtain

$$\begin{aligned}\omega_{N,0}[(\mathbf{m}_N - \bar{\mathbf{z}}, \Delta(\mathbf{m}_N - \bar{\mathbf{z}}))] = \sum_{p\neq l,1}^{K} \Delta_{pl}\omega_{N,0}[(m_p - \bar{z}_p)]\omega_{N,0}[(m_{N_l} - \bar{z}_l)] \\ + \sum_{p=1}^{K} \Delta_{pp}\omega_{N,0}[(m_p - \bar{z}_p)^2]\,.\end{aligned} \tag{39}$$

Now, observe that

$$\begin{aligned}\omega_{N,0}[(m_p - \bar{z}_p)^2] &= \omega_{N,0}[(m_p)^2] - 2\omega_{N,0}[m_p]\bar{z}_p + (\bar{z}_p)^2 \\ &= \frac{1}{N_p^2}\left(\sum_{i\in\Lambda_p} \omega_{N,0}[\sigma_i^2] + \sum_{i\neq j}\omega_{N,0}[\sigma_i\sigma_j]\right) - 2\omega_{N,0}[m_p]\bar{z}_p + (\bar{z}_p)^2\,.\end{aligned} \tag{40}$$

For $i \neq j$, the measure $\omega_{N,0}$ factorizes, and naturally $\sum_{i\in\Lambda_p} \sigma_i^2 \leq N_p$. Therefore,

$$\omega_{N,0}[(m_p - \bar{z}_p)^2] = \omega_{N,0}^2[(m_p - \bar{z}_p)] + \mathcal{O}(N^{-1}). \tag{41}$$

This finally implies

$$\omega_{N,0}\Big[(\mathbf{m}_N - \bar{\mathbf{z}}, \Delta(\mathbf{m}_N - \bar{\mathbf{z}}))\Big] = \mathcal{O}\left(\frac{1}{N}\right) + (\omega_{N,0}(\mathbf{m}_N - \bar{\mathbf{z}}), \Delta\omega_{N,0}(\mathbf{m}_N - \bar{\mathbf{z}}))\,. \tag{42}$$

The last term vanishes because $\omega_{N,0}(\mathbf{m}_N - \bar{\mathbf{z}}) \in \mathrm{Ker}(\Delta)$. □

The proof of Theorem 1 thus follows from Lemma 1 and Lemma 2.

4.2. Proof of Theorem 2

The proof of the CLT in the case of Ising spins, where $\rho = \frac{1}{2}(\delta_{-1} + \delta_{+1})$, is based on careful control of the asymptotic expansion of the partition function Z_N, following the methods of [43,44]. For any integer N and $x \in [-1, 1]$, we define the quantity

$$A_N(x) = \mathrm{card}\Big\{\sigma \in \{-1,1\}^N : \frac{1}{N}\sum_{i=1}^{N}\sigma_i = x\Big\} = \binom{N}{\frac{N(1+x)}{2}}. \tag{43}$$

We state the following useful standard bounds on A_N, whose proof can found in [36].

Lemma 3. *For any $x \in [-1, 1]$, the following inequality holds:*

$$\frac{1}{C\sqrt{N}} e^{-NI(x)} \le A_N(x) \le e^{-NI(x)} \tag{44}$$

where C is a universal constant and,

$$I(x) = \frac{1-x}{2}\log\left(\frac{1-x}{2}\right) + \frac{1+x}{2}\log\left(\frac{1+x}{2}\right). \tag{45}$$

Moreover, for any $x \in (-1, 1)$, one has

$$A_N(x) = \sqrt{\frac{2}{\pi N(1-x^2)}}\exp\left(-NI(x)\right)\cdot\Big(1 + \mathcal{O}\Big(N^{-1}\Big)\Big). \tag{46}$$

We begin by dividing the configuration space $\{-1, 1\}^N$ into microstates of equal local magnetization. For a given species $l \le K$ with spins configuration $\sigma^{(l)}$, the local magnetization m_l takes values in $S_l = \{-1 + \frac{2n}{N_l}, n = 0, \ldots, N_l\}$, with $|S_l| = (N_l + 1)$. The magnetization vector $\mathbf{m}_N = (m_l)_{l\le K}$ thus takes values in $S_N = \bigtimes_{l=1}^{K} S_l$. Thus, the partition function can be expressed as

$$Z_N = \sum_{\mathbf{x}\in S_N}\prod_{l=1}^{K} A_{N_l}(x_l)\exp\left(-H_N(\mathbf{x})\right). \tag{47}$$

Here, $A_{N_l}(x_l)$ counts the number of configurations of $\sigma^{(l)} \in \{-1, 1\}^{N_l}$ that share the same magnetization x_l. From Lemma 3, we can obtain the following bound for the generating function (8), by substituting (44) into (47):

$$-\frac{1}{N}\left(\log C + \frac{1}{2}\sum_{l=1}^{K}\log N_l\right) + \max_{\mathbf{x}} f_N(\mathbf{x}) \le p_N \le \frac{1}{N}\sum_{l=1}^{K}\log(N_l+1) + \max_{\mathbf{x}} f_N(\mathbf{x}) \tag{48}$$

where

$$f_N(\mathbf{x}) = \frac{1}{2}\sum_{p,l=1}^{K} \alpha_{N,l} J_{pl} \alpha_{N,p} x_p x_l + \sum_{p=1}^{K} \alpha_{N,p} h_p x_p - \sum_{p=1}^{K} \alpha_{N,p} I(x_p). \tag{49}$$

Therefore,

$$\lim_{N\to\infty} p_N = \max_{\mathbf{x}\in[-1,1]^K} f(\mathbf{x}) \tag{50}$$

where

$$f(\mathbf{x}) = \lim_{N\to\infty} f_N(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, \Delta\mathbf{x}) + (\tilde{\mathbf{h}}, \mathbf{x}) - (\hat{\boldsymbol{\alpha}}, I(\mathbf{x})) \tag{51}$$

and $I(\mathbf{x}) = (I(x_l))_{l\leq K}$ is defined in (45). The stationarity conditions are

$$x_l = \tanh\left(h_l + \sum_{p=1}^{K} \alpha_p J_{lp} x_p\right) \qquad \text{for} \quad l = 1, \ldots, K. \tag{52}$$

The solutions of the fixed point Equation (52) correspond to the stationary points of f, among which we are interested in those that achieve the supremum. There may be more than one global maximizer, depending on the parameters $(\mathbf{J}, \mathbf{h})$ of the model. In general, even for $K = 2$, the landscape of critical points of f, depending on the values $(\mathbf{J}, \mathbf{h})$, can be rather complicated; see [48,49] for a detailed study of this case. The Hessian of f can be expressed as

$$\mathcal{H}_f(\mathbf{x}) = \boldsymbol{\alpha}\left(\mathbf{J} - \boldsymbol{\alpha}^{-1}\mathrm{diag}(1 - x_p^2)_{p\leq K}^{-1}\right)\boldsymbol{\alpha}\,.$$

Thus, the Hessian is negative definite whenever $\mathbf{J} - \boldsymbol{\alpha}^{-1}\mathrm{diag}(1 - x_p^2)_{p\leq K}^{-1}$ is. Using a test vector, it is not difficult to see that if $\mathbf{J} - \boldsymbol{\alpha}^{-1}$ is negative definite, then $\mathcal{H}_f(\mathbf{x})$ remains negative definite (and hence f concave) over the entire optimization domain. This regime, in which Equation (52) has only one solution, can be identified with the case where interactions are not strong enough to induce a phase transition. On the contrary, if $\mathbf{J} - \boldsymbol{\alpha}^{-1}$ has even a single positive eigenvalue, the variational potential develops an unstable saddle around the origin, and there exists some $\mathbf{x} \in [-1, 1]^K$ where convexity and concavity are exchanged again. What we can say for sure, however, is that for $\mathbf{h} = 0$, the point $\mathbf{x} = 0$ is always critical, and is *the* solution to (52) when $\mathbf{J} - \boldsymbol{\alpha}^{-1} \prec 0$. On the other hand, if $\mathbf{J} - \boldsymbol{\alpha}^{-1}$ has a positive eigenvalue, $\mathbf{x} = 0$ becomes a saddle, and the potential necessarily develops other maxima. Therefore, for $\mathbf{h} = 0$ for instance, a phase transition must necessarily occur when the maximum eigenvalue (with sign, not spectral radius) of $\mathbf{J} - \boldsymbol{\alpha}^{-1}$ vanishes.

Here, we assume that $(\mathbf{J}, \mathbf{h})$ are such that f admits a unique maximizer $\boldsymbol{\mu} = (\mu^{(l)})_{l\leq K}$ with negative definite Hessian, i.e., $\mathcal{H}_f(\boldsymbol{\mu}) \prec 0$. Moreover, f is smooth around $\boldsymbol{\mu}$ since the latter lies in the interior of $[-1, 1]^K$ for any choice of $(\mathbf{J}, \mathbf{h})$ (see Lemma (A1)).

The main idea is to derive the asymptotic expansion of the partition function for a perturbed system, where a small external field is added. For any $\mathbf{t} \in \mathbb{R}^K$, we define

$$H_{N,\mathbf{t}}(\mathbf{m}_N) = H_N(\mathbf{m}_N) - \sqrt{N}(\mathbf{t}, \sqrt{\boldsymbol{\alpha}_N}\,\mathbf{m}_N) \tag{53}$$

where $\sqrt{\boldsymbol{\alpha}_N} = \mathrm{diag}(\sqrt{N_p/N})_{p\leq K}$, and $H_N(\mathbf{m}_N)$ is the unperturbed Hamiltonian defined in (5). The associated partition function is denoted by $Z_{N,\mathbf{t}}$. We also define the function

$$f_{N,\mathbf{t}}(\mathbf{x}) = f_N(\mathbf{x}) + \frac{1}{\sqrt{N}}(\mathbf{t}, \sqrt{\boldsymbol{\alpha}_N}\mathbf{x})\,. \tag{54}$$

Observe that from (54), the function $f_{N,\mathbf{t}}$, and all its partial derivatives, with respect to $\mathbf{x}$, of any order converge uniformly to those of f on the interior of $[-1,1]^K$. Hence, by Lemma A2, for large enough N, the function $f_{N,\mathbf{t}}$ admits a unique maximizer $\boldsymbol{\mu}_{N,\mathbf{t}} = (\mu^{(l)}_{N,\mathbf{t}})_{l\leq K} \in \mathbb{R}^K$ that converges to $\boldsymbol{\mu}$, and satisfies $\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}}) \prec 0$. Next, we demonstrate that the magnetization vector concentrates around $\boldsymbol{\mu}_{N,\mathbf{t}}$ with overwhelming probability under the Boltzmann–Gibbs measure $\mathcal{G}_{N,\mathbf{t}}$ induced by $H_{N,\mathbf{t}}$. For $\delta > 0$, we denote by $B_{N,\delta}$ a poly-interval centered at $\boldsymbol{\mu}_{N,\mathbf{t}}$, with each coordinate $\mu^{(l)}_{N,\mathbf{t}}$ as the center and $N_l^{-\frac{1}{2}+\delta}$ as the length of the interval in the l^{th} direction:

$$B_{N,\delta} = \left\{ \mathbf{x} \in \mathbb{R}^K : \left| x_l - \mu^{(l)}_{N,\mathbf{t}} \right| < N_l^{-\frac{1}{2}+\delta},\ \forall\, l \in \{1,\ldots,K\} \right\}. \tag{55}$$

Lemma 4. *Let $\mathbf{t} \in \mathbb{R}^K$ and assume that $f(\mathbf{x})$ admits a unique global maximizer $\boldsymbol{\mu}$ satisfying $\mathcal{H}_f(\boldsymbol{\mu}) \prec 0$. Then, for a sufficiently large N, $f_{N,\mathbf{t}}$ has a unique maximizer $\boldsymbol{\mu}_{N,\mathbf{t}} \xrightarrow{N\to\infty} \boldsymbol{\mu}$ and $\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}}) \prec 0$. Moreover, for $\delta \in (0, \frac{1}{2K+4})$, we have that*

$$\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta}(\boldsymbol{\mu}_{N,\mathbf{t}})) \leq \exp\left\{\frac{1}{2}N^{2\delta}\lambda_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right). \tag{56}$$

where $\lambda_{N,\mathbf{t}} < 0$ and the partition function (47) can be expanded as

$$Z_{N,\mathbf{t}} = \frac{e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})}}{\sqrt{\det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}})\right)\prod_{l=1}^K(1-(\mu^{(l)}_{N,\mathbf{t}})^2)}} \cdot \left(1 + O\left(N^{-\frac{1}{2}+(K+2)\delta}\right)\right). \tag{57}$$

Proof. By Lemma A2, for a sufficiently large N, $f_{N,\mathbf{t}}$ has a unique non-degenerate global maximizer. For the concentration bound, we observe

$$\begin{aligned}
\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta}) &= \frac{\sum_{\mathbf{x}\in S_N\cap B^c_{N,\delta}} \prod_{l=1}^K A_{N_l}(x_l)\exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right)}{\sum_{\mathbf{x}\in S_N} \prod_{l=1}^K A_{N_l}(x_l)\exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right)} \\
&\leq \frac{C\prod_{l=1}^K \sqrt{N_l}(N_l+1)\sup_{\mathbf{x}\in B^c_{N,\delta}} e^{Nf_{N,\mathbf{t}}(\mathbf{x})}}{\sup_{\mathbf{x}\in[-1,1]^K} e^{Nf_{N,\mathbf{t}}(\mathbf{x})}} \\
&= \exp\left\{N\left(\sup_{\mathbf{x}\in B^c_{N,\delta}} f_{N,\mathbf{t}}(\mathbf{x}) - f_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})\right)\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right).
\end{aligned} \tag{58}$$

The sup in the denominator of the inequality appearing in the second line has been extended over $[-1,1]^K$ instead of S_N due to the fact that $f_{N,\mathbf{t}}$ is Lipschitz continuous away from the boundaries of $[-1,1]^K$. The difference between $\sup_{\mathbf{x}\in S_N} f_{N,\mathbf{t}}(\mathbf{x})$ and $\sup_{\mathbf{x}\in[-1,1]^K} f_{N,\mathbf{t}}(\mathbf{x})$ is small due to the Lipschitz property of $f_{N,\mathbf{t}}$ and bounded by a constant due to the mesh size of S_N. Let us consider the simplest case, where $K = 1$. Then, the set $S_N = S_1$ and $S_1 = \left\{-1 + \frac{2n}{N_1}, n = 0,\ldots,N_1\right\}$ with mesh size $\frac{2}{N_1}$ and $\mathbf{t} \in \mathbb{R}^K = t$. Now, the Lipschitz continuity of $f_{N,t}$ means there exists a constant L such that $|f_{N,t}(x) - f_{N,t}(y)| \leq L|x-y|$ for all x, y away from ± 1. Given that the maximum distance between any point $x \in S_N$ and the nearest point in $(-1,1)$ is at most $\frac{2}{N_1}$, the difference between the supremum of $f_{N,t}$ over S_N and over $(-1,1)$ can be bounded by $\left|\sup_{x\in S_N} f_{N,t}(x) - \sup_{x\in(-1,1)} f_{N,t}(x)\right| \leq \frac{2L}{N_1}$. Thus, the difference is small and goes to zero as $N_1 \to \infty$. This justifies substituting the sum over the discrete set S_N with the supremum over $[-1,1]$, with the error being controlled by the mesh size and the Lipschitz constant L.

The supremum of $f_{N,\mathbf{t}}$ restricted to $B^c_{N,\delta}$, in the last equality of (58), is attained at some $\mathbf{y}_{N,\mathbf{t},\delta}$ on the boundary of $B_{N,\delta}$. A formal proof can be obtained through a straightforward generalization of Lemma B.11 in [43]. In fact, if $\mathbf{x} \in B^c_{N,\delta}$, then at least one of the coordinates x_l is outside $\left(\mu^{(l)}_{N,\mathbf{t}} - N_l^{-\frac{1}{2}+\delta}, \mu^{(l)}_{N,\mathbf{t}} + N_l^{-\frac{1}{2}+\delta}\right)$. The function $f_{N,\mathbf{t}}$ decreases outside $B_{N,\delta}$, and hence the supremum must be on the boundary. Bearing this in mind, we have

$$|\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}}|^2 \leq \sum_{l \leq K} N_l^{-1+2\delta} = \sum_{l \leq K} (N\alpha_{N,l})^{-1+2\delta} = N^{-1+2\delta} \underbrace{\sum_{l \leq K} (\alpha_{N,l})^{-1+2\delta}}_{=c_{\alpha_N,\delta}} = N^{-1+2\delta} c_{\alpha_N,\delta} \quad (59)$$

and, hence, using a Taylor expansion up to third order with Lagrange type remainder,

$$\begin{aligned} f_{N,\mathbf{t}}(\mathbf{y}_{N,\mathbf{t},\delta}) &= f_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) + \frac{1}{2}\Big((\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}})\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}}), (\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}})\Big) \\ &\quad + \frac{1}{6}\sum_{l,p,s=1}^{K} \frac{\partial^3 f_{N,\mathbf{t}}}{\partial x_l \partial x_p \partial x_s}(\vartheta_{N,\mathbf{t}})(\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}})_l(\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}})_p(\mathbf{y}_{N,\mathbf{t},\delta} - \boldsymbol{\mu}_{N,\mathbf{t}})_s \\ &\leq f_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) + \frac{1}{2}N^{-1+2\delta}\underbrace{c_{\alpha_N,\delta}}_{>1}\lambda_{N,\mathbf{t}} + \frac{1}{6}\sum_{l,p,s=1}^{K}\left|\frac{\partial^3 f_{N,\mathbf{t}}}{\partial x_l \partial x_p \partial x_s}(\vartheta_{N,\mathbf{t}})\right| \cdot N^{-3/2+3\delta}(c_{\alpha_N,\delta})^{3/2} \\ &\leq f_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) + \frac{1}{2}N^{-1+2\delta}\lambda_{N,\mathbf{t}} + \mathcal{O}\left(N^{-3/2+3\delta}\right) \end{aligned} \quad (60)$$

where $\vartheta_{N,\mathbf{t}}$ is an intermediate point of the segment $[\boldsymbol{\mu}_{N,\mathbf{t}}, \mathbf{y}_{N,\mathbf{t},\delta}]$, $\lambda_{N,\mathbf{t}} < 0$ is the largest eigenvalue of $\mathcal{H}_{f_{N,\mathbf{t}}}$, and in the third term of the last inequality of (60) we have that $\alpha_{N,l} \to \alpha_l$ as $N \to \infty$. Hence, (56) follows from (58) and (60).

We now prove the asymptotic expansion of the partition function $Z_{N,\mathbf{t}}$ (57). Observe that the concentration results, Equation (56), imply that almost all the contribution to $Z_{N,\mathbf{t}}$ comes from spin configurations with magnetization $\mathbf{m}_N$ in a vanishing neighborhood of $\boldsymbol{\mu}_{N,\mathbf{t}}$, i.e., $\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B_{N,\delta}) = 1 - \mathcal{O}(e^{-cN^{2\delta}})$ for some $c > 0$. Hence,

$$Z_{N,\mathbf{t}} = \left(1 + \mathcal{O}\left(e^{-cN^{2\delta}}\right)\right) \sum_{\mathbf{x} \in S_N \cap B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}), \quad \zeta_{N,\mathbf{t}}(\mathbf{x}) := \prod_{l=1}^{K} \binom{N_l}{\frac{N_l(1+x_l)}{2}} \exp(-H_{N,\mathbf{t}}(\mathbf{x})). \quad (61)$$

Following the same argument as [43,44], we approximate the sum in (61) by an integral using Lemma A5 over the set $B_{N,\delta}$ with shrinking interval containing the unique vector of the global maximizer of $f_{N,\mathbf{t}}$, which are elements of S_N:

$$\begin{aligned} \left| \int_{B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}) d\mathbf{x} - \frac{2^K}{\prod_{l=1}^{K} N_l} \sum_{\mathbf{x} \in S_N \cap B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}) \right| &\leq N^{\left(\frac{1}{2}+\delta\right)(K-1)} \cdot N^{-\frac{1}{2}+\delta} \cdot N^{-K} \sup_{\mathbf{x} \in B_{N,\delta}} |\nabla \zeta_{N,t}(\mathbf{x})| \\ &= \mathcal{O}\left(N^{\left(-\frac{1}{2}+\delta\right)(K+1)}\right) \zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) \end{aligned} \quad (62)$$

where $\sup_{\mathbf{x} \in B_{N,\delta}} |\nabla \zeta_{N,\mathbf{t}}(\mathbf{x})|$ is bounded in Lemma A3. Now, following from (62) and applying the results of Lemma A6 to approximate the integral, we have that

$$\sum_{\mathbf{x}\in S_N\cap B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}) = \frac{\prod_{l=1}^K N_l}{2^K}\int_{B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x})d\mathbf{x} + \mathcal{O}\left(N^{\left(-\frac{1}{2}+\delta\right)(K+1)+K}\right)\zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})$$
$$= \frac{\prod_{l=1}^K N_l}{2^K}\sqrt{\frac{2^K}{\pi^K\prod_{l=1}^K N_l(1-(\mu_{N,\mathbf{t}}^{(l)})^2)}}\sqrt{\frac{(2\pi)^K}{\prod_{l=1}^K N_l \det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}})\right)}}e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})}.$$
$$\cdot\left(1+\mathcal{O}\left(N^{-1/2+(K+2)\delta}\right)\right)$$
$$= \frac{e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})}}{\sqrt{\det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}})\right)\prod_{l=1}^K(1-(\mu_{N,\mathbf{t}}^{(l)})^2)}}\cdot\left(1+\mathcal{O}\left(N^{-\frac{1}{2}+(K+2)\delta}\right)\right) \tag{63}$$

for $\delta \in \left(0, \frac{1}{2K+4}\right)$. Therefore, we obtain the asymptotic form of the partition function:

$$Z_{N,\mathbf{t}} = \frac{e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})}}{\sqrt{\det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}_{N,\mathbf{t}})\right)\prod_{l=1}^K(1-(\mu_{N,\mathbf{t}}^{(l)})^2)}}\cdot\left(1+\mathcal{O}\left(N^{-\frac{1}{2}+(K+2)\delta}\right)\right). \tag{64}$$

This completes the proof. □

Now we are ready to proof Theorem 2. In order to approximate the distribution of the scaled difference between the vector of global species magnetization and the limiting global maximizers, we will compute the limiting moment generating function of $\mathbf{m}_N = (m_l)_{l\leq K}$ for some $\mathbf{t}\in\mathbb{R}^K$ using the expanded form of the partition function in (64):

$$\mathbb{E}\left[e^{\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N-\boldsymbol{\mu}))}\right] = e^{-\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu})}\int_{\mathbb{R}^K} e^{\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathbf{m}_N)}\mathcal{G}_N(\mathbf{m}_N)d\mathbf{m}_N$$
$$= e^{-\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu})}\frac{Z_{N,\mathbf{t}}}{Z_N}. \tag{65}$$

From the last equality on the right of (65) and using (64),

$$\frac{Z_{N,\mathbf{t}}}{Z_N} \sim \frac{\exp\left(N\max_{\mathbf{x}} f_{N,\mathbf{t}}(\mathbf{x})\right)}{\exp\left(N\max_{\mathbf{x}} f_N(\mathbf{x})\right)}$$
$$= \exp\left(N[f_N(\boldsymbol{\mu}_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}_N)] + \sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_{N,\mathbf{t}}\right)\right) \tag{66}$$

where $\sim$ denotes an equality up to $(1+o(1))$. In the last equality above, we had that $\boldsymbol{\mu}_{N,\mathbf{t}}$ and $\boldsymbol{\mu}_N$ are unique global maximizers of $f_{N,\mathbf{t}}(\mathbf{x})$ and $f_N(\mathbf{x})$, respectively. Now, following from Lemma A4, Equation (66) becomes

$$\frac{Z_{N,\mathbf{t}}}{Z_N} \sim \exp\left(-\frac{1}{2}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\right) + \sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_N)\right). \tag{67}$$

Therefore,

$$\mathbb{E}\left[e^{\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N-\boldsymbol{\mu}))}\right] \sim e^{-\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}\right)}\cdot e^{\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_N\right)}\cdot e^{-\frac{1}{2}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\right)}$$
$$= e^{\sqrt{N}\left[\mathbf{t}\sqrt{\boldsymbol{\alpha}_N}(\boldsymbol{\mu}_N-\boldsymbol{\mu})\right]}\cdot e^{-\frac{1}{2}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\right)}. \tag{68}$$

Now, from (49) we have that

$$\boldsymbol{\mu}_N = \tanh\left(\mathbf{J}\boldsymbol{\alpha}_N\boldsymbol{\mu}_N + \mathbf{h}\right). \tag{69}$$

Let $\boldsymbol{\beta} = (\beta_p)_{p\leq K}$ with $|\beta_p| < \infty$, and $\theta \in \left[\frac{1}{2}, \infty\right)$. Assume that $\boldsymbol{\alpha}_N \equiv \boldsymbol{\alpha}(\boldsymbol{\beta}) = \boldsymbol{\alpha} + N^{-\theta}\mathrm{diag}(\boldsymbol{\beta})$. If we set $\mathbf{b}_N = \mathbf{J}\boldsymbol{\alpha}_N\boldsymbol{\mu}_N + \mathbf{h}$, then

$$\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}} = \frac{\partial\boldsymbol{\mu}_N}{\partial\mathbf{b}_N}\cdot\frac{\partial\mathbf{b}_N}{\partial\boldsymbol{\beta}} \tag{70}$$

where

$$\frac{\partial\boldsymbol{\mu}_N}{\partial\mathbf{b}_N} = \mathrm{diag}(1-\mu_{N,l}^2)_{l\leq K} \quad \text{and} \quad \frac{\partial\mathbf{b}_N}{\partial\boldsymbol{\beta}} = \mathbf{J}\,\frac{\partial\boldsymbol{\alpha}_N}{\partial\boldsymbol{\beta}}\boldsymbol{\mu}_N + \mathbf{J}\,\boldsymbol{\alpha}_N\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}}. \tag{71}$$

Let $\mathbf{M} = \mathrm{diag}(1-\mu_{N,l}^2)_{l\leq K}$ then

$$\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}} = \mathbf{M}\left[\mathbf{J}\,\frac{\partial\boldsymbol{\alpha}_N}{\partial\boldsymbol{\beta}}\boldsymbol{\mu}_N + \mathbf{J}\,\boldsymbol{\alpha}_N\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}}\right] \tag{72}$$

which entails

$$\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}}\left[\mathbf{M}^{-1} - \mathbf{J}\,\boldsymbol{\alpha}_N\right] = N^{-\theta}\,\mathbf{J}\,\mathrm{diag}(\boldsymbol{\mu}_N). \tag{73}$$

Since $\mathcal{H}_{f_N}(\boldsymbol{\mu}_N) = -\left[\mathbf{M}^{-1} - \mathbf{J}\,\boldsymbol{\alpha}_N\right]\cdot\boldsymbol{\alpha}_N$, hence, (72) yields

$$\frac{\partial\boldsymbol{\mu}_N}{\partial\boldsymbol{\beta}} = \left[\mathbf{M}^{-1} - \mathbf{J}\boldsymbol{\alpha}_N\right]^{-1} N^{-\theta}\,\mathbf{J}\,\mathrm{diag}(\boldsymbol{\mu}_N) = -N^{-\theta}\,\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\,\boldsymbol{\alpha}_N\,\mathbf{J}\,\mathrm{diag}(\boldsymbol{\mu}_N). \tag{74}$$

Finally, a Taylor expansion of $\boldsymbol{\mu}_N$ around $\boldsymbol{\beta} = \mathbf{0}$ gives

$$\boldsymbol{\mu}_N - \boldsymbol{\mu} = -N^{-\theta}\,\mathcal{H}_f^{-1}(\boldsymbol{\mu})\,\boldsymbol{\alpha}\,\mathbf{J}\,\mathrm{diag}(\boldsymbol{\mu})\boldsymbol{\beta} + \mathcal{O}\left(N^{-2\theta}\right). \tag{75}$$

Therefore, Theorem 2 follows from (75) and taking the limit as $N\to\infty$ of (68).

4.3. Proof of Theorem 3

We will now address the case where f reaches its maximum in more than one point.

Lemma 5. *Suppose $f(\mathbf{x})$ has n global maximizers $\boldsymbol{\mu}^i$ for $i\leq n$ satisfying $\mathcal{H}_f(\boldsymbol{\mu}^i) \prec 0$. For each $i\leq n$, let $A_i \subset [-1,1]^K$ be a poly-interval such that $\boldsymbol{\mu}^i \in int(A_i)$ is the unique maximizer of f on $cl(A_i)$. Then, for each $i\leq n$ and for a sufficiently large N, $f_{N,\mathbf{t}}$ has a unique global maximizer $\boldsymbol{\mu}^i_{N,\mathbf{t}} \to \boldsymbol{\mu}$ on A_i with $\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}^i_{N,\mathbf{t}}) < 0$. Moreover, for $\delta \in \left(0, \frac{1}{2K+4}\right)$, one has*

$$\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}})|\mathbf{m}_N \in A_i) = \exp\left\{\frac{1}{2}N^{2\delta}\lambda^i_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right) \tag{76}$$

where $\lambda^i_{N,\mathbf{t}} < 0$ is the largest eigenvalue of $\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}^i_{N,\mathbf{t}})$ for $\boldsymbol{\mu}^i_{N,\mathbf{t}} \in int(A_i)$ and

$$\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta,n}) = \exp\left\{\frac{1}{2}N^{2\delta}\max_{1\leq i\leq n}\lambda^i_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right) \tag{77}$$

where $B^c_{N,\delta,n} = \bigcup_{i\leq n} B^c_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}})$. The partition function restricted to the interval A_i can be expanded as

$$
\begin{aligned}
Z_{N,\mathbf{t}}\big|_{A_i} &= \sum_{\mathbf{x}\in S_N\cap A_i}\prod_{l=1}^{K} A_{N_l}(x_l)\exp N(f_{N,\mathbf{t}}(\mathbf{x})) \\
&= \frac{e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}^i_{N,\mathbf{t}})}}{\sqrt{\det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}^i_{N,\mathbf{t}})\right)\prod_{l=1}^{K}\left(1-\left(\mu^{i,(l)}_{N,\mathbf{t}}\right)^2\right)}}\cdot\left(1+O\left(N^{-1/2+(K+2)\delta}\right)\right).
\end{aligned}
\tag{78}
$$

Proof. For a sufficiently large N, $B_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}}) \subset A_i$ for all $i \leq n$ and equation (76) is obtained following the step-by-step argument used to prove Equation (56). Hence, it follows that

$$
\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}})|\mathbf{m}_N \in A_i) = \exp\left\{\frac{1}{2}N^{2\delta}\lambda^i_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right).
\tag{79}
$$

Now, let us observe that, for a sufficiently large N, $A_i \setminus B_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}}) = A_i \setminus B_{N,\delta,n}$ for $i \leq n$. Hence, for all $1 \leq i \leq n$ and N large, it follows that

$$
\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}})|\mathbf{m}_N \in A_i) = \mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta,n}|\mathbf{m}_N \in A_i).
\tag{80}
$$

Therefore, we have that

$$
\begin{aligned}
\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta,n}) &= \sum_{1\leq i\leq n}\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B^c_{N,\delta,n}|\mathbf{m}_N \in A_i)\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in A_i) \\
&\leq \exp\left\{\frac{1}{2}N^{2\delta}\max_{1\leq i\leq n}\lambda^i_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right)\sum_{1\leq i\leq n}\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in A_i) \\
&= \exp\left\{\frac{1}{2}N^{2\delta}\max_{1\leq i\leq n}\lambda^i_{N,\mathbf{t}}\right\}\mathcal{O}\left(N^{\frac{3K}{2}}\right).
\end{aligned}
\tag{81}
$$

This complete the result in (77) following from (81).

The proof for the asymptotic expansion of the partition function when there are multiple vectors of global maximizers of $f_{N,\mathbf{t}}$ follows exactly the same argument for the case with a unique vector of global maximizer when conditioned on an interval containing only one of the global maximizers. Notice that for fixed $i \leq n$ and N large, $\mathbf{m}_N$ concentrates around $\boldsymbol{\mu}^i_{N,\mathbf{t}} \in A_i$ as stated in Equation (76). Hence,

$$
\mathcal{G}_{N,\mathbf{t}}(\mathbf{m}_N \in B_{N,\delta}(\boldsymbol{\mu}^i_{N,\mathbf{t}})|\mathbf{m}_N \in A_i) = \frac{1}{Z_{N,\mathbf{t}}\big|_{A_i}}\sum_{\mathbf{x}\in S_N\cap B_{N,\delta}}\prod_{l=1}^{K}\binom{N_l}{\frac{N_l(1+x_l)}{2}}\exp\{-H_{N,\mathbf{t}}(\mathbf{x})\}.
\tag{82}
$$

Now, following the exact computation and argument in the uniqueness case, the restricted partition function for each of the global maximizers $\boldsymbol{\mu}^i_{N,\mathbf{t}}$ for $i \leq n$ can be expanded as

$$
Z_{N,\mathbf{t}}\big|_{A_i} = \frac{e^{Nf_{N,\mathbf{t}}(\boldsymbol{\mu}^i_{N,\mathbf{t}})}}{\sqrt{\det\left(-\mathcal{H}_{f_{N,\mathbf{t}}}(\boldsymbol{\mu}^i_{N,\mathbf{t}})\right)\prod_{l=1}^{K}\left(1-\left(\mu^{i,(l)}_{N,\mathbf{t}}\right)^2\right)}}\cdot\left(1+\mathcal{O}\left(N^{-1/2+(K+2)\delta}\right)\right).
\tag{83}
$$

This completes the proof of Lemma 5. □

The conditional moment generating function for a certain parameter $\mathbf{t} \in \mathbb{R}^K$ can be computed as

$$\mathbb{E}\left[e^{\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N-\boldsymbol{\mu}^i))}\middle|\left\{\mathbf{m}_N\in A_i\right\}\right]$$
$$= e^{-\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}^i\right)}\int_{\mathbb{R}^K} e^{-\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathbf{m}_N\right)}\mathcal{G}_N(\mathbf{m}_N)|_{\{\mathbf{m}_N\in A_i\}}d\mathbf{m}_N$$
$$= e^{-\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}^i\right)}\frac{Z_{N,\mathbf{t}|A_i}}{Z_{N|A_i}}. \quad (84)$$

Using the asymptotic expansion of the perturbed partition function in (83), the proof follows the same arguments as in Theorem 2. Hence,

$$\frac{Z_{N,\mathbf{t}|A_i}}{Z_{N|A_1}} \sim \exp\left(N[f_N(\boldsymbol{\mu}^i_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}^i_N)] + \sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}^i_{N,\mathbf{t}}\right)\right) \quad (85)$$

where $\boldsymbol{\mu}^i_{N,\mathbf{t}} \in int(A_i)$ is the unique maximizer of $f_{N,\mathbf{t}}$ and $\boldsymbol{\mu}^i_N \in int(A_i)$ is the unique maximizer of f_N. Now, following the argument of Lemma A4 and the proof of Theorem 2, we have that

$$\mathbb{E}\left[e^{\sqrt{N}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}(\mathbf{m}_N-\boldsymbol{\mu}^i))}\middle|\left\{\mathbf{m}_N\in A_i\right\}\right] \sim e^{-\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}^i\right)}\cdot e^{\sqrt{N}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}^i_N\right)}\cdot e^{-\frac{1}{2}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}^{-1}_{f_N}(\boldsymbol{\mu}^i_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\right)}$$
$$= e^{\sqrt{N}\left[\mathbf{t}\sqrt{\boldsymbol{\alpha}_N}\left(\boldsymbol{\mu}^i_N-\boldsymbol{\mu}^i\right).\right]}\cdot e^{-\frac{1}{2}\left(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}^{-1}_{f_N}(\boldsymbol{\mu}^i_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\right)}. \quad (86)$$

Similarly, if we set $\boldsymbol{\alpha}_N \equiv \boldsymbol{\alpha}(\boldsymbol{\beta}) = \boldsymbol{\alpha} + N^{-\theta}\mathrm{diag}(\boldsymbol{\beta})$, it follows from (75) that

$$\boldsymbol{\mu}^i_N - \boldsymbol{\mu}^i = -N^{-\theta}\,\mathcal{H}^{-1}_f(\boldsymbol{\mu}^i)\,\boldsymbol{\alpha}\,\mathbf{J}\,\mathrm{diag}(\boldsymbol{\mu}^i)\boldsymbol{\beta} + \mathcal{O}\left(N^{-2\theta}\right) \quad (87)$$

for $\boldsymbol{\beta} = (\beta_p)_{p\leq K}$ where $0 < \beta_p < \infty$ and $\theta \in \left[\frac{1}{2},\infty\right)$. Hence, Theorem 3 follows from (87) by taking the limit as $N \to \infty$ of (86). This completes the proof of Theorem 3.

5. Conclusions

In this work, we introduced a generalized version of the multispecies Curie–Weiss model featuring arbitrary spins and a non-definite interaction matrix. Specifically, we computed the large-number limit of the generating functional associated with the Hamiltonian (3) and established the validity of the Central Limit Theorem (CLT) for a suitably rescaled vector of global magnetization in the Ising spin case. Our results demonstrate that the rescaled global magnetization vector follows either a centered or non-centered multivariate normal distribution, depending on the rate at which the relative particle densities converge to their limiting values as the system size tends to infinity.

This generalized framework holds significant promise for applications in systems where spin particles or units can exhibit arbitrary states, moving beyond traditional binary or discrete-valued models. Such flexibility is particularly relevant in fields like statistical physics, network theory, and social dynamics, where elements often have a continuum of possible states or operate under complex interactions. Furthermore, the inclusion of an indefinite interaction matrix enables the modeling of systems with both cooperative and antagonistic interactions [50], broadening the scope of potential applications to include meta-magnets, financial systems, and multi-agent interactions.

Future work could explore extending these results to models with more complex topologies or additional constraints, such as time-varying interaction matrices or non-equilibrium dynamics [51–53]. Additionally, investigating the robustness of the CLT under perturbations in the interaction matrix or in the distribution of spin states could yield further insights into the stability and universality of the model's behavior and broaden its applicability [54].

Author Contributions: Conceptualization, F.C., E.M. and G.O.; methodology, F.C., E.M. and G.O.; validation, F.C., E.M. and G.O.; formal analysis, F.C., E.M. and G.O.; investigation, F.C., E.M. and G.O.; writing—original draft preparation, F.C., E.M. and G.O.; writing—review and editing, F.C., E.M. and G.O.; supervision, E.M.; project administration, G.O. All authors have read and agreed to the published version of the manuscript.

Funding: E.M. was supported by the EU H2020 ICT48 project Humane AI Net contract number 952026; by the Italian Extended Partnership PE01—FAIR Future Artificial Intelligence Research—Proposal code PE00000013 under the MUR National Recovery and Resilience Plan. E.M. and G.O. were supported by a grant for the project PRIN22CONTUCCI, 2022B5LF52 "Boltzmann Machines beyond the "Independent Identically Distributed" Paradigm: a Mathematical Physics Approach", CUP J53D23003690006, under the MUR National Recovery and Resilience Plan.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank Pierluigi Contucci for inspiring this collaboration. They are also grateful to the anonymous reviewers for their valuable suggestions.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Technical Tools

Lemma A1. *The maximizer $\boldsymbol{\mu}$ of $f(\mathbf{x})$ belongs to the interior of $[-1,1]^K$.*

Proof. Observe that the function $f(\mathbf{x})$ satisfies

$$\frac{\partial f}{\partial x_l} = \sum_{p=1}^{K} \Delta_{l,p} x_p + \tilde{h}_l - \alpha_l \left(\frac{1}{2} \log \left(\frac{1+x_l}{1-x_l} \right) \right) \quad \text{for} \quad l = 1, ..., K. \tag{A1}$$

Suppose that by contradiction $\mu_l^2 = 1$ for some $l \leq K$. Then, if we consider the function $f(x_l) = f(\mathbf{x})\Big|_{x_j=\mu_j, j\neq l}$, it follows that $f(x_l)$ attains its maximum in at least one point $x_l \in (-1,1)$ which satisfies (52). Indeed, from (A1), $\lim_{x_l \to -1^+} \frac{\partial}{\partial x_l} f(x_l) = +\infty$ and $\lim_{x_l \to 1^-} \frac{\partial}{\partial x_l} f(x_l) = -\infty$. Therefore, there exists $\epsilon > 0$ such that $f(x_l)$ is strictly increasing on $[-1,-1+\epsilon]$ and strictly decreasing on $[1-\epsilon, 1]$. Since this is true for all $l = 1, \ldots, K$, then the thesis follows. □

Lemma A2. *Let $\boldsymbol{\mu}$ be a point on the interior of $[-1,1]^K$ and let A be an open neighborhood of $\boldsymbol{\mu}$. Let $f : cl(A) \to \mathbb{R}$ and assume that $\boldsymbol{\mu}$ is the unique global maximum point of f and the Hessian $\mathcal{H}_f(\boldsymbol{\mu})$ is negative definite. Let (f_N) be a sequence of functions with bounded partial derivatives up to order 2 converging uniformly to those of f. Then, for a sufficiently large N, f_N has a unique maximizer $\boldsymbol{\mu}_N \to \boldsymbol{\mu}$ and $\mathcal{H}_{f_N}(\boldsymbol{\mu}_N) \prec 0$.*

Proof. Suppose that $\{\mathbf{x}_N\}$ is a sequence of any maximizer of f_N which exists since $cl(A)$ is compact. Then, there exists a subsequence $\{N_n\}_{n\geq 1}$ such that $\{\mathbf{x}_{N_n}\}$ converges to some $\mathbf{y}$. Clearly, $f_{N_n}(\mathbf{x}_{N_n}) \geq f_{N_n}(\mathbf{x})$ for all $\mathbf{x} \in cl(A)$. Therefore, by uniform convergence and taking the limit as $n \to \infty$, we obtain that $f(\mathbf{y}) \geq f(\mathbf{x})$ for all $\mathbf{x} \in cl(A)$. This implies that $\mathbf{y} = \boldsymbol{\mu}$ by uniqueness of the global maximizers of f and, therefore, $\mathcal{H}_f(\boldsymbol{\mu}) \prec 0$. Now, since f_N converges uniformly to f, one has that for a sufficiently large N, the maximizer $\boldsymbol{\mu}_N$ of f_N is unique and $\mathcal{H}_{f_N}(\boldsymbol{\mu}_N) \prec 0$. □

Lemma A3. *For $\delta \in (0, 1/6]$, the following bound holds*

$$\sup_{\mathbf{x} \in B_{N,\delta}} |\nabla \zeta_{N,\mathbf{t}}(\mathbf{x})| \leq \zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) \mathcal{O}\Big(N^{1/2+\delta}\Big). \tag{A2}$$

Proof. The proof is carried in two steps, the first is to compute $\nabla \zeta_N(\mathbf{x})$ and the second finds the supremum.

Step 1: Let us recall from (61) that

$$\zeta_{N,\mathbf{t}}(\mathbf{x}) = \prod_{l=1}^{K} \binom{N_l}{\frac{N_l(1+x_l)}{2}} \exp\big\{-H_{N,\mathbf{t}}(\mathbf{x})\big\}. \tag{A3}$$

Suppose that $\frac{N_l(1+x_l)}{2}$ is any real number, then the binomial coefficient becomes a continuous binomial coefficient and can be expanded using the arguments of [43,55]. Now, using gamma functions, we have that for each $l = 1, \ldots, K$:

$$A_{N_l}(x_l) = \binom{N_l}{\frac{N_l(1+x_l)}{2}} = \frac{\Gamma(N_l + 1)}{\Gamma\Big(\frac{N_l(1+x_l)}{2} + 1\Big)\Gamma\Big(\frac{N_l(1-x_l)}{2} + 1\Big)}. \tag{A4}$$

For a given component l, differentiating with respect to x_l gives

$$\frac{\partial A_{N_l}(x_l)}{\partial x_l} = A_{N_l}(x_l)\bigg(-\psi\Big(\frac{N_l(1+x_l)}{2} + 1\Big) \cdot \frac{N_l}{2} + \psi\Big(\frac{N_l(1-x_l)}{2} + 1\Big) \cdot \frac{N_l}{2}\bigg). \tag{A5}$$

Here, $\psi(z)$ is the digamma function, the derivative of $\log \Gamma(z)$. Now, using the asymptotic expansion of ψ and the properties of Γ, we have that

$$\begin{aligned} \psi\Big(\frac{N_l(1+x_l)}{2} + 1\Big) &= \log\Big(\frac{N_l(1+x_l)}{2}\Big) + \frac{1}{N_l(1+x_l)} + \mathcal{O}\Big(N_l^{-2}\Big) \quad \text{and} \\ \psi\Big(\frac{N_l(1-x_l)}{2} + 1\Big) &= \log\Big(\frac{N_l(1-x_l)}{2}\Big) + \frac{1}{N_l(1-x_l)} + \mathcal{O}\Big(N_l^{-2}\Big). \end{aligned} \tag{A6}$$

Now, by the product and chain rule, we have that

$$\begin{aligned} \frac{\partial \zeta_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} &= \frac{\partial A_{N_l}(x_l)}{\partial x_l} \prod_{p \neq l} A_{N_p}(x_p) \cdot \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right) + A_{N_l}(x_l) \cdot \prod_{p \neq l} A_{N_p}(x_p) \cdot \frac{\partial}{\partial x_l} \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right) \\ &= \prod_{p \leq K} A_{N_p}(x_p) \frac{N_l}{2}\bigg(\log\Big(\frac{N_l(1-x_l)}{2}\Big) - \log\Big(\frac{N_l(1+x_l)}{2}\Big) + \frac{1}{N_l(1-x_l)} - \frac{1}{N_l(1+x_l)} + \mathcal{O}\Big(N_l^{-2}\Big)\bigg) \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right) \\ &\quad + \prod_{p \leq K} A_{N_p}(x_p) A_{N_l}(x_l) \cdot N\left\{\sum_{p=1}^{K} \Delta_{l,p} x_p + \tilde{h}_l + \frac{t_l\sqrt{N_l}}{N}\right\} \cdot \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right) \\ &= \prod_{p \leq K} A_{N_p}(x_p) \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right)\bigg(-N_l \operatorname{arctanh}(x_l) + \frac{x_l}{(1-x_l^2)} + N\left\{\sum_{p=1}^{K} \Delta_{l,p} x_p + \tilde{h}_l + \frac{t_l\sqrt{N_l}}{N}\right\} + \mathcal{O}\Big(N_l^{-1}\Big)\bigg) \\ &= \underbrace{\prod_{p \leq K} A_{N_p}(x_p) \exp\left(-H_{N,\mathbf{t}}(\mathbf{x})\right)}_{=\zeta_{N,\mathbf{t}}(\mathbf{x})}\left(N \underbrace{\left\{\sum_{p=1}^{K} \Delta_{l,p} x_p + \tilde{h}_l - \alpha_{N,l} \operatorname{arctanh}(x_l) + \frac{t_l\sqrt{N_l}}{N}\right\}}_{=\frac{\partial f_{N,\mathbf{t}}}{\partial x_l}} + \frac{x_l}{(1-x_l^2)} + \mathcal{O}\Big(N_l^{-1}\Big)\right) \\ &= \zeta_{N,\mathbf{t}}(\mathbf{x})\bigg(N \frac{\partial f_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} + \frac{x_l}{(1-x_l^2)} + \mathcal{O}\Big(N^{-1}\Big)\bigg). \end{aligned} \tag{A7}$$

Step 2: Observe from the last equality in (A7) that

$$\sup_{\mathbf{x}\in B_{N,\delta}} |\nabla\zeta_N(\mathbf{x})| = \sup_{\mathbf{x}\in B_{N,\delta}} \left\{ \left| \zeta_{N,\mathbf{t}}(\mathbf{x}) \left(N\frac{\partial f_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} + \frac{x_l}{(1-x_l^2)} + \mathcal{O}\left(N^{-1}\right) \right) \right| \right\}_{l\le K}. \tag{A8}$$

Now, for each $l \in \{1,\ldots,K\}$, by the Mean Value Theorem, there exists a point $\mathbf{c}$ in the line segment connecting $[\mathbf{x}, \boldsymbol{\mu}_{N,\mathbf{t}}]$ such that

$$\begin{aligned} \frac{\partial f_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} &= \frac{\partial f_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})}{\partial x_l} + \sum_{p=1}^{K} \frac{\partial^2 f_{N,\mathbf{t}}(\mathbf{c})}{\partial x_l \partial x_p}(x_p - \mu_{N,\mathbf{t}}^{(p)}) \\ &= \sum_{p=1}^{K} \frac{\partial^2 f_{N,\mathbf{t}}(\mathbf{c})}{\partial x_l \partial x_p}(x_p - \mu_{N,\mathbf{t}}^{(p)}). \end{aligned} \tag{A9}$$

Hence,

$$\sup_{\mathbf{x}\in B_{N,\delta}} \left| \frac{\partial f_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} \right| \le \sup_{\mathbf{c}\in B_{N,\delta}} \sum_{p=1}^{K} \left| \frac{\partial^2 f_{N,\mathbf{t}}(\mathbf{c})}{\partial x_l \partial x_p} \right| N^{-\frac{1}{2}+\delta}. \tag{A10}$$

Therefore,

$$\sup_{\mathbf{x}\in B_{N,\delta}} |\nabla f_{N,\mathbf{t}}(\mathbf{x})| = \mathcal{O}\left(N^{-1/2+\delta}\right) \tag{A11}$$

and using (46) we have that

$$\sup_{\mathbf{x}\in B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}) \le (1+\mathcal{O}(N^{-1}))\zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}}) \sup_{\mathbf{x}\in B_{N,\delta}} \sqrt{\frac{1-\boldsymbol{\mu}_{N,\mathbf{t}}^2}{1-\mathbf{x}^2}} = \zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})\mathcal{O}(1). \tag{A12}$$

This implies that

$$\sup_{\mathbf{x}\in B_{N,\delta}} \left| \frac{\partial \zeta_{N,\mathbf{t}}(\mathbf{x})}{\partial x_l} \right| \le \zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})\mathcal{O}\left(N^{1/2+\delta}\right). \tag{A13}$$

Hence,

$$\begin{aligned} \sup_{\mathbf{x}\in B_{N,\delta}} |\nabla\zeta_N(\mathbf{x})| \le \max_{l\in\{1,\ldots,K\}} &\left\{ \sup_{\mathbf{x}\in B_{N,\delta}} \zeta_{N,\mathbf{t}}(\mathbf{x}) \cdot N \cdot \sup_{\mathbf{c}\in B_{N,\delta}} \sum_{p=1}^{K} \left| \frac{\partial^2 f_{N,\mathbf{t}}(\mathbf{c})}{\partial x_l \partial x_p} \right| N^{-\frac{1}{2}+\delta} \right\} \\ &= \zeta_{N,\mathbf{t}}(\boldsymbol{\mu}_{N,\mathbf{t}})\mathcal{O}\left(N^{1/2+\delta}\right). \end{aligned} \tag{A14}$$

□

Lemma A4. *Let Ω be a bounded open set in $\mathbb{R}^K$. Let $f_N : \Omega \to \mathbb{R}$ be a sequence of functions such that for a sufficiently large N, it has a unique global maximum point $\boldsymbol{\mu}_N \in \Omega$ and $\mathcal{H}_{f_N}(\boldsymbol{\mu}_N) \prec 0$ with bounded partial derivatives up to order 3. For any $\mathbf{t} \in \mathbb{R}^K$, consider the function*

$$g_N(\mathbf{t},\mathbf{x}) = f_N(\mathbf{x}) + \frac{1}{\sqrt{N}}(\mathbf{t}, \sqrt{\boldsymbol{\alpha}_N}\mathbf{x}). \tag{A15}$$

Then, for a sufficiently large N, the function $g_N(\mathbf{t},\mathbf{x})$ also has a unique global maximizer with $\mathcal{H}_{g_N}(\boldsymbol{\mu}_{N,\mathbf{t}}) \prec 0$ and $\boldsymbol{\mu}_{N,\mathbf{t}} \to \boldsymbol{\mu}$ as $N \to \infty$, the following expansion holds:

$$g_N(\mathbf{t},\boldsymbol{\mu}_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}_N) = -\frac{1}{2N}(\mathbf{t}, \sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}) + \frac{1}{\sqrt{N}}(\mathbf{t}, \sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_N) + \mathcal{O}\left(N^{-3/2}\right). \tag{A16}$$

Proof. In order to prove (A16), let us start with

$$\nabla_{\mathbf{x}} g_N(\mathbf{t},\mathbf{x}) = \nabla_{\mathbf{x}} f_N(\mathbf{x}) + \frac{1}{\sqrt{N}}\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}. \tag{A17}$$

Now, since $\boldsymbol{\mu}_{N,\mathbf{t}}$ is a maximizer of $g_N(\mathbf{t},\mathbf{x})$, then

$$\nabla_{\mathbf{x}} g_N(\mathbf{t},\mathbf{x})\big|_{\mathbf{x}=\boldsymbol{\mu}_{N,\mathbf{t}}} = \mathbf{0} = \nabla_{\mathbf{x}} f_N(\boldsymbol{\mu}_{\mathbf{N},\mathbf{t}}) + \frac{1}{\sqrt{N}}\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}. \tag{A18}$$

Now, we take the gradient on both sides, obtaining

$$\nabla_{\mathbf{t}}\boldsymbol{\mu}_{N,\mathbf{t}} = -\frac{1}{\sqrt{N}}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_{N,\mathbf{t}})\sqrt{\boldsymbol{\alpha}_N}. \tag{A19}$$

Let us notice from Equation (A15), and using the fact that $\boldsymbol{\mu}_{N,\mathbf{t}}$ and $\boldsymbol{\mu}_N$ are the global maximizers of $g_N(\mathbf{t},\mathbf{x})$ and $f_N(\mathbf{x})$, respectively, we have that

$$g_N(\mathbf{t},\boldsymbol{\mu}_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}_N) = \underbrace{f_N(\boldsymbol{\mu}_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}_N)}_{=\Phi_N} + \frac{1}{\sqrt{N}}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_{N,\mathbf{t}}). \tag{A20}$$

By an application of Taylor's expansion of $f_N(\boldsymbol{\mu}_{N,\mathbf{t}})$ around $\boldsymbol{\mu}_N$:

$$\Phi_N = \frac{1}{2}\Big(\Big(\boldsymbol{\mu}_{N,\mathbf{t}} - \boldsymbol{\mu}_N\Big), \mathcal{H}_{f_N}(\boldsymbol{\mu}_N)\Big(\boldsymbol{\mu}_{N,\mathbf{t}} - \boldsymbol{\mu}_N\Big)\Big) + \mathcal{O}(N^{-3/2}). \tag{A21}$$

Now to compute Φ_N, we need $\boldsymbol{\mu}_{N,\mathbf{t}} - \boldsymbol{\mu}_N$. To begin with, let us first observe from Equation (A15) that, when $\mathbf{t} = \mathbf{0}$: $g_N(\mathbf{0},\mathbf{x}) = f_N(\mathbf{x})$ and hence $\boldsymbol{\mu}_{N,\mathbf{0}} = \boldsymbol{\mu}_N$. Therefore, taking Taylor's expansion of $\boldsymbol{\mu}_{N,\mathbf{t}}$ around $\mathbf{t} = \mathbf{0}$ and using (A19):

$$\begin{aligned}\boldsymbol{\mu}_{N,\mathbf{t}} - \boldsymbol{\mu}_{N,\mathbf{0}} =& \nabla_{\mathbf{t}}\boldsymbol{\mu}_{N,\mathbf{0}}\mathbf{t} + \mathcal{O}\left(\frac{1}{N}\right)\\ &= -\frac{1}{\sqrt{N}}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t} + \mathcal{O}\left(\frac{1}{N}\right).\end{aligned} \tag{A22}$$

It now follows from (A21) that

$$\begin{aligned}\Phi_N =& \frac{1}{2N}\Big(\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}, \mathcal{H}_{f_N}(\boldsymbol{\mu}_N)\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big) + \mathcal{O}(N^{-3/2})\\ =& \frac{1}{2N}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big) + \mathcal{O}(N^{-3/2})\end{aligned} \tag{A23}$$

and, therefore,

$$\begin{aligned}g_N(\mathbf{t},\boldsymbol{\mu}_{N,\mathbf{t}}) - f_N(\boldsymbol{\mu}_N) =& \frac{1}{2N}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big)\\ &+ \frac{1}{\sqrt{N}}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\Big(\boldsymbol{\mu}_N - \frac{1}{\sqrt{N}}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big)\Big) + \mathcal{O}(N^{-3/2})\\ =& \frac{1}{2N}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big) + \frac{1}{\sqrt{N}}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_N)\\ &- \frac{1}{N}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big) + \mathcal{O}(N^{-3/2})\\ =& -\frac{1}{2N}\Big(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\mathcal{H}_{f_N}^{-1}(\boldsymbol{\mu}_N)\sqrt{\boldsymbol{\alpha}_N}\mathbf{t}\Big) + \frac{1}{\sqrt{N}}(\mathbf{t},\sqrt{\boldsymbol{\alpha}_N}\boldsymbol{\mu}_N) + \mathcal{O}(N^{-3/2}).\end{aligned} \tag{A24}$$

This completes the proof. □

Appendix B. Approximation Lemmas

In this section, standard mathematical approximations which played a crucial role in the asymptotic expansion of the partition function are given.

Lemma A5 (Multidimensional Riemann Approximation)**.** *Let $Q = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_K, b_K]$ be a rectangular domain in $\mathbb{R}^K$, and let $P = \{(x_{1,0}, \ldots, x_{K,0}), (x_{1,1}, \ldots, x_{K,1}), \ldots, (x_{1,n}, \ldots, x_{K,n})\}$ be any partition of Q, where $a_i = x_{i,0} < x_{i,1} < \ldots < x_{i,n} = b_i$ for each $i = 1, 2, \ldots, K$. Assume that g has continuous partial derivatives $\frac{\partial g}{\partial x_i}$ on Q for all $i = 1, 2, \ldots, K$. Let $\epsilon_i = \max_{1 \le j \le n}(x_{i,j} - x_{i,j-1})$ denote the mesh size of the partition along the i-th variable. Then,*

$$\left| \int_Q g(\mathbf{x})\, d\mathbf{x} - \sum_{j_1, j_2, \ldots, j_K = 1}^{n} g(\mathbf{c}_{j_1, j_2, \ldots, j_K}) \cdot \prod_{i=1}^{K} (x_{i,j_i} - x_{i,j_i - 1}) \right| \le Kn^{K-1} \max_{i \le K}(b^{(i)} - a^{(i)}) \max_{\boldsymbol{\xi} \in Q} \|\nabla g(\boldsymbol{\xi})\| \prod_{i=1}^{K} \epsilon_i \quad \text{(A25)}$$

where $\mathbf{c}_{j_1, j_2, \ldots, j_r}$ is any point in the j_1-th subinterval along the first variable, j_2-th subinterval along the second variable, and so on, up to the j_K-th subinterval along the K-th variable.

Proof. To begin with, we can decompose the integral into a summation of integrals over all the poly-intervals $Q_{j_1, j_2, \ldots, j_K}$ constituting the mesh grid:

$$\int_Q g(\mathbf{x})\, d\mathbf{x} = \sum_{j_1, j_2, \ldots, j_K = 1}^{n} \int_{Q_{j_1, j_2, \ldots, j_K}} g(\mathbf{x}) d\mathbf{x}. \quad \text{(A26)}$$

Using the fact the g is continuous and that each poly-interval is compact, we can use the integral mean value theorem. Therefore, for any $j_1, j_2, \ldots, j_K$, there exists a $\boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}$ such that

$$\int_{Q_{j_1, j_2, \ldots, j_K}} g(\mathbf{x}) d\mathbf{x} = g(\boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}) \prod_{i=1}^{K} (x_{i,j_i} - x_{i,j_i - 1})\,. \quad \text{(A27)}$$

This allows us to rewrite the l.h.s. of (A25) as a unique summation:

$$\sum_{j_1, j_2, \ldots, j_K = 1}^{n} \left(g(\boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}) - g(\mathbf{c}_{j_1, j_2, \ldots, j_K}) \right) \prod_{i=1}^{K} (x_{i,j_i} - x_{i,j_i - 1}). \quad \text{(A28)}$$

In order to bound its absolute value, we use the triangular inequality, Cauchy–Schwartz inequality and the multivariate version of Lagrange's mean value theorem:

$$\begin{aligned} &\left| \sum_{j_1, j_2, \ldots, j_K = 1}^{n} \left(g(\boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}) - g(\mathbf{c}_{j_1, j_2, \ldots, j_K}) \right) \prod_{i=1}^{K} (x_{i,j_i} - x_{i,j_i - 1}) \right| \\ &\quad \le \sum_{j_1, j_2, \ldots, j_K = 1}^{n} \max_{\boldsymbol{\xi} \in Q_{j_1, j_2, \ldots, j_K}} \|\nabla g(\boldsymbol{\xi})\| \|\mathbf{c}_{j_1, j_2, \ldots, j_K} - \boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}\| \prod_{i=1}^{K} \left| x_{i,j_i} - x_{i,j_i - 1} \right| \quad \text{(A29)} \\ &\quad \le \max_{\boldsymbol{\xi} \in Q} \|\nabla g(\boldsymbol{\xi})\| \prod_{i=1}^{K} \epsilon_i \sum_{j_1, j_2, \ldots, j_K = 1}^{n} \|\mathbf{c}_{j_1, j_2, \ldots, j_K} - \boldsymbol{\tau}_{j_1, j_2, \ldots, j_K}\|\,. \end{aligned}$$

Let us briefly focus on the last factor, and let us consider only the summation with regards to, say, j_1:

$$\sum_{j_1=1}^{n} \|\mathbf{c}_{j_1,j_2,\dots,j_K} - \boldsymbol{\tau}_{j_1,j_2,\dots,j_K}\| \leq \sum_{i=1}^{K}\sum_{j_1=1}^{n} \left| c^{(i)}_{j_1,j_2,\dots,j_K} - \tau^{(i)}_{j_1,j_2,\dots,j_K} \right|$$
$$\leq \sum_{i=1}^{K}\sum_{j_1=1}^{n} \left| x^{(i)}_{j_1} - x^{(i)}_{j_1-1} \right| = \sum_{i=1}^{K} (b^{(i)} - a^{(i)}) \leq K \max_{i\leq K}(b^{(i)} - a^{(i)}) \quad \text{(A30)}$$

where we used the standard $L^2 - L^1$ norm inequality and $^{(i)}$ denotes the i-th component. Hence, the difference between the Riemann sum and the integral is controlled by

$$Kn^{K-1} \max_{i\leq K}(b^{(i)} - a^{(i)}) \max_{\boldsymbol{\xi}\in Q} \|\nabla g(\boldsymbol{\xi})\| \prod_{i=1}^{K} \epsilon_i \,. \quad \text{(A31)}$$

□

Remark A1. *Notice that if all $\epsilon_i = \frac{b^{(i)} - a^{(i)}}{n}$, then the above is still of order*

$$Nn^{K-1} \max_{i\leq K}(b^{(i)} - a^{(i)})^{K+1} \frac{1}{n^K} = \mathcal{O}\left(\frac{1}{n}\right) \quad \text{(A32)}$$

which means it still vanishes when the decomposition is fine enough ($n \to \infty$), and if the dimension K is not diverging.

Lemma A6 (Multivariate Laplace Approximation). *Let $f_N : Q \subset \mathbb{R}^K \longrightarrow \mathbb{R}$ be a differentiable sequence of function bounded away from the boundary of $Q = \bigtimes_{l=1}^{K} \left[\mu_{N,l} - N_l^{-\frac{1}{2}+\delta}, \mu_{N,l} + N_l^{-\frac{1}{2}+\delta}\right]$, satisfying $\nabla f_N(\boldsymbol{\mu}_N) = \mathbf{0}$ and $\mathcal{H}_{f_N}(\boldsymbol{\mu}_N) \prec 0$, such that $f_N(\boldsymbol{\mu}_N) > f_N(\mathbf{x})$ for all $\mathbf{x} \in Q$. Let $g(\mathbf{x})$ be analytic function in a neighborhood of $\boldsymbol{\mu}_N$, then for $\delta \in (0, \frac{1}{6})$ the following holds:*

$$\int_Q g(\mathbf{x}) e^{Nf_N(\mathbf{x})} d\mathbf{x} = \sqrt{\frac{(2\pi)^K}{\prod_{l=1}^{K} N_l \det(-\mathcal{H}_{f_N}(\boldsymbol{\mu}_N))}} g(\boldsymbol{\mu}_N) e^{Nf_N(\boldsymbol{\mu}_N)} (1 + \mathcal{O}(N^{-\frac{1}{2}+\delta})). \quad \text{(A33)}$$

Here, ∇f_N is the gradient vector.

Proof. Observe from the left hand side of Equation (A33) that

$$\mathcal{L} = \int_Q g(\mathbf{x}) e^{Nf_N(\mathbf{x})} d\mathbf{x} = \int_{\mu_{N,1} - N_1^{-\frac{1}{2}+\delta}}^{\mu_{N,1} + N_1^{-\frac{1}{2}+\delta}} \cdots \int_{\mu_{N,K} - N_K^{-\frac{1}{2}+\delta}}^{\mu_{N,K} + N_K^{-\frac{1}{2}+\delta}} g(x_1, \dots, x_K) e^{Nf_N(x_1,\dots,x_K)} dx_1 \cdots dx_K \,. \quad \text{(A34)}$$

Let us consider the following change of variables $t_l = \sqrt{N_l}(x_l - \mu_{N,l})$ for $l = 1, \dots, K$. Then, it follows that $x_l = \frac{t_l}{\sqrt{N_l}} + \mu_{N,l}$ and $dx_l = \frac{dt_l}{\sqrt{N_l}}$. Now, the bounds on $\mathbf{x} \in Q$ become $t_l \in [-N_l^{\delta}, N_l^{\delta}]$ and the integral transforms into

$$\mathcal{L} = \int_{-N_1^{\delta}}^{N_1^{\delta}} \cdots \int_{-N_K^{\delta}}^{N_K^{\delta}} g\left(\frac{t_1}{\sqrt{N_1}} + \mu_{N,1}, \dots, \frac{t_K}{\sqrt{N_K}} + \mu_{N,K}\right) e^{Nf_N\left(\frac{t_1}{\sqrt{N_1}} + \mu_{N,1}, \dots, \frac{t_K}{\sqrt{N_K}} + \mu_{N,K}\right)} \prod_{l=1}^{K} \frac{dt_l}{\sqrt{N_l}}. \quad \text{(A35)}$$

By an application of Taylor expansion of f_N and g around the vector $\boldsymbol{\mu}_N$, we have that

$$e^{Nf_N\left(\frac{t_1}{\sqrt{N_1}}+\mu_{N,1},\dots,\frac{t_K}{\sqrt{N_K}}+\mu_{N,K}\right)} = e^{Nf_N(\boldsymbol{\mu}_N)+\frac{1}{2}(\mathbf{t},\mathcal{H}_{f_N}(\boldsymbol{\mu}_N)\mathbf{t})}\left(1+\mathcal{O}\left(N\left(\frac{N^\delta}{\sqrt{N}}\right)^3\right)\right) \quad \text{and}$$

$$\begin{aligned} g\left(\frac{t_1}{\sqrt{N_1}}+\mu_{N,1},\dots,\frac{t_K}{\sqrt{N_K}}+\mu_{N,K}\right) &= g(\boldsymbol{\mu}_N)+\nabla g(\boldsymbol{\mu}_N)\left(\frac{N^\delta}{\sqrt{N}}\right) \\ &= g(\boldsymbol{\mu}_N)\left(1+\mathcal{O}\left(N^{\delta-1/2}\right)\right). \end{aligned} \quad \text{(A36)}$$

Now, following from (A36), the right side of (A35) becomes

$$\begin{aligned} \mathcal{L} &= \prod_{l=1}^K \frac{1}{\sqrt{N_l}}\left(1+\mathcal{O}\left(N^{\delta-1/2}\right)\right)\left(1+\mathcal{O}\left(N^{3\delta-1/2}\right)\right)g(\boldsymbol{\mu}_N)e^{Nf_N(\boldsymbol{\mu}_N)}\int_{-N_1^\delta}^{N_1^\delta}\cdots\int_{-N_K^\delta}^{N_K^\delta} e^{\frac{1}{2}(\mathbf{t},\mathcal{H}_{f_N}(\boldsymbol{\mu}_N)\mathbf{t})}d\mathbf{t} \\ &= \left(1+\mathcal{O}\left(N^{3\delta-1/2}\right)\right)\sqrt{\frac{(2\pi)^K}{\prod_{l=1}^K N_l \det\left(-\mathcal{H}_{f_N}(\boldsymbol{\mu}_N)\right)}}g(\boldsymbol{\mu}_N)e^{Nf_N(\boldsymbol{\mu}_N)}. \end{aligned} \quad \text{(A37)}$$

Notice that we have bounded t_l by its limit N_l^δ. This completes the proof of Lemma A6. □

References

1. Kac, M. MATHEMATICAL MECHANISMS OF PHASE TRANSITIONS. In *Statistical Physics, Phase Transitions, and Superfluidity*; Chretien, M., Gross, E.P., Deser, S., Eds.; Gordon and Breach, Science Publishers: New York, NY, USA, 1968; Volume I, pp. 241–305.
2. Kirsch, W.; Langner, J., The fate of the square root law for correlated voting. In *Studies in Choice and Welfare*; Springer International Publishing: Cham, Switzerland, 2014; pp. 147–158.
3. Kirsch, W., A mathematical view on voting and power. In *Mathematics and Society*; EMS Press: Zuerich, Switzerland, 2016; pp. 251–279.
4. Gsänger, M.; Hösel, V.; Mohamad-Klotzbach, C.; Müller, J. Opinion models, election data, and political theory. *Entropy* **2024**, *26*, 212. [CrossRef] [PubMed]
5. Brock, W.A.; Durlauf, S.N. Discrete choice with social interactions. *Rev. Econ. Stud.* **2001**, *68*, 235–260. [CrossRef]
6. Blume, L.; Durlauf, S. Equilibrium concepts for social interaction models. *Int. Game Theory Rev.* **2003**, *5*, 193–209. [CrossRef]
7. Marsman, M.; Tanis, C.C.; Bechger, T.M.; Waldorp, L.J., Network psychometrics in educational practice: Maximum likelihood estimation of the Curie-Weiss model. In *Theoretical and Practical Advances in Computer-based Educational Measurement*; Springer International Publishing: Cham, Switzerland, 2019; pp. 93–120.
8. Contucci, P.; Ghirlanda, S. Modelling society with statistical mechanics: An application to cultural contact and immigration. *Qual. Quant.* **2007**, *41*, 569–578. [CrossRef]
9. Burioni, R.; Contucci, P.; Fedele, M.; Vernia, C.; Vezzani, A. Enhancing participation to health screening campaigns by group interactions. *Sci. Rep.* **2015**, *5*, 9904. [CrossRef]
10. Opoku, A.A.; Osabutey, G.; Kwofie, C. Parameter evaluation for a statistical mechanical model for binary choice with social interaction. *J. Probab. Stat.* **2019**, *2019*, 3435626. [CrossRef]
11. Frøyen, S.; Sudbø, A.S.; Hemmer, P. Ising models with two-and three-spin interactions: Mean field equation of state. *Phys. A Stat. Mech. Its Appl.* **1976**, *85*, 399–408. [CrossRef]
12. Bidaux, R.; Boccara, N.; Forgàcs, G. Three-spin interaction Ising model with a nondegenerate ground state at zero applied field. *J. Stat. Phys.* **1986**, *45*, 113–134. [CrossRef]
13. Kincaid, J.M.; Cohen, E.G.D. Phase diagrams of liquid helium mixtures and metamagnets: Experiment and mean field theory. *Phys. Rep.* **1975**, *22*, 57–143. [CrossRef]
14. Galam, S.; Yokoi, C.S.O.; Salinas, S.R. Metamagnets in uniform and random fields. *Phys. Rev. B* **1998**, *57*, 8370–8374. [CrossRef]
15. Gallo, I.; Contucci, P. Bipartite mean field spin systems. Existence and solution. *Math. Phys. Electron. J.* **2008**, *14*, 1–21.
16. Genovese, G.; Barra, A. A certain class of Curie-Weiss models. *arXiv* **2009**, arXiv:0906.4673.
17. Barra, A.; Genovese, G.; Guerra, F. Equilibrium statistical mechanics of bipartite spin systems. *J. Phys. A Math. Theor.* **2011**, *44*, 245002. [CrossRef]
18. Genovese, G.; Tantari, D. Non-convex multipartite ferromagnets. *J. Stat. Phys.* **2016**, *163*, 492–513. [CrossRef]
19. Fedele, M.; Contucci, P. Scaling limits for multi-species statistical mechanics mean-field models. *J. Stat. Phys.* **2011**, *144*, 1186–1205. [CrossRef]

20. Ellis, R.S.; Newman, C.M. The statistics of Curie-Weiss models. *J. Stat. Phys.* **1978**, *19*, 149–161. [CrossRef]
21. Löwe, M.; Schubert, K. Fluctuations for block spin Ising models. *Electron. Commun. Probab.* **2018**, *23*, 1–12. [CrossRef]
22. Knöpfel, H.; Löwe, M.; Schubert, K.; Sinulis, A. Fluctuation results for general block spin Ising models. *J. Stat. Phys.* **2020**, *178*, 1175–1200. [CrossRef]
23. Kirsch, W.; Toth, G. Two groups in a Curie-Weiss model. *Math. Phys. Anal. Geom.* **2020**, *23*, 17. [CrossRef]
24. Fleermann, M.; Kirsch, W.; Toth, G. Local central limit theorem for multi-group Curie–Weiss models. *J. Theor. Probab.* **2022**, *35*, 2009–2019. [CrossRef]
25. Kirsch, W.; Toth, G. Limit theorems for multi-group Curie–Weiss models via the method of moments. *Math. Phys. Anal. Geom.* **2022**, *25*, 24. [CrossRef]
26. Barra, A.; Contucci, P.; Mingione, E.; Tantari, D. Multi-species mean field spin glasses. Rigorous results. In *Proceedings of the Annales Henri Poincaré*; Springer: Basel, Switzerland, 2015; Volume 16, pp. 691–708. [CrossRef]
27. Panchenko, D. The free energy in a multi-species Sherrington–Kirkpatrick model. *Ann. Probab.* **2015**, *43*, 3494–3513. [CrossRef]
28. Baik, J.; Lee, J.O. Free energy of bipartite spherical Sherrington–Kirkpatrick model. *Ann. Inst. Henri Poincaré* **2020**, *56*, 2897–2934. [CrossRef]
29. Subag, E. TAP approach for multispecies spherical spin glasses II: The free energy of the pure models. *Ann. Probab.* **2023**, *51*, 1004–1024. [CrossRef]
30. Alberici, D.; Camilli, F.; Contucci, P.; Mingione, E. The Multi-species Mean-Field Spin-Glass on the Nishimori Line. *J. Stat. Phys.* **2021**, *182*, 2. [CrossRef]
31. Alberici, D.; Camilli, F.; Contucci, P.; Mingione, E. The Solution of the Deep Boltzmann Machine on the Nishimori Line. *Commun. Math. Phys.* **2021**, *387*, 1191–1214. [CrossRef]
32. Mourrat, J.C. Nonconvex interactions in mean-field spin glasses. *arXiv* **2020**, arXiv:2004.01679.
33. Hong-Bin Chen, J.C.M. On the free energy of vector spin glasses with non-convex interactions. *arXiv* **2023**, arXiv:2311.08980.
34. Guerra, F.; Lucio Toninelli, F. Central limit theorem for fluctuations in the high temperature region of the Sherrington–Kirkpatrick spin glass model. *J. Math. Phys.* **2002**, *43*, 6224–6237. [CrossRef]
35. Guerra, F.; Toninelli, F.L. Quadratic replica coupling in the Sherrington–Kirkpatrick mean field spin glass model. *J. Math. Phys.* **2002**, *43*, 3704–3716. [CrossRef]
36. Talagrand, M. *Spin Glasses: A Challenge for Mathematicians: Cavity and Mean Field Models*; Springer: Berlin/Heidelberg, Germany; Springer Science & Business Media: New York, NY, USA, 2003; Volume 46.
37. Camilli, F.; Contucci, P.; Mingione, E. Central limit theorem for the overlaps on the Nishimori line. *arXiv* **2023**, arXiv:2305.19943.
38. Dey, P.; Wu, Q. Fluctuation Results for Multi-species Sherrington-Kirkpatrick Model in the Replica Symmetric Regime. *J. Stat. Phys.* **2021**, *185*. [CrossRef]
39. Zimmaro, F.; Galam, S.; Javarone, M.A. Asymmetric games on networks: Mapping to Ising models and bounded rationality. *Chaos Solitons Fractals* **2024**, *181*, 114666. [CrossRef]
40. Agliari, E.; Migliozzi, D.; Tantari, D. Non-convex multi-species Hopfield models. *J. Stat. Phys.* **2018**, *172*, 1247–1269. [CrossRef]
41. Contucci, P.; Vernia, C. On a statistical mechanics approach to some problems of the social sciences. *Front. Phys.* **2020**, *8*, 585383. [CrossRef]
42. Ignacio, G.; Barra, A.; Contucci, P. Parameter evaluation of a simple mean-field model of social interaction. *Math. Models Methods Appl. Sci* **2009**, *19*, 1427–1439.
43. Mukherjee, S.; Son, J.; Bhattacharya, B.B. Fluctuations of the magnetization in the p-spin Curie-Weiss model. *Comm. Math. Phys.* **2021**, *387*, 681–728. [CrossRef]
44. Contucci, P.; Mingione, E.; Osabutey, G. Limit theorems for the cubic mean-field Ising model. *Ann. Henri Poincare* **2024**, *25*, 5019–5044. [CrossRef]
45. Dembo, A.; Zeitouni, O. *Large Deviations Techniques and Applications*; Springer: Berlin, Germany, 2010.
46. Opoku, A.A.; Osabutey, G. Multipopulation spin models: A view from large deviations theoretic window. *J. Math.* **2018**, *2018*, 1–13. [CrossRef]
47. Eisele, T.; Ellis, R.S. Multiple phase transitions in the generalized Curie-Weiss model. *J. Stat. Phys.* **1988**, *52*, 161–202. [CrossRef]
48. Collet, F. Macroscopic limit of a bipartite Curie–Weiss model: A dynamical approach. *J. Stat. Phys.* **2014**, *157*, 1301–1319. [CrossRef]
49. Fedele, M.; Unguendoli, F. Rigorous results on the bipartite mean-field model. *J. Phys. A Math. Theor.* **2012**, *45*, 385001. [CrossRef]
50. Zhao, K.; Bianconi, G. Percolation on interdependent networks with a fraction of antagonistic interactions. *J. Stat. Phys.* **2013**, *152*, 1069–1083. [CrossRef]
51. Andreis, L.; Tovazzi, D. Coexistence of stable limit cycles in a generalized Curie–Weiss model with dissipation. *J. Stat. Phys.* **2018**, *173*, 163–181. [CrossRef]
52. Ayi, N.; Pouradier Duteil, N. Mean-field and graph limits for collective dynamics models with time-varying weights. *J. Differ. Equations* **2021**, *299*, 65–110. [CrossRef]

53. Giacomin, G.; Poquet, C. Noise, interaction, nonlinear dynamics and the origin of rhythmic behaviors. Brazilian Jounal of Probab. *Stat* **2015**, *29*, 460–493.
54. Collet, F.; Formentin, M.; Tovazzi, D. Rhythmic behavior in a two-population mean-field Ising model. *Phys. Rev. E* **2016**, *94*, 042139. [CrossRef]
55. Salwinski, D. The Continuous Binomial Coefficient: An Elementary Approach. *Am. Math. Mon.* **2018**, *125*, 231–244. [CrossRef]