On the mixed even-spin Sherrington-Kirkpatrick model with ferromagnetic interaction

Wei-Kuo Chen*

July 13, 2012

Abstract

We study a spin system with both mixed even-spin Sherrington-Kirkpatrick (SK) couplings and Curie-Weiss (CW) interaction. Our main results are: (i) The thermodynamic limit of the free energy is given by a variational formula involving the free energy of the SK model with a change in the external field. (ii) In the presence of a centered Gaussian external field, the positivity of the overlap and the extended Ghirlanda-Guerra identities hold on a dense subset of the temperature parameters. (iii) We establish a general inequality between the magnetization and overlap. (iv) We construct a temperature region in which the magnetization can be quantitatively controlled and deduce different senses of convergence for the magnetization depending on whether the external field is present or not. Our approach is based on techniques from the study of the CW and SK models and results in convex analysis.

Keywords: Ferromagnetic interaction, Ghirlanda-Guerra identities, Parisi formula, Sherrington-Kirkpatrick model, Ultrametricity

1 Introduction

The Sherrington-Kirkpatrick (SK) model formulated by Sherrington and Kirkpatrick [18] is one of the most important mean field spin glasses with the aim of understanding strange magnetic properties of certain alloys. In the recent decades, many essential conjectures proposed by physicists have been intensively studied in the mathematical community, including the validity of the Parisi formula and the ultrametricity of the overlap. In this paper, we are interested in the SK model coupled with the familiar Curie-Weiss (CW) ferromagnetic interaction. There have been a few studies of this model so far [3, 6, 25] (one may also refer to [7] for a much more difficult coupling, the SK model with Ising interaction). However, rigorous results are very limited and mainly restricted to the high temperature regime. The main reason is that in this case, the effect of the (non random) ferromagnetic interaction can be linearly approximated. The model

*Department of Mathematics, University of California at Irvine, 340 Rowland Hall, Irvine, CA 92697-3875, USA, email: weikuoc@uci.edu

then becomes essentially the SK model with a slight perturbation on the external field. Therefore as one might expect, in the high temperature region, many properties of the SK model are also valid in our model. Indeed, following the same arguments as [23] or [24], one can prove (see [3]) that for this model, the thermodynamic limit of the free energy exists, the magnetization and overlap in the limit concentrate on a singleton, the central limit theorem for the free energy holds, and the Thouless-Anderson-Palmer system of equations is valid.

We will be concerned with the more general mean field model with both the mixed even-spin SK couplings and ferromagnetic interaction (SKFI) and address the following questions: (i) How can one compute the thermodynamic limit of the free energy of the SKFI model? (ii) Which properties does the SKFI model inherit from the CW and SK models? (iii) Is there any general relation between the magnetization and overlap? (iv) Can one give a quantitative control on the magnetization or the overlap in general? Our answers to these problems will be stated in Section 2 and will also cover the situation in the low temperature regime.

Let us now give the description of the SKFI model, which depends on two quantities: the (inverse) temperature parameter $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ and external field h, where $\mathcal{B} = \{(\beta, (\beta_p)_{p\geq 1}) : \beta \geq 0,\ \sum_{p\geq 1} 2^p \beta_p^2 < \infty\}$ and h is a Gaussian random variable (possibly degenerate). One may think of β as the temperature for the CW interaction and $\boldsymbol{\beta}$ as the temperature for the SK couplings. Let us emphasize that β_p may take negative values, while since we are concerned with the ferromagnetic interaction, the CW temperature β only takes nonnegative values. For each positive integer N, set the configuration space $\Sigma_N = \{-1, +1\}^N$. Let $(h_i)_{i\leq N}$ be i.i.d. copies of h. For a given temperature $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ and external field h, the SKFI model has Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \frac{\beta N}{2} m(\boldsymbol{\sigma})^2 + H_N^{SK}(\boldsymbol{\sigma}) + \sum_{i\leq N} h_i \sigma_i, \tag{1.1}$$

where the quantity $m = m(\boldsymbol{\sigma}) := N^{-1} \sum_{i\leq N} \sigma_i$ is called the magnetization per site. Here, H_N^{SK} is the mixed even p-spin interactions for the SK model, that is,

$$H_N^{SK}(\boldsymbol{\sigma}) = \sum_{p\geq 1} \frac{\beta_p}{N^{p-1/2}} \sum_{1\leq i_1,\ldots,i_{2p}\leq N} g_{i_1,\ldots,i_{2p}} \sigma_{i_1} \cdots \sigma_{i_{2p}}, \tag{1.2}$$

where $\boldsymbol{g} = (g_{i_1,\ldots,i_{2p}} : 1 \leq i_1, \ldots, i_{2p} \leq N, p \geq 1)$ are i.i.d. standard Gaussian r.v.s independent of $(h_i)_{i\leq N}$. One may see easily that the covariance of H_N^{SK} is a function of the overlap $R_{1,2} = R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) := N^{-1} \sum_{i\leq N} \sigma_i^1 \sigma_i^2$ through $\mathbb{E} H_N^{SK}(\boldsymbol{\sigma}^1) H_N^{SK}(\boldsymbol{\sigma}^2) = N\xi(R_{1,2})$, where $\xi(x) := \sum_{p\geq 1} \beta_p^2 x^{2p}$.

We define the partition function, Gibbs measure, and free energy for the SKFI model, respectively, by $Z_N = Z_N(\beta, \boldsymbol{\beta}, h) = \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp H_N(\boldsymbol{\sigma})$, $G_N(\boldsymbol{\sigma}) = \exp H_N(\boldsymbol{\sigma})/Z_N$, and $F_N = F_N(\beta, \boldsymbol{\beta}, h) = N^{-1}\mathbb{E} \ln Z_N$. We will use $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$, etc. to denote the *replicas* sampled independently from G_N. For any real-valued function f on Σ_N^n, we define its Gibbs average corresponding to the Gibbs measure G_N as

$$\langle f \rangle = \sum_{\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n \in \Sigma_N} f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) G_N(\boldsymbol{\sigma}^1) \cdots G_N(\boldsymbol{\sigma}^n).$$

In the case of $\beta = 0$, our model is known as the mixed even p-spin SK model (see [24]) and we will use Z_N^{SK}, G_N^{SK}, F_N^{SK}, and $\langle\cdot\rangle^{SK}$ to denote its partition function, Gibbs measure, free energy, and Gibbs average, respectively. On the other hand, if $\boldsymbol{\beta} = \mathbf{0}$, our model reduces to the Curie-Weiss (CW) model and Z_N^{CW}, G_N^{CW}, F_N^{CW}, and $\langle\cdot\rangle^{CW}$ are also defined in the same manner.

2 Main results

Our main results will be stated in this section. Proofs are deferred to Section 3. Throughout the paper, we will use $I(E)$ to denote the indicator function for the event E.

2.1 The thermodynamic limit of the free energy

Let us begin by illustrating the different natures of the SKFI, CW, and SK models. The SK model has been widely studied, see, for example, [18] and [19] for details. By an application of the Gaussian interpolation technique, Guerra and Toninelli proved [8] that $\{\mathbb{E}\ln Z_N^{SK}\}_{N\geq 1}$ is superadditive and as a consequence

$$F^{SK}(\boldsymbol{\beta}, h) := \lim_{N\to\infty} F_N^{SK}(\boldsymbol{\beta}, h) \tag{2.1}$$

exists. Using Jensen's inequality, it is easy to see that $\{\mathbb{E}\ln Z_N^{CW}\}_{N\geq 1}$ is subadditive, which ensures the existence of the thermodynamic limit of the free energy for the CW model. However, if $\beta \neq 0$ and $\boldsymbol{\beta} \neq \mathbf{0}$, neither superadditivity nor subadditivity obviously holds for $\{\mathbb{E}\ln Z_N\}_{N\geq 1}$ in the SKFI model. The existence of the thermodynamic limit of the free energy for the SKFI model was firstly shown in [9]. Our first main result regarding the formula of the thermodynamic limit of the free energy for the SKFI model is stated as follows:

Theorem 1. *For any $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$, we have*

$$F(\beta, \boldsymbol{\beta}, h) := \lim_{N\to\infty} F_N(\beta, \boldsymbol{\beta}, h) = \max_{\mu\in[-1,1]} \left\{ F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2} \right\}. \tag{2.2}$$

For any given $(\beta, \boldsymbol{\beta})$ and h, we set

$$\Omega = \Omega(\beta, \boldsymbol{\beta}, h) = \mathrm{Argmax}_{\mu\in[-1,1]} \left\{ F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2} \right\}. \tag{2.3}$$

The following proposition says that the magnetization is essentially supported on $\Omega(\beta, \boldsymbol{\beta}, h)$.

Proposition 1. *For any open subset U of $[-1, 1]$ with*

$$\inf\{|x - y| : x \in U, y \in \Omega\} > 0,$$

we have for every N,

$$\mathbb{E}\langle I(m \in U)\rangle \leq K \exp\left(-\frac{N}{K}\right), \tag{2.4}$$

where K is a constant independent of N. In particular,

$$\lim_{N\to\infty} \langle I(m \in U)\rangle = 0 \quad a.s. \tag{2.5}$$

2.2 Positivity of the overlap

In the SK model with external field, Talagrand [24] proved that the overlap is essentially greater than a positive constant with high probability, and deduced from this fact that the extended Ghirlanda-Guerra identities hold. In this section, we prove that these results are "typically" valid in the SKFI model.

Before we state our main results, let us recall the formulation of the Parisi formula and some known results regarding the differentiability of the Parisi measure. Let $\mathcal{M}_0$ be the collection of probability measures on $[0,1]$ that consist of a finite number of point masses. For each $\nu \in \mathcal{M}_0$, we consider a function $\Phi_\nu(x,q)$ defined on $\mathbb{R} \times [0,1]$ with $\Phi_\nu(x,1) = \ln\cosh(x)$ and satisfying the PDE

$$\frac{\partial \Phi_\nu}{\partial q} = -\frac{1}{2}\xi''(q)\left(\frac{\partial^2 \Phi_\nu}{\partial x^2} + \nu([0,q])\left(\frac{\partial \Phi_\nu}{\partial x}\right)^2\right). \tag{2.6}$$

The Parisi formula states that the thermodynamic limit of the free energy of the SK model with temperature $\boldsymbol{\beta}$ and external field h can be represented as

$$F^{SK}(\boldsymbol{\beta}, h) = \inf_{\nu \in \mathcal{M}_0} \mathcal{P}(\boldsymbol{\beta}, h, \nu), \tag{2.7}$$

where

$$\mathcal{P}(\boldsymbol{\beta}, h, \nu) := \ln 2 + \mathbb{E}\Phi_\nu(h,0) - \frac{1}{2}\theta(1) + \frac{1}{2}\int_0^1 \theta(q)\nu(dq), \quad \nu \in \mathcal{M}_0$$

and $\theta(q) := q\xi'(q) - \xi(q)$. The validity of this formula was firstly verified in the work of Talagrand [20] and was later extended to the general mixed p-spin SK model [15] and the spherical SK model [4, 22]. Let $\mathcal{M}$ be the space of all probability measures on $[0,1]$ endowed with the weak topology. Since $\mathcal{P}(\boldsymbol{\beta}, h, \cdot)$ is Lipschitz with respect to the metric (see [10] and [21]):

$$d(\nu_1, \nu_2) := \int_0^1 |\nu_1([0,q]) - \nu_2([0,q])| dq, \quad \nu_1, \nu_2 \in \mathcal{M}_0, \tag{2.8}$$

$\mathcal{P}$ can be extended continuously to $\mathcal{M}$. From the compactness of $\mathcal{M}$, the infimum (2.7) is achieved and any $\nu \in \mathcal{M}$ that achieves the infimum is called a Parisi measure. Arguments of [12] and [21] imply the differentiability of the Parisi formula

$$\frac{\partial}{\partial \beta_p} F^{SK}(\boldsymbol{\beta}, h) = \beta_p\left(1 - \int_0^1 q^{2p}\nu_{\boldsymbol{\beta},h}(dq)\right) \tag{2.9}$$

and give the moment computation for $|R_{1,2}|$ via

$$\lim_{N\to\infty} \mathbb{E}\left\langle R_{1,2}^{2p}\right\rangle^{SK} = \int_0^1 q^{2p}\nu_{\boldsymbol{\beta},h}(dq) \tag{2.10}$$

provided $\beta_p \neq 0$, where $\nu_{\boldsymbol{\beta},h}$ is a Parisi measure. In the case of $\beta_p \neq 0$ for all $p \geq 1$, (2.10) implies that $\nu_{\boldsymbol{\beta},h}$ is the limiting distribution of $|R_{1,2}|$; if, in addition, h is nondegenerate, it is well-known that the Parisi measure takes nonnegative values and, again, from (2.10),

the Parisi measure is the limiting distribution of the overlap $R_{1,2}$ (see Chapter 14 of [24] for detailed discussions).

Recall the set Ω from (2.3). Let us denote by $\mathcal{B}_d$ the collection of all $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ that satisfy $\beta > 0$ and

$$\text{either } |\Omega(\beta, \boldsymbol{\beta}, h)| = 1 \text{ or } \Omega(\beta, \boldsymbol{\beta}, h) = \{\mu, -\mu\} \text{ for some } 0 < \mu < 1. \tag{2.11}$$

The following proposition gives the connection between the set $\mathcal{B}_d$ and the differentiability of $F(\beta, \boldsymbol{\beta}, h)$ with respect to β.

Proposition 2. *Suppose that $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ with $\beta > 0$. Then $\frac{\partial F}{\partial \beta}(\beta, \boldsymbol{\beta}, h)$ exists if and only if $(\beta, \boldsymbol{\beta}) \in \mathcal{B}_d$.*

Notice that since $(\beta, \boldsymbol{\beta}) \mapsto F(\beta, \boldsymbol{\beta}, h)$ is a continuous convex function on the space of all $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ with $\beta > 0$, which is open in the separable Banach space $\{(\beta, \boldsymbol{\beta}) : \beta^2 + \sum_{p \geq 1} 2^p \beta_p^2 < \infty\}$ endowed with the norm $|(\beta, \boldsymbol{\beta})| = (\beta^2 + \sum_{p \geq 1} 2^p \beta_p^2)^{1/2}$, it follows, by Mazur's theorem (Theorem 1.20 [11]), that the set where $F(\cdot, \cdot, h)$ is Gâteaux-differentiable (in the sense that the directional derivative exists in all directions) is a dense G_δ set in $\mathcal{B}$ contained in $\mathcal{B}_d$. This means that typically the magnetization concentrates either on a singleton or two distinct values, which are symmetric with respect to the origin. This property coincides with the behavior of the CW model.

We prove that analogues of (2.9) and (2.10) also hold for the SKFI model in certain temperature region. For technical purposes, we assume that h is centered. Let us denote by $\mathcal{B}'$ the collection of all $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$ with $\beta > 0$ and $\beta_p \neq 0$ for all $p \geq 1$. Set $\mathcal{B}'_d = \mathcal{B}_d \cap \mathcal{B}'$. Notice that $\mathcal{B}'_d$ is a G_δ subset in $\mathcal{B}'$ and that, concluding from the convexity of $F(\cdot, \boldsymbol{\beta}, h)$ for every fixed $(\boldsymbol{\beta}, h)$, $F(\cdot, \boldsymbol{\beta}, h)$ is differentiable for all but countably many β. Thus, using Proposition 2, $\mathcal{B}'_d$ forms a dense G_δ subset in $\mathcal{B}'$.

Theorem 2. *If $(\beta, \boldsymbol{\beta}) \in \mathcal{B}'_d$, then for every $p \geq 1$, we have*

$$\frac{\partial F}{\partial \beta_p}(\beta, \boldsymbol{\beta}, h) = \beta_p \left(1 - \int_0^1 q^{2p} \nu_{\boldsymbol{\beta}, \beta\mu + h}(dq)\right) \tag{2.12}$$

and

$$\lim_{N \to \infty} \mathbb{E} \left\langle R_{1,2}^{2p} \right\rangle = \int_0^1 q^{2p} \nu_{\boldsymbol{\beta}, \beta\mu + h}(dq), \tag{2.13}$$

where $\nu_{\boldsymbol{\beta}, \beta\mu + h}$ is the unique Parisi measure for the SK model with temperature $\boldsymbol{\beta}$, external field $\beta\mu + h$, and $\mu \in \Omega(\beta, \boldsymbol{\beta}, h)$.

From (2.12), it means that the limiting distribution of $|R_{1,2}|$ is determined by the Parisi measure $\nu_{\boldsymbol{\beta}, \beta\mu + h}$. If $\mathbb{E}h^2 \neq 0$, we will prove that $\nu_{\boldsymbol{\beta}, \beta\mu + h}$ gives the limiting distribution of the overlap $R_{1,2}$ relying on Talagrand's positivity of the overlap in the SK model. The precise statement of the latter is described as follows: Let $\mathbb{E}h^2 \neq 0$. Consider a Parisi measure ν and the smallest point c in the support of ν. Then $c > 0$ and for any $c' < c$, there exists some constant K independent of N such that

$$\mathbb{E} \left\langle I \left(R_{1,2} \leq c'\right)\right\rangle^{SK} \leq K \exp\left(-\frac{N}{K}\right). \tag{2.14}$$

As for the SKFI model, we have a weaker version of Talagrand's positivity.

Theorem 3. *Suppose that $\mathbb{E}h^2 \neq 0$. Let $(\beta, \boldsymbol{\beta}) \in \mathcal{B}'_d$ and $\nu_{\boldsymbol{\beta},\beta\mu+h}$ be the Parisi measure of the SK model stated in Theorem 2. Suppose that c is the smallest value in the support of $\nu_{\boldsymbol{\beta},\beta\mu+h}$. Then $c > 0$ and for every $0 < c' < c$, we have*

$$\lim_{N\to\infty} \mathbb{E}\left\langle I\left(R_{1,2} \leq c'\right)\right\rangle = 0 \tag{2.15}$$

and for every continuous function f on $[-1, 1]$,

$$\lim_{N\to\infty} \mathbb{E}\left\langle f(R_{1,2})\right\rangle = \int_0^1 f(q)\nu_{\boldsymbol{\beta},\beta\mu+h}(dq). \tag{2.16}$$

The equation (2.15) implies that in our model the overlap is greater than or equal to a positive constant c with high probability, mirroring the same phenomenon in the SK model with Gaussian external field. More importantly, (2.16) means that the limiting law of the overlap of the SKFI model is the same as that of the SK model with a shifted external field $\beta\mu + h$.

Proposition 3. *Let $(\beta, \boldsymbol{\beta}) \in \mathcal{B}'_d$. If $\mathbb{E}h^2 \neq 0$, then the sequence (G_N) of Gibbs measures of the SKFI model satisfies the extended Ghirlanda-Guerra (EGG) identities, that is, for each n and each continuous function ψ on $\mathbb{R}$, we have*

$$\lim_{N\to\infty} \sup_f \left| n\mathbb{E}\left\langle \psi(R_{1,n+1})f\right\rangle - \mathbb{E}\left\langle \psi(R_{1,2})\right\rangle \mathbb{E}\left\langle f\right\rangle - \sum_{2\leq l\leq n} E\left\langle \psi(R_{1,l})f\right\rangle \right| = 0, \tag{2.17}$$

where the supremum is taken over all (non random) functions f on Σ_N^n with $|f| \leq 1$.

These identities were firstly discovered by Ghirlanda and Guerra in the context of the SK model with $\psi(x) = x$. Later, they were generalized to the mixed p-spin SK models and also mixed p-spin spherical SK models, see Chapter 12 [24] for details. The importance of the EGG identities are due to the conjecture that they yield the ultrametric property of the overlaps, that is, under the Gibbs measure, the event

$$R_{1,2} \geq \min(R_{1,3}, R_{2,3}) \tag{2.18}$$

has probability nearly one. This conjecture was recently confirmed by Panchenko [14]. Thus, from the Baffionni-Rosati theorem, the limiting behavior of the Gibbs measure can be characterized by the Poisson-Dirichlet cascades, which is closely related to the replica symmetry breaking scheme in the computation of the Parisi formula, see Chapter 15 [24]. For the applications of the EGG identities, the readers are referred to [5] and [13].

In the same fashion, Proposition 3 implies the ultrametric structure (2.18) of the overlap in the SKFI model. The proof of Proposition 3 is based on the concentration of the Hamiltonian and the positivity of the overlap. As the argument has been explained in great detail in Chapter 12 in [24], the proof will be omitted in this paper. We will present an immediate application of the EGG identities in Theorem 4 below that yields a general inequality between the magnetization and overlap.

2.3 An inequality between the magnetization and overlap

In this section, we present an inequality between the magnetization and overlap. Again, for technical purposes, we assume that the external field h is centered throughout this section. Let us first motivate our idea by considering the original SK model ($\beta_p = 0$ for every $p \geq 2$) with ferromagnetic interaction. It is well-known that in this case the magnetization and overlap of the SKFI model in the high temperature regime (β_1 and β are very small) are concentrated essentially at single values in the sense that

$$\mathbb{E}\left\langle (m-\mu)^{2k} \right\rangle \leq \frac{K}{N^k}$$
$$\mathbb{E}\left\langle (R_{1,2}-q)^{2k} \right\rangle \leq \frac{K}{N^k}$$

for every $k \geq 1$, where K is a constant independent of N and (μ, q) is the unique solution to

$$\mu = \mathbb{E}\tanh(\beta_1 z\sqrt{2q} + \beta\mu + h)$$
$$q = \mathbb{E}\tanh^2(\beta_1 z\sqrt{2q} + \beta\mu + h)$$

for some standard Gaussian r.v. z independent of h. For the proof, one may follow the same argument as [3]. As one can see immediately from the Cauchy-Schwarz inequality, $\mu^2 \leq q$, that is, the overlap is essentially bounded from below by the square of the magnetization. It is natural to ask whether in general a similar relation between the magnetization and overlap holds or not. Using the fundamental property (2.11) of the magnetization and the EGG identities for the overlaps, we will prove that the answer is in the affirmative. Recall that $\mathcal{B}'_d$ is a dense G_δ set in $\mathcal{B}'$ and from (2.4) and (2.11), if $(\beta, \boldsymbol{\beta}) \in \mathcal{B}'_d$, there exists some $0 \leq \mu < 1$ such that

$$\lim_{N\to\infty} \mathbb{E}\left\langle I(||m|-\mu| \leq \varepsilon)\right\rangle = 1 \tag{2.19}$$

for all $\varepsilon > 0$. Our first main result is stated as follows.

Theorem 4. *Let $(\beta, \boldsymbol{\beta}) \in \mathcal{B}'_d$. We have that*

1. *if $\mathbb{E}h^2 = 0$, $\lim_{N\to\infty} \mathbb{E}\left\langle I(\mu^2 - \varepsilon \leq |R_{1,2}|)\right\rangle = 1$ for every $\varepsilon > 0$;*
2. *if $\mathbb{E}h^2 \neq 0$, $\lim_{N\to\infty} \mathbb{E}\left\langle I(\mu^2 - \varepsilon \leq R_{1,2})\right\rangle = 1$ for every $\varepsilon > 0$.*

In other words, μ^2 provides a lower bound for the support of the Parisi measure $\nu_{\boldsymbol{\beta},\beta\mu+h}$. From (2.19), Theorem 4 also means that for $(\boldsymbol{\sigma}, \boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ sampled from $EG_N^{\otimes 3}$, essentially $m(\boldsymbol{\sigma})^2 \leq |R_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)|$ if $\mathbb{E}h^2 = 0$ and $m(\boldsymbol{\sigma})^2 \leq R_{1,2}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)$ if $\mathbb{E}h^2 \neq 0$.

2.4 A quantitative control on the magnetization

We will construct a temperature region where the effect of the ferromagnetic interaction is much stronger than the effect of the mixed even p-spin interactions. In this region, we can control the magnetization quantitatively away from the origin and deduce different senses

of convergence of the magnetization depending on whether the external field is present or not. Suppose, throughout this section, that the external field h is centered satisfying

$$\mathbb{E}e^{2|h|} < \frac{1}{\max_{\beta\geq 0}\frac{\beta}{\cosh^2\beta}}. \tag{2.20}$$

Notice that $\max_{\beta\geq 0}\beta/\cosh^2\beta < 1$. This ensures the existence of h. The assumption (2.20) is just for technical purposes that might possibly be omitted (see the remark right after Lemma 6 below). The description of the temperature region involves the function f in the variational formula for the thermodynamic limit of the free energy of the CW model,

$$f(\mu,\beta) := F^{SK}(\mathbf{0},\beta\mu+h) - \frac{\beta\mu^2}{2} = \ln 2 + \mathbb{E}\ln\cosh(\beta\mu+h) - \frac{\beta\mu^2}{2} \tag{2.21}$$

for $\mu\in[-1,1]$ and $\beta\in(\alpha,\infty)$, where α satisfies $\alpha\mathbb{E}1/\cosh^2 h = 1$. Some basic properties of f can be summarized in the following technical proposition.

Proposition 4. *For each fixed $\beta\in(\alpha,\infty)$, the global maximum of $f(\cdot,\beta)$ over $[0,1]$ is uniquely achieved at some $\mu(\beta)\in(0,1)$. As functions of β, $\mu(\beta)$ and $f(\mu(\beta),\beta)$ are strictly increasing, continuous, and differentiable such that*

$$\lim_{\beta\to\alpha+}\mu(\beta) = 0,\ \lim_{\beta\to\infty}\mu(\beta) = 1,\ \textit{and}\ \lim_{\beta\to\infty} f(\mu(\beta),\beta) = \infty. \tag{2.22}$$

Suppose that u is any number satisfying $0<u<1$. From Proposition 4, there exists a unique $\beta_u\in(\alpha,\infty)$ such that $\mu(\beta_u)=u$. Define $\delta_u:[\beta_u,\infty)\to[0,\infty)$ by

$$\delta_u(\beta) = f(\mu(\beta),\beta) - f(u,\beta). \tag{2.23}$$

Proposition 5. *δ_u is strictly increasing and $\lim_{\beta\to\infty}\delta_u(\beta)=\infty$.*

Recall the definition of $\mathcal{B}'_d$ from (2.11) and also $\xi(x)=\sum_{p\geq 1}\beta_p^2x^{2p}$. Suppose that u is any number satisfying $0<u<1$. We define a temperature region,

$$\mathcal{R}_u = \{(\beta,\boldsymbol{\beta})\in\mathcal{B}'_d : \beta>\beta_u \text{ and } \xi(1)<2\delta_u(\beta)\}. \tag{2.24}$$

Recall μ from (2.11). Notice that $\nu_{\boldsymbol{\beta},\beta\mu+h}$ is the limiting distribution of the overlap in the SKFI model with temperature $(\beta,\boldsymbol{\beta})$ and external field h and also in the SK model with temperature $\boldsymbol{\beta}$ and external field $\beta\mu+h$. In the case of the original SK model ($\beta_p=0$ for all $p\geq 2$) with external field $\beta\mu+h$, if $(\beta,\boldsymbol{\beta})$ satisfies $\beta>\beta_u$ and $\xi(1)<2\delta_u(\beta)$, one sees, from Proposition 5, the definition (2.11) of μ, and our main results in Theorem 5 below, that $\beta>\beta_u$ can be arbitrary large and β_1 lies very likely inside the conjectured high temperature region (below the Almeida-Thouless line) of the original SK model, that is,

$$\mathbb{E}\frac{2\beta_1^2}{\cosh^4(\beta_1 z\sqrt{2q}+\beta\mu+h)} < 1,$$

which means that $\nu_{\boldsymbol{\beta},\beta\mu+h}$ is expected to present essentially high temperature behavior, that is, $\nu_{\boldsymbol{\beta},\beta\mu+h}$ consists of a single point mass, where z is a standard Gaussian r.v. independent of h and q is the unique solution to $q=\mathbb{E}\tanh^2(\beta_1 z\sqrt{2q}+\beta\mu+h)$. Therefore,

heuristically in the region $\mathcal{R}_u$, the SKFI model has low CW and high SK temperatures. The idea of the region $\mathcal{R}_u$ comes from the observation that since $\xi(1)$ is very small comparing to β, the magnetization in the SKFI model behaves very much the same as in the CW model. Thus, if the magnetization in the CW model is away from the origin, it will also be the case in the SKFI model. Now our main result is stated as follows. Recall Ω from (2.3).

Theorem 5. *For $0 < u < 1$, we have $\Omega(\beta, \boldsymbol{\beta}, h) \subset [-1, -u) \cup (u, 1]$ for all $(\beta, \boldsymbol{\beta}) \in \mathcal{R}_u$.*

In other words, from Proposition 1, the magnetization is basically bounded away from the set $[-u, u]$. As an immediate consequence of the symmetry of the magnetization and the positivity of the overlap, we have the following proposition.

Proposition 6. *The following statements hold:*

1. *Let $0 < u < 1$ and $\mathbb{E}h^2 = 0$. For $(\beta, \boldsymbol{\beta}) \in \mathcal{R}_u$, there exists some $\mu \in (u, 1)$ such that $\langle I(|m - \mu| \leq \varepsilon)\rangle$ and $\langle I(|m + \mu| \leq \varepsilon)\rangle$ converge to $1/2$ a.s. for all $0 < \varepsilon < \mu$.*

2. *Let $0 < u < 1/2$ and $\mathbb{E}h^2 \neq 0$. For $(\beta, \boldsymbol{\beta}) \in \mathcal{R}_u$, there exists some $\mu \in (u, 1)$ such that $\langle I(|m - \mu| \leq \varepsilon)\rangle$ and $\langle I(|m + \mu| \leq \varepsilon)\rangle$ converge to Bernoulli$(1/2)$ r.v.s for all $0 < \varepsilon < \mu$.*

The first statement is well-known in the CW model without external field. The proof follows immediately from the symmetry of the magnetization under the Gibbs measure. When $\mathbb{E}h^2 \neq 0$ this symmetry does not hold, which leads to a different sense of convergence. One may also refer to [1] for the conditional self-averaging property of the magnetization that naturally leads to a similar result as the second statement of Proposition 6 in the case of the CW model with random external field. However, since the SKFI model contains SK couplings, it seems not applicable to deduce the second statement of Proposition 6 in the same approach as [1]. As will be seen in the proof, we control the magnetization using the overlap and conclude the announced result via the positivity of the overlap.

3 Proofs

In Section 3.1, we prove the main results in Section 2.1 via the usual approach in the CW model. We proceed to study the differentiability of the thermodynamic limit of the free energy of the SKFI model in Section 3.2 and conclude the results in Section 2.2. Section 3.3 is devoted to proving Theorem 4 using the EGG identities. Finally, in Section 3.4, we demonstrate how to control the magnetization quantitatively on the temperature region R_u and deduce Theorem 5 and Proposition 6. For convenience, throughout the paper, for any given $a, b \in \mathbb{R}$, we define $\delta_{a,b} = 1$ if $a = b$ and $\delta_{a,b} = 0$ if $a \neq b$; for any given set P, $|P|$ denotes the cardinality of P.

3.1 Approaches from the Curie-Weiss model

We will prove Theorem 1 and Proposition 1 by the usual approaches in the CW model. Lemma 1 is a consequence of a classical result in convex analysis, while Lemma 2 is a standard application of Gaussian concentration of measure, see [19]. These will play essential roles in our proofs.

Lemma 1. *For fixed $(\beta, \boldsymbol{\beta}) \in \mathcal{B}$, $\{F_N^{SK}(\boldsymbol{\beta}, \beta \cdot + h)\}_{N\geq 1}$ is a sequence of convex functions converging to $F^{SK}(\boldsymbol{\beta}, \beta \cdot + h)$ uniformly on $[-1, 1]$ and $F^{SK}(\boldsymbol{\beta}, \beta \cdot + h)$ is continuous and convex.*

Proof. Define $p(\mu) = F^{SK}(\boldsymbol{\beta}, \beta\mu + h)$ and $p_N(\mu) = F_N^{SK}(\boldsymbol{\beta}, \beta\mu + h)$ on $\mathbb{R}$ for each $N \geq 1$. Since

$$p_N'' = N\beta^2 \left(\mathbb{E} \left\langle m^2 \right\rangle^{SK} - \left(\mathbb{E} \left\langle m \right\rangle^{SK} \right)^2 \right) \geq 0,$$

$\{p_N\}$ is a sequence of convex functions on $\mathbb{R}$ and converges pointwise from (2.1). Note that here $\langle \cdot \rangle^{SK}$ is the Gibbs average of the SK model with temperature $\boldsymbol{\beta}$ and external field $\beta\mu + h$. A classical result in convex analysis, which can be found in [17], finishes our proof: Let $\{p_N\}_{N\geq 1}$ be a sequence of convex functions on $\mathbb{R}$ converging to p pointwise. Then p is a continuous and convex function and the convergence of $\{p_N\}_{N\geq 1}$ to p is uniform on any bounded interval. □

The proof of Lemma 2 is left to the reader.

Lemma 2. *For each N, we set $\Theta_N = \left\{-1, -1 + \frac{2}{N}, \ldots, 1 - \frac{2}{N}, 1\right\}$ and*

$$\triangle_\mu = \frac{1}{N} \ln Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{1}{N} \mathbb{E} \ln Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h)$$

for $\mu \in [-1, 1]$. Then for every $N \geq 1$,

$$P\left(\max_{\mu \in \Theta_N} |\triangle_\mu| \geq t \right) \leq K \exp\left(-\frac{t^2 N}{K} \right), \quad t \geq 0 \tag{3.1}$$

and

$$\mathbb{E} \max_{\mu \in \Theta_N} |\triangle_\mu| \leq \frac{K}{N^{1/4}}, \tag{3.2}$$

where K is a constant independent of N.

Proof of Theorem 1 : Let μ be any real number. Since $m^2 \geq 2\mu m - \mu^2$, it is easy to see

$$Z_N(\beta, \boldsymbol{\beta}, h) \geq Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) \exp\left(-\frac{N\mu^2\beta}{2} \right) \tag{3.3}$$

and this implies

$$\liminf_{N\to\infty} F_N(\beta, \boldsymbol{\beta}, h) \geq \max_{\mu \in [-1,1]} \left\{ F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\mu^2\beta}{2} \right\}.$$

On the other hand, let us observe that $m \in \Theta_N$ can take only $N+1$ distinct values. Write $1 = \sum_{\mu \in \Theta_N} \delta_{\mu,m}$. If $m = \mu$, then $m^2 = 2\mu m - \mu^2$. So by exchanging the order of summations,

$$\begin{aligned} Z_N(\beta, \boldsymbol{\beta}, h) &= \sum_{\mu \in \Theta_N} \exp\left(-\frac{N\beta\mu^2}{2}\right) \sum_{\boldsymbol{\sigma}} \delta_{\mu,m} \exp\left(H_N^{SK}(\boldsymbol{\sigma}) + \sum_{i \le N} (\beta\mu + h_i)\sigma_i\right) \\ &\le \sum_{\mu \in \Theta_N} \exp\left(-\frac{N\beta\mu^2}{2}\right) \sum_{\boldsymbol{\sigma}} \exp\left(H_N^{SK}(\boldsymbol{\sigma}) + \sum_{i \le N} (\beta\mu + h_i)\sigma_i\right) \\ &= \sum_{\mu \in \Theta_N} \exp\left(-\frac{N\beta\mu^2}{2} + Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h)\right). \end{aligned}$$

Therefore,

$$\begin{aligned} F_N(\beta, \boldsymbol{\beta}, h) &\le \frac{\ln(N+1)}{N} + \mathbb{E}\left[\max_{\mu \in \Theta_N} \left\{-\frac{\beta\mu^2}{2} + \frac{1}{N} \ln Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h)\right\}\right] \\ &\le \frac{\ln(N+1)}{N} + \max_{\mu \in \Theta_N} \left\{-\frac{\beta\mu^2}{2} + F_N^{SK}(\boldsymbol{\beta}, \beta\mu + h)\right\} + \mathbb{E} \max_{\mu \in \Theta_N} |\triangle_\mu|. \end{aligned}$$

From (3.2), we obtain

$$\limsup_{N \to \infty} F_N(\beta, \boldsymbol{\beta}, h) \le \max_{\mu \in [-1,1]} \left\{F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2}\right\} + \limsup_{N \to \infty} \mathbb{E} \max_{\mu \in \Theta_N} |\triangle_\mu|.$$

and by using Lemma 2, we are done.

□

Proof of Proposition 1 : It is easy to see that if (2.4) holds, then using the exponential bound of (2.4), (2.5) follows immediately. So we only prove (2.4). As in Theorem 1, by exchanging the order of summations, we obtain

$$\begin{aligned} &\langle \{m \in U\} \rangle \, Z_N(\beta, \boldsymbol{\beta}, h) \\ &= \sum_{\mu \in \Theta_N} 1_{\{\mu \in U\}} \exp\left(-\frac{N\beta\mu^2}{2}\right) \sum_{\boldsymbol{\sigma} \in \Sigma_N} \delta_{\mu,m} \exp\left(-H_N^{SK}(\boldsymbol{\sigma}) + \sum_{i \le N} (\beta\mu + h_i)\sigma_i\right) \\ &\le \sum_{\mu \in \Theta_N} 1_{\{\mu \in U\}} \exp\left(-\frac{N\beta\mu^2}{2}\right) Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) \\ &\le (N+1) \exp\left(N \sup_{\mu \in \Theta_N \cap U} \left\{\frac{1}{N} \ln Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2}\right\}\right). \end{aligned}$$

From (3.3),

$$\langle \{m \in U\} \rangle \le (N+1) \exp N \left(\max_{\mu \in \Theta_N \cap U} W_\mu - \max_{\mu \in \Theta_N} W_\mu\right), \tag{3.4}$$

where

$$W_\mu := \frac{1}{N} \ln Z_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2}.$$

We claim that $\lim_{N\to\infty} \sup_{\mu\in\Theta_N\cap O} W_\mu = \sup_{\mu\in O} \left\{F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \beta\mu^2/2\right\}$ for every open subset O of $[-1, 1]$. For convenience, we set

$$\Gamma_\mu = F_N^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2},$$
$$F_\mu = F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2}.$$

For any μ, observe that $-|\triangle_\mu| + \Gamma_\mu \le W_\mu = \triangle_\mu + \Gamma_\mu \le |\triangle_\mu| + \Gamma_\mu$ and thus,

$$-\max_{\mu\in\Theta_N\cap O} |\triangle_\mu| + \max_{\mu\in\Theta_N\cap O} \Gamma_\mu \le \max_{\mu\in\Theta_N\cap O} W_\mu \le \max_{\mu\in\Theta_N\cap O} |\triangle_\mu| + \max_{\mu\in\Theta_N\cap O} \Gamma_\mu. \tag{3.5}$$

Using Lemma 1 and 2, this completes the proof of our claim since

$$\sup_{\mu\in O} F_\mu = \liminf_{N\to\infty} \max_{\mu\in\Theta_N\cap O} \Gamma_\mu \le \liminf_{N\to\infty} \max_{\mu\in\Theta_N\cap U} |W_\mu|$$

and

$$\sup_{\mu\in O} F_\mu = \limsup_{N\to\infty} \max_{\mu\in\Theta_N\cap O} \Gamma_\mu \ge \limsup_{N\to\infty} \max_{\mu\in\Theta_N\cap U} |W_\mu|.$$

To obtain (2.4), we write from (3.4),

$$\mathbb{E}\left\langle \{m \in U\} \right\rangle \le (N+1)\mathbb{P}(A_N) + (N+1)\exp\left(-N\varepsilon\right)\mathbb{P}(A_N^c),$$

where

$$\varepsilon = 2\left(\max_{[-1,1]} F_\mu - \sup_U F_\mu\right) > 0,$$
$$A_N = \left\{\max_{\Theta_N\cap U} W_\mu - \max_{\Theta_N} W_\mu \le -\varepsilon\right\}.$$

Since

$$\max_{\Theta_N} \Gamma_\mu - \max_{\Theta_N\cap U} \Gamma_\mu \to \max_{[-1,1]} F_\mu - \sup_U F_\mu = \frac{\varepsilon}{2},$$

it follows that for sufficiently large N,

$$\max_{\Theta_N} \Gamma_\mu - \max_{\Theta_N\cap U} \Gamma_\mu \le \frac{3\varepsilon}{4}.$$

Now use (3.5) to obtain

$$\begin{aligned}
\mathbb{P}(A_N) &\le \mathbb{P}\left(\max_{\Theta_N\cap U} \Gamma_\mu - \max_{\Theta_N} \Gamma_\mu \le -\varepsilon + \max_{\Theta_N\cap U} |\triangle_\mu| + \max_{\Theta_N} |\triangle_\mu|\right) \\
&\le \mathbb{P}\left(\max_{\Theta_N} |\triangle_\mu| \ge \frac{1}{2}\left(\varepsilon + \max_{\Theta_N\cap U} \Gamma_\mu - \max_{\Theta_N} \Gamma_N\right)\right) \\
&\le \mathbb{P}\left(\max_{\Theta_N} |\triangle_\mu| \ge \frac{\varepsilon}{8}\right).
\end{aligned}$$

So for large enough N and from equation (3.1), we get

$$\mathbb{E}\left\langle\{m \in U\}\right\rangle \leq (N+1)\left(K\exp\left(-\frac{\varepsilon^2 N}{64K}\right)+\exp(-N\varepsilon)\right)$$

and this establishes (2.4).

□

3.2 The differentiability of $F(\beta, \boldsymbol{\beta}, h)$ in $(\beta, \boldsymbol{\beta})$

We will study the differentiability of $F(\beta, \boldsymbol{\beta}, h)$ with respect to β and β_p for every $p \geq 1$ in this section using the standard results in convex analysis. From this, we deduce the main results in Section 2.2. First let us recall that the thermodynamic limit $F^{SK}(\boldsymbol{\beta}, h)$ of the free energy in the mixed even p-spin SK model can be characterized by the Parisi formula for any $\boldsymbol{\beta}$ with $\sum_{p\geq 1} 2^p\beta_p^2 < \infty$ and Gaussian r.v. h (possibly degenerate). Using this variational formula and the usual trick concerning the differentiability of the convex functions, it is well-known [12, 21] that $F^{SK}(\boldsymbol{\beta}, h)$ is differentiable with respect to β_p for every $p \geq 1$. For each $x \in \mathbb{R}$, we consider the mixed even p-spin SK model with temperature $\boldsymbol{\beta}$ and external field $x + h$. One may see that following a similar argument as [12, 21], the function $F^{SK}(\boldsymbol{\beta}, x + h)$ is differentiable with respect to x. More precisely, the following statement holds.

Proposition 7. *Let $\{W_t\}_{t\geq 0}$ be a standard Brownian motion. For every fixed $\boldsymbol{\beta}$ and h, $F^{SK}(\boldsymbol{\beta}, x + h)$ is differentiable in x and*

$$\frac{\partial F^{SK}}{\partial x}(\boldsymbol{\beta}, x+h) = \mathbb{E}\left[\tanh(x+h+W_{\xi'(1)})\exp S(x)\right], \tag{3.6}$$

where $S(x)$ is some r.v. depending only on the Parisi measure $\nu_{\boldsymbol{\beta},x+h}$ and ξ such that $\mathbb{E}\exp S(x) = 1$ for every $x \in \mathbb{R}$.

Now let us turn to the study of the differentiability of $F(\beta, \boldsymbol{\beta}, h)$ in $(\beta, \boldsymbol{\beta})$. Recall from (2.2) that the thermodynamic limit of the free energy of the SKFI model, $F(\beta, \boldsymbol{\beta}, h)$, is obtained by maximizing

$$f(\mu, \beta, \boldsymbol{\beta}) := F^{SK}(\boldsymbol{\beta}, \beta\mu + h) - \frac{\beta\mu^2}{2} \tag{3.7}$$

over all $\mu \in [-1, 1]$. Let us observe that for fixed μ, f is convex in β and β_p for each $p \geq 1$. Such an optimization problem is of great importance in the analysis of convex optimization. The differentiability of $F(\beta, \boldsymbol{\beta}, h)$ in β and β_p for each $p \geq 1$ relies on the following classical theorem in convex analysis.

Theorem 6 (Danskin [2])**.** *Let I_1 be an open interval and I_2 be a compact interval. Suppose that g is a continuous function defined from $I_1 \times I_2$ to $\mathbb{R}$ such that for every fixed y, $g(\cdot, y)$ is convex and $\frac{\partial g}{\partial x}(x, y)$ exists for every $(x, y) \in I_1 \times I_2$. Define $G : I_1 \to \mathbb{R}$ by $G(x) = \max_{y\in I_2} g(x, y)$. Then*

$$\frac{dG}{dx+}(x) = \max_{y\in\Omega_g(x)} \frac{\partial g}{\partial x}(x, y)$$

and

$$\frac{dG}{dx-}(x) = \min_{y\in\Omega_g(x)} \frac{\partial g}{\partial x}(x,y),$$

where $\frac{dG}{dx+}$ *and* $\frac{dG}{dx-}$ *are the right and left partial derivatives of* G *with respect to* x*, respectively, and* $\Omega_g(x)$ *is the argmax of* $g(x,\cdot)$ *on* I_2 *for each* $x \in I_1$*. In particular, if* $\Omega_g(x)$ *consists of a single element, then* G *is differentiable at* x*.*

Before turning to the proof of our main results, we need two technical lemmas.

Lemma 3 (Griffith)**.** *Suppose that* $\{g_n\}$ *is a sequence of differentiable convex functions defined on an open interval* I*. If* $\{g_n\}$ *converges pointwise to* g *and* g *is differentiable at* x*, then* $\lim_{n\to\infty} g_n'(x) = g'(x)$*.*

Lemma 4. *Let* $(\beta,\boldsymbol{\beta}) \in \mathcal{B}$*. If* $\beta > 0$*, then* $\Omega(\beta,\boldsymbol{\beta},h) \subset (-1,1)$*.*

Proof. Notice that for fixed $(\beta,\boldsymbol{\beta}) \in \mathcal{B}$, $f(\cdot,\beta,\boldsymbol{\beta})$ is a well-defined function on $\mathbb{R}$. Since $|m(\boldsymbol{\sigma})| \le 1$ for every $\boldsymbol{\sigma} \in \Sigma_N$, this implies

$$\begin{aligned} F^{SK}(\boldsymbol{\beta},\beta\mu+h) &= \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\ln\sum_{\boldsymbol{\sigma}\in\Sigma_N}\exp\left(-H_N^{SK}(\boldsymbol{\sigma})+\sum_{i\le N}h_i\sigma_i+\beta\mu\sum_{i\le N}\sigma_i\right)\right] \\ &\le F^{SK}(\boldsymbol{\beta},h)+\beta|\mu| \end{aligned}$$

and $f(\mu,\beta,\boldsymbol{\beta}) \to -\infty$ as $|\mu| \to \infty$. So the global maximum of $f(\cdot,\beta,\boldsymbol{\beta})$ is achieved. Suppose that μ is any maximizer. Then $\frac{\partial f}{\partial\mu}(\mu,\beta,\boldsymbol{\beta}) = 0$ and (3.6) together yield

$$\mathbb{E}\left[\tanh(\beta\mu+h+W_{\xi'(1)})\exp S(\beta\mu)\right] = \mu,$$

where S is defined in Proposition 7. Since $|\tanh| < 1$ and $\mathbb{E}\left[\exp S(\beta\mu)\right] = 1$, it means $\mu \in (-1,1)$. So

$$\Omega(\beta,\boldsymbol{\beta},h) = \mathrm{Argmax}_{\mu\in[-1,1]} f(\mu,\beta,\boldsymbol{\beta}) = \mathrm{Argmax}_{\mu\in\mathbb{R}} f(\mu,\beta,\boldsymbol{\beta}) \subset (-1,1).$$

□

Proof of Proposition 2 : For fixed μ and $\boldsymbol{\beta}$, since $F^{SK}(\boldsymbol{\beta},\beta\mu+h) - \beta\mu^2/2$ is convex and differentiable in β, it follows by Danskin's theorem that

$$\frac{\partial F}{\partial\beta+}(\beta,\boldsymbol{\beta},h) = \max_{\mu\in\Omega(\beta,\boldsymbol{\beta},h)}\left(\mu\left.\frac{\partial F^{SK}}{\partial y}(\boldsymbol{\beta},y+h)\right|_{y=\beta\mu} - \frac{\mu^2}{2}\right) \tag{3.8}$$

and

$$\frac{\partial F}{\partial\beta-}(\beta,\boldsymbol{\beta},h) = \min_{\mu\in\Omega(\beta,\boldsymbol{\beta},h)}\left(\mu\left.\frac{\partial F^{SK}}{\partial y}(\boldsymbol{\beta},y+h)\right|_{y=\beta\mu} - \frac{\mu^2}{2}\right). \tag{3.9}$$

Suppose that $F(\beta,\boldsymbol{\beta},h)$ is differentiable at β. If $|\Omega(\beta,\boldsymbol{\beta},h)| \ge 3$, then from Proposition 7 and Lemma 4, there exist some $\mu_1,\mu_2 \in \Omega(\beta,\boldsymbol{\beta},h) \subset (-1,1)$ with $|\mu_1| < |\mu_2|$ such that

$$\mu_1 = \left.\frac{\partial F^{SK}}{\partial y}(\boldsymbol{\beta},y+h)\right|_{y=\beta\mu_1} \quad\text{and}\quad \mu_2 = \left.\frac{\partial F^{SK}}{\partial y}(\boldsymbol{\beta},y+h)\right|_{y=\beta\mu_2}.$$

From these two equations, (3.8), and (3.9), we obtain

$$\frac{\partial F}{\partial \beta+}(\beta,\boldsymbol{\beta},h) \geq \frac{1}{2}\mu_2^2 > \frac{1}{2}\mu_1^2 \geq \frac{\partial F}{\partial \beta-}(\beta,\boldsymbol{\beta},h), \tag{3.10}$$

which contradicts to our assumption that F is differentiable. Hence, $|\Omega(\beta,\boldsymbol{\beta},h)| \leq 2$ and if $\mu_1,\mu_2 \in \Omega(\beta,\boldsymbol{\beta},h)$ are distinct, then $\mu_1 = -\mu_2$. So $(\beta,\boldsymbol{\beta}) \in \mathcal{B}_d$. Conversely, suppose that $(\beta,\boldsymbol{\beta}) \in \mathcal{B}_d$. If $|\Omega(\beta,\boldsymbol{\beta},h)| = 1$, then we are done. If $|\Omega(\beta,\boldsymbol{\beta},h)| = 2$ and $\mu_1,\mu_2 \in \Omega(\beta,\boldsymbol{\beta},h)$ with $\mu_1 = -\mu_2$, then from Lemma 4, (3.8), (3.9), and (3.10), we have $\frac{\partial F}{\partial \beta+}(\beta,\boldsymbol{\beta},h) = \frac{\partial F}{\partial \beta-}(\beta,\boldsymbol{\beta},h)$. So $F(\beta,\boldsymbol{\beta},h)$ is differentiable at β and this completes the proof.

□

Proof of Theorem 2 : For any $(\beta,\boldsymbol{\beta}) \in \mathcal{B}'_d$, since h is centered, $f(\mu,\beta,\boldsymbol{\beta})$ is symmetric in μ and we may represent $F(\beta,\boldsymbol{\beta},h)$ as

$$F(\beta,\boldsymbol{\beta},h) = \max_{\mu\in[0,1]} f(\mu,\beta,\boldsymbol{\beta}).$$

Let $(\beta,\boldsymbol{\beta}) \in \mathcal{B}'_d$. Then either $\Omega(\beta,\boldsymbol{\beta},h) = \{0\}$ or $\Omega(\beta,\boldsymbol{\beta},h) = \{\mu,-\mu\}$ for some $\mu \neq 0$. This means that $\Omega(\beta,\boldsymbol{\beta},h)\cap[0,1]$ consists of a single element, say μ. By Danskin's theorem we obtain

$$\frac{\partial F}{\partial \beta_p+}(\beta,\boldsymbol{\beta},h) = \frac{\partial F^{SK}}{\partial \beta_p}(\boldsymbol{\beta},\beta\mu+h) = \frac{\partial F}{\partial \beta_p-}(\beta,\boldsymbol{\beta},h).$$

This proves that $F(\beta,\boldsymbol{\beta},h)$ is differentiable with respect to every β_p and from (2.9) the equation (2.12) follows. Using Gaussian integration by parts, we have

$$\frac{\partial}{\partial \beta_p}\frac{1}{N}\mathbb{E}\ln Z_N(\beta,\boldsymbol{\beta},h) = \beta_p(1-\mathbb{E}\left\langle R_{1,2}^{2p}\right\rangle).$$

By Griffith's lemma, this implies that

$$\frac{\partial F}{\partial \beta_p}(\beta,\boldsymbol{\beta},h) = \beta_p\left(1-\lim_{N\to\infty}\mathbb{E}\left\langle R_{1,2}^{2p}\right\rangle\right)$$

and from (2.12), we get (2.13).

□

Proof of Theorem 3 : Note that by Talagrand's positivity, $c > 0$. From (2.13) and a continuity argument, for every continuous function f on $[0,1]$,

$$\lim_{N\to\infty}\mathbb{E}\left\langle f(|R_{1,2}|)\right\rangle = \int_0^1 f(q)\nu_{\boldsymbol{\beta},\beta\mu+h}(dq). \tag{3.11}$$

In particular, let $f_0 : [0,1] \to \mathbb{R}$ be the continuous function satisfying $f_0(x) = 1$ if $0 \leq x \leq c'$, $f_0(x) = (c-c')^{-1}(c-x)$ if $c' < x < c$, and $f_0(x) = 0$ if $c \leq x \leq 1$. Then from (3.11),

$$\lim_{N\to\infty}\mathbb{E}\left\langle I(|R_{1,2}| \leq c')\right\rangle \leq \lim_{N\to\infty}\mathbb{E}\left\langle f_0(|R_{1,2}|)\right\rangle \leq \nu_{\boldsymbol{\beta},\beta\mu+h}([0,c)) = 0. \tag{3.12}$$

Define $F(t) = F(\beta, \boldsymbol{\beta}, th)$ and $F^{SK}(t) = F(\boldsymbol{\beta}, th)$ for $t \in \mathbb{R}$. Recall that since h is centered Gaussian, $F^{SK}(t)$ is differentiable in t by [12]. Thus, the same argument as Theorem 2 implies that $F(t)$ is differentiable at $t = 1$ and so

$$\lim_{N\to\infty} \mathbb{E}\langle R_{1,2}\rangle = \int_0^1 q\nu_{\boldsymbol{\beta},\beta\mu+h}(dq).$$

On the other hand, letting $f(x) = x$ and using (3.11),

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\left\langle R_{1,2}^-\right\rangle &= \frac{1}{2}\left(\lim_{N\to\infty} \mathbb{E}\langle |R_{1,2}|\rangle - \lim_{N\to\infty} \mathbb{E}\langle R_{1,2}\rangle\right) \\ &= \frac{1}{2}\left(\int_0^1 q d\nu_{\boldsymbol{\beta},\beta\mu+h}(q) - \int_0^1 q d\nu_{\boldsymbol{\beta},\beta\mu+h}(q)\right) \\ &= 0.\end{aligned} \tag{3.13}$$

Thus, from (3.12) and (3.13) and applying the Markov inequality, we obtain (2.15) since

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\langle I(R_{1,2} \le c')\rangle &\le \limsup_{N\to\infty} \mathbb{E}\langle I(|R_{1,2}| \le c')\rangle + \limsup_{N\to\infty} \mathbb{E}\left\langle I(R_{1,2}^- > c')\right\rangle \\ &\le \limsup_{N\to\infty} \frac{1}{c'}\mathbb{E}\left\langle R_{1,2}^-\right\rangle \\ &= 0.\end{aligned}$$

From this and (3.11), we conclude (2.16) since for any continuous function f on $[0,1]$,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\langle f(|R_{1,2}|)\rangle &= \lim_{N\to\infty} \mathbb{E}\langle f(-R_{1,2})I(R_{1,2} < 0)\rangle + \mathbb{E}\langle f(R_{1,2})I(R_{1,2} \ge 0)\rangle \\ &= \lim_{N\to\infty} \mathbb{E}\langle f(R_{1,2})I(R_{1,2} \ge 0)\rangle \\ &= \lim_{N\to\infty} \mathbb{E}\langle f(R_{1,2})I(R_{1,2} < 0)\rangle + \mathbb{E}\langle f(R_{1,2})I(R_{1,2} \ge 0)\rangle \\ &= \lim_{N\to\infty} \mathbb{E}\langle f(R_{1,2})\rangle\end{aligned}$$

for every continuous function f on $[-1,1]$.

□

3.3 An application of the extended Ghirlanda-Guerra identities

This section is devoted to proving Theorem 4 using the EGG identities. Let $(\beta, \boldsymbol{\beta}) \in \mathcal{B}_d'$. Recall that from Proposition 3, the EGG identities (2.17) hold under the assumption $\mathbb{E}h^2 \neq 0$. In the case of $\mathbb{E}h^2 = 0$, we have the following weaker identities that can be derived in the same way as Proposition 3: for each n and each continuous function ψ on $\mathbb{R}$,

$$\lim_{N\to\infty} \sup_f \left| n\mathbb{E}\langle \psi(|R_{1,n+1}|f)\rangle - \mathbb{E}\langle \psi(|R_{1,2}|)\rangle \mathbb{E}\langle f\rangle - \sum_{2\le\ell\le n} \mathbb{E}\langle \psi(|R_{1,\ell}|)f\rangle \right| = 0, \tag{3.14}$$

where the supremum is taken over all (non random) functions f on Σ_N^n with $|f| \leq 1$. Let us remark that (2.17) obviously implies (3.14).

Recall from Theorems 2 and 3 that the Parisi measure $\nu_{\boldsymbol{\beta},\beta\mu+h}$ is a probability measure defined $[0,1]$ that describes the limiting distribution of $|R_{1,2}|$ for both cases $\mathbb{E}h^2 = 0$ and $\mathbb{E}h^2 \neq 0$. Let ν_N be the distribution of the array of all overlaps $(|R_{\ell,\ell'}|)_{\ell,\ell'\geq 1}$ under the Gibbs average $\mathbb{E}\langle\cdot\rangle$. By compactness, the sequence (ν_N) converges weakly over subsequences but, for simplicity of notation, we will assume that ν_N converges weakly to the limit ν. We will still use the notations $(|R_{\ell,\ell'}|)_{\ell,\ell'\geq 1}$ to denote the elements of the overlap array in the limit and, again, for simplicity of notations we will denote by $\mathbb{E}$ the expectation with respect to the measure ν. Using these notations, (3.14) implies

$$\mathbb{E}\psi(|R_{1,n+1}|)f = \frac{1}{n}\mathbb{E}\psi(|R_{1,2}|)\mathbb{E}f + \frac{1}{n}\sum_{\ell=2}^{n}\mathbb{E}\psi(|R_{1,\ell}|)f \tag{3.15}$$

for all bounded measurable functions f of the overlaps on n replicas and bounded measurable function ψ on $\mathbb{R}$. We will need the following essential lemma.

Lemma 5. *Let $(\beta,\boldsymbol{\beta}) \in \mathcal{B}_d'$. Suppose that A is any measurable subset of $[0,1]$. Set $A_n = \{|R_{\ell,\ell'}| \in A,\ \forall \ell \neq \ell' \leq n\}$. Then $\nu(A_n) \geq \nu_{\boldsymbol{\beta},\beta\mu+h}(A)^n$.*

Proof. For any $n \geq 1$, observe that

$$I_{A_{n+1}} \geq I_{A_n} - \sum_{\ell\leq n} I(|R_{\ell,n+1}| \notin A)I_{A_n}. \tag{3.16}$$

For all $1 \leq \ell \leq n$, applying (3.15) and using symmetry of the overlaps,

$$\begin{aligned}\mathbb{E}I(|R_{\ell,n+1}| \notin A)I_{A_n} &= \frac{1}{n}\nu_{\boldsymbol{\beta},\beta\mu+h}(A^c)\nu(A_n) + \frac{1}{n}\sum_{\ell'\neq\ell}^{n}\mathbb{E}I(|R_{\ell,\ell'}| \notin A)I_{A_n} \\ &= \frac{1}{n}\nu_{\boldsymbol{\beta},\beta\mu+h}(A^c)\nu(A_n)\end{aligned}$$

and, therefore, from (3.16), $\nu(A_{n+1}) \geq \nu_{\boldsymbol{\beta},\beta\mu+h}(A)\nu(A_n)$. Thus, an induction argument yields the result.

□

Proof of Theorem 4 : If $\mathbb{E}h^2 \neq 0$, then from the positivity of the overlap and the first statement, the second statement follows immediately. So we only need to prove the first statement. If $\varepsilon \geq \mu^2$, we are obviously done. Suppose that $\varepsilon < \mu^2$ and the announced result fails. Then $\liminf_{N\to\infty}\mathbb{E}\langle I(|R_{1,2}| \geq \mu^2 - \varepsilon)\rangle < 1$ for some $\varepsilon > 0$ or equivalently, $0 < \limsup_{N\to\infty}\mathbb{E}\langle I(|R_{1,2}| < \mu^2 - \varepsilon)\rangle$. Without loss of generality, we may assume that $\nu_{\boldsymbol{\beta},\beta\mu+h}$ is continuous at $\mu^2 - \varepsilon$. Then $\nu_{\boldsymbol{\beta},\beta\mu+h}([0,\mu^2-\varepsilon)) > 0$ and from Lemma 5, $\nu(A_n) > 0$ for every n, where A_n is defined in the statement of Lemma 5 using $A = [0,\mu^2-\varepsilon)$. Let $\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n$ be n replicas and $a_1,\ldots,a_n \in \{-1,1\}$ such that $a_\ell m(\boldsymbol{\sigma}^\ell) = |m(\boldsymbol{\sigma}^\ell)|$ for $1 \leq \ell \leq n$. From the Cauchy-Schwarz inequality,

$$N\sum_{\ell\leq n}|m(\boldsymbol{\sigma}^\ell)| = N\sum_{\ell\leq n}m(a_\ell\boldsymbol{\sigma}^\ell) = \mathbf{1}\cdot\sum_{\ell\leq n}a_\ell\boldsymbol{\sigma}^\ell \leq \sqrt{N}\left\|\sum_{\ell\leq n}a_\ell\boldsymbol{\sigma}^\ell\right\|, \tag{3.17}$$

where $\|\cdot\|$ is the Eculidean distance in $\mathbb{R}^N$. Notice that

$$\left\|\sum_{\ell\le n} a_\ell \boldsymbol{\sigma}^\ell\right\|^2 = nN + \sum_{\ell\neq\ell'\le n} a_\ell a_{\ell'} \boldsymbol{\sigma}^\ell\cdot\boldsymbol{\sigma}^{\ell'} \le Nn + N\sum_{\ell\neq\ell'\le n}|R_{\ell,\ell'}|. \tag{3.18}$$

Combining (3.17) and (3.18),

$$\sum_{\ell\le n}|m(\boldsymbol{\sigma}^\ell)| \le \left(n + \sum_{\ell\neq\ell'\le n}|R_{\ell,\ell'}|\right)^{1/2} \tag{3.19}$$

From this inequality, applying $\nu(A_n)>0$ together with the openness of A_n, we obtain

$$\liminf_{N\to\infty}\mathbb{E}\left\langle I\left(|m(\boldsymbol{\sigma}^1)|+\cdots+|m(\boldsymbol{\sigma}^n)|<(n+(\mu^2-\varepsilon)n(n-1))^{1/2}\right)\right\rangle>0. \tag{3.20}$$

On the other hand, let us pick $0<\varepsilon'<\varepsilon$ and notice that for each $1\le\ell\le n$,

$$\lim_{N\to\infty}\mathbb{E}\left\langle I\left(\left||m(\boldsymbol{\sigma}^\ell)|-\mu\right|<\mu-(\mu^2-\varepsilon')^{1/2}\right)\right\rangle=1.$$

We conclude, from this, (3.20), and triangular inequality, that for each $n\ge1$, with nonzero probability,

$$n(\mu^2-\varepsilon')^{1/2} = n(\mu-\mu+(\mu^2-\varepsilon')^{1/2}) \le \sum_{\ell\le n}|m(\boldsymbol{\sigma}^\ell)| < (n+(\mu^2-\varepsilon)n(n-1))^{1/2},$$

and this means $\varepsilon<\varepsilon'$, a contradiction. □

3.4 Controlling the magnetization using the CW free energy

In this section, we will demonstrate how to control the magnetization quantitatively using the thermodynamic limit of the free energy of the CW model. From this, we conclude the main results in Section 2.4. Recall that the external field h in Section 2.4 is a centered Gaussian r.v. satisfying (2.20). First, let us establish a technical lemma that will be used in Proposition 4.

Lemma 6. *Suppose that h is centered Gaussian satisfying* (2.20). *Then $\beta\mathbb{E}1/\cosh^2(\beta+h)<1$ for every $\beta\ge0$.*

Let us remark that the technical condition (2.20) is only used here throughout the paper, while the inequality will play a crucial role that ensures the validity of our main results. According to the simulation data, the inequality in Lemma 6 should be also valid even without the assumption (2.20). However, the proof for this general case seems much more involved and too distracted. For clarity, we will only focus on the h satisfying (2.20).

Proof of Lemma 6 : Let $\beta \geq 0$. We claim that $\cosh^2\beta/\cosh^2(\beta+x) < \exp(2|x|)$ for all $x \neq 0$. To see this, define $g(x) = 2(\ln\cosh\beta - \ln\cosh(\beta+x))$. Then $g(0) = 0$ and $g'(x) = -2\tanh(\beta+x)$. For each x, using mean value theorem, we obtain

$$g(x) = g(0) + g'(x')x = -2x\tanh(\beta+x') \leq 2|x|$$

for some $x' \in (0,x)$ if $x > 0$ or $x' \in (x,0)$ if $x < 0$. This completes the proof of our claim and consequently, Lemma 6 follows from the assumption on h,

$$\beta\mathbb{E}\frac{1}{\cosh^2(\beta+h)} = \frac{\beta}{\cosh^2\beta}\mathbb{E}\frac{\cosh^2\beta}{\cosh^2(\beta+h)} \leq \frac{\beta}{\cosh^2\beta}\mathbb{E}\exp(2|h|) < 1.$$

□

Proof of Proposition 4 : Recall from (2.21) that $f(\mu,\beta)$ is defined for $\mu \in [-1,1]$ and $\beta \in (\alpha,\infty)$ for some α satisfying $\alpha\mathbb{E}1/\cosh^2 h = 1$. A simple computation yields the first three partial derivatives of $f(\mu,\beta)$ with respect to μ :

$$\begin{aligned}
\frac{\partial f}{\partial\mu}(\mu,\beta) &= \beta\left(\mathbb{E}\tanh(\beta\mu+h) - \mu\right), \\
\frac{\partial^2 f}{\partial\mu^2}(\mu,\beta) &= \beta\left(\beta\mathbb{E}\frac{1}{\cosh^2(\beta\mu+h)} - 1\right), \\
\frac{\partial^3 f}{\partial\mu^3}(\mu,\beta) &= -2\beta^3\mathbb{E}\frac{\tanh(\beta\mu+h)}{\cosh^2(\beta\mu+h)}.
\end{aligned} \tag{3.21}$$

Let us recall a useful lemma from the proof of Proposition A.14.1 in [24]: Let ϕ be an increasing bounded function on $\mathbb{R}$ satisfying $\phi(-y) = -\phi(y)$ and $\phi''(y) < 0$ for $y > 0$. Then for every $\mu \geq 0$ and center Gaussian random variable z,

$$\mathbb{E}\phi(z+\mu)\phi'(z+\mu) \geq 0.$$

Applying this lemma to $\phi(y) = \tanh(y)$, we have $\frac{\partial^3 f}{\partial\mu^3} < 0$ for every $\mu > 0$ from (3.21). It implies that $\frac{\partial^2 f}{\partial\mu^2}(\cdot,\beta)$ is strictly decreasing on $[0,1]$. By the definition of α and Lemma 6, we also know that $\frac{\partial^2 f}{\partial\mu^2}(0,\beta) > 0$ and $\frac{\partial^2 f}{\partial\mu^2}(1,\beta) < 0$. So $\frac{\partial^2 f}{\partial\mu^2}(\cdot,\beta)$ has a unique zero in $(0,1)$ and so does $\frac{\partial f}{\partial\mu}(\cdot,\beta)$ since $\frac{\partial f}{\partial\mu}(0,\beta) = 0$ and $\frac{\partial f}{\partial\mu}(1,\beta) < 0$. Let $\mu(\beta) \in (0,1)$ be the zero of $\frac{\partial f}{\partial\mu}(\cdot,\beta)$. Hence, $\frac{\partial f}{\partial\mu}(\cdot,\beta) > 0$ on $(0,\mu(\beta))$ and $\frac{\partial f}{\partial\mu}(\cdot,\beta) < 0$ on $(\mu(\beta),1)$, which implies that in $[0,1]$, $f(\cdot,\beta)$ attains its unique global maximum at $\mu(\beta)$.

The continuity and differentiability of $\mu(\cdot)$ follow from the implicit function theorem. It is then clear that $f(\mu(\cdot),\cdot)$ is continuous and differentiable. Since

$$\mathbb{E}\tanh(\beta\mu(\beta)+h) = \mu(\beta), \tag{3.22}$$

by taking derivative on both sides, we obtain

$$(\mu(\beta)+\beta\mu'(\beta))\,\mathbb{E}\frac{1}{\cosh^2(\beta\mu(\beta)+h)} = \mu'(\beta)$$

and so

$$\mu'(\beta) = -\frac{\beta\mu(\beta)}{\frac{\partial^2 f}{\partial \mu^2}(\mu(\beta),\beta)} E \frac{1}{\cosh^2(\beta\mu(\beta)+h)}.$$

Since $\mu(\beta)$ is greater than the unique zero of $\frac{\partial^2 f}{\partial \mu^2}(\cdot,\beta)$ in $(0,1)$, $\frac{\partial^2 f}{\partial \mu^2}(\mu(\beta),\beta) < 0$ and this means that $\mu(\cdot)$ is a strictly increasing function. We also show the monotonicity of $f(\mu(\cdot),\cdot)$ by using (3.22),

$$\begin{aligned}
\frac{df}{d\beta}(\mu(\beta),\beta) &= (\mu(\beta)+\beta\mu'(\beta))\mathbb{E}\tanh(\beta\mu(\beta)+h) - \frac{\mu(\beta)^2}{2} - \beta\mu'(\beta)\mu(\beta) \\
&= (\mu(\beta)+\beta\mu'(\beta))\mu(\beta) - \frac{\mu(\beta)^2}{2} - \beta\mu'(\beta)\mu(\beta) \\
&= \frac{1}{2}\mu(\beta)^2.
\end{aligned}$$

Finally, we check (2.22). First notice that the solution of $\mathbb{E}\tanh(\alpha x + h) = x$ for $x \in [0,1]$ is unique and equals 0. This can be verified by the same argument as in the first part of our proof. Thus, from (3.22),

$$\mathbb{E}\tanh\left(\alpha \lim_{\beta\to\alpha+} \mu(\beta) + h\right) = \lim_{\beta\to\alpha+} \mu(\beta)$$

implies $\lim_{\beta\to\alpha+}\mu(\beta) = 0$. Since $\beta\mu(\beta) \to \infty$ as $\beta \to \infty$, we obtain, by the dominated convergence theorem,

$$\lim_{\beta\to\infty} \mu(\beta) = \lim_{\beta\to\infty} \mathbb{E}\tanh(\beta\mu(\beta)+h) = 1.$$

Since by the monotonicity of $\mu(\cdot)$ and the mean value theorem

$$f(\mu(\beta),\beta) - f(\mu(\beta'),\beta') \geq \frac{\mu(\beta')^2}{2}(\beta-\beta')$$

for $\beta > \beta' > \alpha$, this implies that $\lim_{\beta\to\infty} f(\mu(\beta),\beta) = \infty$ and completes our proof. □

Proof of Proposition 5 : Notice that $\mathbb{E}\tanh(\beta\mu+h)$ is a strictly increasing function in μ since $\frac{d}{d\mu}\mathbb{E}\tanh(\beta\mu+h) = \beta\mathbb{E}1/\cosh^2(\beta\mu+h) > 0$ and that $\mu(\beta) = \mathbb{E}\tanh(\beta\mu(\beta)+h)$ since $\mu(\beta) \in (0,1)$ is the maximizer of $f(\cdot,\beta)$ on $[0,1]$. Thus, for $\beta > \beta_u$,

$$\begin{aligned}
\frac{d}{d\beta}\left(f(\mu(\beta),\beta) - f(u,\beta)\right) &= \frac{1}{2}(\mu(\beta)^2+u^2) - u\mathbb{E}\tanh(\beta u+h) \\
&> \frac{1}{2}\left(\mu(\beta)^2+u^2\right) - u\mathbb{E}\tanh(\beta\mu(\beta)+h) \\
&= \frac{1}{2}\left(\mu(\beta)^2+u^2\right) - u\mu(\beta) \\
&= \frac{1}{2}(\mu(\beta)-u)^2 \\
&> 0
\end{aligned}$$

and this implies that δ_u is strictly increasing. Since $\mu(\cdot)$ is strictly increasing, from this inequality, we can further conclude that $\lim_{\beta\to\infty}\delta_u(\beta)=\infty$. $\square$

Proof of Theorem 5 : Recall the definitions for $f(\mu,\beta)$ and $f(\mu,\beta,\boldsymbol{\beta})$ from (2.21) and (3.7). Then $f(\mu,\beta,\mathbf{0})=f(\mu,\beta)$. We claim that for every $(\beta,\boldsymbol{\beta})$ and $\mu\in[-1,1]$, we have

$$f(\mu,\beta)\le f(\mu,\beta,\boldsymbol{\beta})\le f(\mu,\beta)+\frac{1}{2}\xi(1).$$

To prove this, let $\mathbb{E}_{\boldsymbol{g}}$ be the expectation on the randomness of the disorder $\boldsymbol{g}$ and $\mathbb{E}_h$ be the expectation on the randomness of $(h_i)_{i\le N}$. Then we can rewrite

$$\frac{1}{N}\mathbb{E}\ln Z_N^{SK}(\beta,\boldsymbol{\beta},h)-\frac{1}{N}\mathbb{E}\ln Z_N^{SK}(\beta,\mathbf{0},h)=\frac{1}{N}\mathbb{E}_h\mathbb{E}_{\boldsymbol{g}}\ln\left\langle\exp H_N^{SK}(\boldsymbol{\sigma})\right\rangle^{CW},$$

where $\langle\cdot\rangle^{CW}$ is the Gibbs average for the CW model. From Jensen's inequality and using $\mathbb{E}_{\boldsymbol{g}}\exp H_N^{SK}(\boldsymbol{\sigma})=\exp\left(N\xi(1)/2\right)$ and $\mathbb{E}_{\boldsymbol{g}}H_N^{SK}(\boldsymbol{\sigma})=0$ for every $\boldsymbol{\sigma}\in\Sigma_N$, the proof for our claim is completed since

$$\mathbb{E}_h\mathbb{E}_{\boldsymbol{g}}\ln\left\langle\exp H_N^{SK}(\boldsymbol{\sigma})\right\rangle^{CW}\le\mathbb{E}_h\ln\left\langle\mathbb{E}_{\boldsymbol{g}}\exp H_N^{SK}(\boldsymbol{\sigma})\right\rangle^{CW}=\frac{1}{2}N\xi(1)$$

and

$$\mathbb{E}_h\mathbb{E}_{\boldsymbol{g}}\ln\left\langle\exp H_N^{SK}(\boldsymbol{\sigma})\right\rangle^{CW}\ge\mathbb{E}_h\mathbb{E}_{\boldsymbol{g}}\left\langle H_N^{SK}(\boldsymbol{\sigma})\right\rangle^{CW}=0.$$

Now, suppose $(\beta,\boldsymbol{\beta})\in\mathcal{R}_u$. Recall from the definition of $\mathcal{R}_u$, $\beta>\beta_u$ and $\xi(1)\le 2\delta_u(\beta)$. From Proposition 4, since $\mu(\cdot)$ is strictly increasing, we have $\mu(\beta)>\mu(\beta_u)=u$ for every $\beta>\beta_u$. On the other hand, since $f(\cdot,\beta)$ is strictly increasing on $[0,\mu(\beta)]$, it follows that from the definition of $\mathcal{R}_u$ and our claim,

$$\begin{aligned}
f(\mu,\beta,\boldsymbol{\beta})&\le f(\mu,\beta)+\frac{1}{2}\xi(1)\\
&< f(u,\beta)+\frac{1}{2}\xi(1)\\
&= f(\mu(\beta),\beta)-\delta_u(\beta)+\frac{1}{2}\xi(1)\\
&\le f(\mu(\beta),\beta,\boldsymbol{\beta})-\delta_u(\beta)+\frac{1}{2}\xi(1)\\
&< f(\mu(\beta),\beta,\boldsymbol{\beta})
\end{aligned}$$

for every $\mu\in[0,u]$. Since h is centered, $f(\cdot,\beta)$ and $f(\cdot,\beta,\boldsymbol{\beta})$ are even functions on $[-1,1]$. Thus, we may also conclude $f(-\mu,\beta,\boldsymbol{\beta})<f(-\mu(\beta),\beta,\boldsymbol{\beta})$ for $\mu\in[0,u]$, which means

$$\Omega(\beta,\boldsymbol{\beta},h)=\mathrm{Argmax}_{\mu\in[-1,1]}f(\mu,\beta,\boldsymbol{\beta})\subset[-1,-u)\cup(u,1]$$

and we are done. $\square$

The following fundamental lemma will be used in the proof of Proposition 6.

Lemma 7. *Suppose that (X_N) is a sequence of random variables with $0 \le X_N \le 1$ for each N. If $\lim_{N\to\infty} \mathbb{E}X_N = 1/2$ and $\lim_{N\to\infty} \mathbb{E}X_N(1-X_N) = 0$, then $\{X_N\}$ converges to a Bernoulli$(1/2)$ r.v. weakly.*

Proof. First we claim that $\mathbb{E}X_N^n \to 1/2$ for each $n \ge 1$ by induction. From the given condition, this holds for $n = 1$. Suppose that this is true for some $n \ge 1$. Then using the fact that $0 \le X_N \le 1$, we obtain

$$\left|\mathbb{E}X_N^{n+1} - \mathbb{E}X_N^n\right| = \mathbb{E}X_N^n(1-X_N) \le \mathbb{E}X_N(1-X_N) \to 0.$$

Therefore, $\lim_{N\to\infty} \mathbb{E}X_N^{n+1} = \lim_{N\to\infty} \mathbb{E}X_N^n = 1/2$ and this completes the proof of our claim. Now, by using the dominated convergence theorem and our claim, the announced statement follows since

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\exp(itX_N) &= \lim_{N\to\infty} \sum_{n=0}^{\infty} \frac{(it)^n}{n!}\mathbb{E}X_N^n = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \lim_{N\to\infty} \mathbb{E}X_N^n \\ &= \frac{1}{2} + \frac{e^{it}}{2} = \mathbb{E}\exp(itX),\end{aligned}$$

where X is Bernoulli$(1/2)$.

□

Proof of Proposition 6 : From the definition of $\mathcal{R}_u$, Lemma 4, and Theorem 5, there exists some $\mu \in (u,1)$ such that $\Omega(\beta, \boldsymbol{\beta}, 0) = \{\mu, -\mu\}$. Since $(-u, u)$ has a positive distance to $\Omega(\beta, \boldsymbol{\beta}, h)$, Proposition 1 implies

$$\lim_{N\to\infty} \langle I(|m| \ge u)\rangle = 1. \tag{3.23}$$

If $\mathbb{E}h^2 = 0$, then $\langle I(m \in A)\rangle = \langle I(m \in -A)\rangle$ for every $A \subset [-1,1]$, where $-A := \{-x : x \in A\}$. Thus, the first statement follows from (3.23). Next, let $\mathbb{E}h^2 \ne 0$ and $1/2 < u < 1$. Recall that $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ are two configurations sampled independently from the Gibbs measure G_N with respect to the same realization $\boldsymbol{g}$. Set

$$m_1 = m_1(\boldsymbol{\sigma}^1) = \frac{1}{N}\sum_{i\le N} \sigma_i^1 \quad \text{and} \quad m_2 = m_2(\boldsymbol{\sigma}^2) = \frac{1}{N}\sum_{i\le N} \sigma_i^2.$$

We claim that

$$\{m_1 \in [u,1], m_2 \in [-1,-u]\} \subset \{R_{1,2} \le 1-2u\}. \tag{3.24}$$

Set

$$\begin{aligned} P_1^+ &= \left\{1 \le i \le N : \sigma_i^1 = 1\right\}, \ P_1^- = \left\{1 \le i \le N : \sigma_i^1 = -1\right\}, \\ P_2^+ &= \left\{1 \le i \le N : \sigma_i^2 = 1\right\}, \ P_2^- = \left\{1 \le i \le N : \sigma_i^2 = -1\right\}. \end{aligned}$$

Suppose $m_1 \in [u,1]$ and $m_2 \in [-1,-u]$. Let k be the smallest integer such that $u \le k/N$. Since $2|P_1^+| - N = |P_1^+| - |P_1^-| \ge k$ and $2|P_2^-| - N = |P_2^-| - |P_2^+| \ge k$, it implies $|P_1^+| \ge (k+N)/2$ and $|P_2^-| \ge (k+N)/2$. Consequently,

$$\begin{aligned}|P_1^+ \cap P_2^-| &= |P_1^+| - |P_1^+ \cap P_2^+| \ge |P_1^+| - |P_2^+| \\ &\ge \frac{k+N}{2} - \left(N - \frac{k+N}{2}\right) \ge k\end{aligned}$$

and our claim (3.24) follows from

$$\begin{aligned}
NR_{1,2} &= \sum_{i\le N} \sigma_i^1\sigma_i^2 \\
&= \sum_{P_1^+\cap P_2^+} \sigma_i^1\sigma_i^2 + \sum_{P_1^-\cap P_2^-} \sigma_i^1\sigma_i^2 + \sum_{P_1^+\cap P_2^-} \sigma_i^1\sigma_i^2 + \sum_{P_1^-\cap P_2^+} \sigma_i^1\sigma_i^2 \\
&= |P_1^+\cap P_2^+| + |P_1^-\cap P_2^-| - \left(|P_1^+\cap P_2^-| + |P_1^-\cap P_2^+|\right) \\
&\le |P_2^+| + |P_1^-| - |P_1^+\cap P_2^-| \\
&\le \left(N - \frac{k+N}{2}\right) + \left(N - \frac{k+N}{2}\right) - k \\
&= N - 2k \\
&\le N(1-2u).
\end{aligned}$$

Now, set $X_N = \langle I(m \ge u)\rangle$. From the independence of m^1 and m^2, $u > 1/2$, (3.24), the positivity of the overlap, and then (3.23), we obtain

$$\begin{aligned}
\mathbb{E}\left[X_N(1-X_N)\right] &= \mathbb{E}\left[\langle I(m_1 \ge u)\rangle \left(\langle I(m_2 \le -u)\rangle + \langle I(|m_2| < u)\rangle\right)\right] \\
&\le \mathbb{E}\left[\langle I(m_1 \ge u)\rangle \langle I(m_2 \le -u)\rangle\right] + \mathbb{E}\left[\langle I(|m_2| < u)\rangle\right] \\
&\le \mathbb{E}\langle I(R_{1,2} \le 1-2u)\rangle + \mathbb{E}\left[\langle I(|m_2| < u)\rangle\right] \\
&\to 0.
\end{aligned}$$

On the other hand, since h is centered, it is easy to derive $\mathbb{E}\langle I(m \ge u)\rangle = \mathbb{E}\langle I(m \le -u)\rangle$ and from (3.23), we deduce $\mathbb{E}X_N \to 1/2$. Consequently, from Lemma 7, (X_N) converges weakly to a Bernoulli(1/2) r.v. Write $X_N = \langle I(|m-\mu| \le \varepsilon)\rangle + Y_N$ for

$$Y_N := -\langle I(|m-\mu| \le \varepsilon, m < u)\rangle + \langle I(|m-\mu| > \varepsilon, m \ge u)\rangle .$$

If $0 < \varepsilon < \mu$, then

$$|Y_N| \le \langle I(0 \le m \le u)\rangle + \langle I(|m-\mu| > \varepsilon, m \ge u)\rangle \to 0 \ a.s.$$

and it follows that $\langle I(|m-\mu| \le \varepsilon)\rangle$ converges weakly to a Bernoulli(1/2) r.v. Since

$$\lim_{N\to\infty} \langle I(|m-\mu| \le \varepsilon, |m+\mu| \le \varepsilon)\rangle = 1$$

a.s., we also obtain that $\langle I(|m+\mu| \le \varepsilon)\rangle$ converges weakly to a Bernoulli(1/2) r.v. and this completes the proof of the second announced result. □

Acknowledgements. The author would like to thank Michel Talagrand and Alexander Vandenberg-Rodes for several helpful suggestions concerning the presentation of this paper. He would also like to thank anonymous referees for giving many crucial comments that lead to Theorem 4.

References

[1] Amaro de Matos, J. M. G., Patrick, A. E., Zagrebnov, V. A. (1992) Random infinite-volume Gibbs states for the Curie-Weiss random field Ising model. *J. Stat. Phys.*, **66**, no. 1-2, 139-164.

[2] Bertsekas, D. (1999) Nonlinear Programming, 2nd ed., Athena Scientific.

[3] Cadel, A., Rovira, C. (2010) The Sherrington Kirkpatrick model with ferromagnetic interaction. *Rocky Mountain J. Math.,* **40**, no. 5, 1441-1471.

[4] Chen, W.-K. (2012) The Aizenman-Sims-Starr scheme and Parisi formula for mixed p-spin spherical models. Arxiv:1204.5115.

[5] Chen, W.-K., Panchenko, D. (2012) An approach to chaos in some mixed p-spin models. Arxiv:1201.2198.

[6] Comets, F., Giacomin, G., Lebowitz, J. L. (1999) The Sherrington-Kirkpatrick model with short range ferromagnetic interactions. *C. R. Acad. Sci. Paris Sér. I Math.*, **328**, no. 1, 57-62.

[7] Comets, F., Guerra, F., Toninelli, F. L. (2005) The Ising-Sherrington-Kirkpatrick model in a magnetic field at high temperature. *J. Stat. Phys.*, **120**, 147-165.

[8] Guerra, F., Toninelli, F.L. (2002) The thermodynamical limit in mean field spin glass model. *Comm. Math. Phys.*, **230**, 71-79.

[9] Guerra, F., Toninelli, F.L. (2003) The infinite volume limit in generalized mean field disordered models, *Markov Proc. Rel. Fields*, **9**, 195-207.

[10] Guerra, F. (2003) Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, **233**, no. 1, 1-12.

[11] Mezard, M., Parisi, G., Virasoro, M. (1987) Spin Glass Theory and Beyond. World Scientific, Singapore.

[12] Panchenko, D. (2008) On the differentiability of the Parisi formula. *Elec. Comm. in Probab.*, **13**, 241-247.

[13] Panchenko, D. (2011)A new representation of the Ghirlanda-Guerra identities with applications. Arxiv:1108.0379.

[14] Panchenko, D. (2011) The Parisi ultrametricity conjecture. Arxiv:1112.1003.

[15] Panchenko, D. (2011) The Parisi formula for mixed p-spin models. The Arxiv:1112.4409.

[16] Phelps, R. (1989) Convex Functions, Monotone Operators and Differentiability. *Lecture Notes in Mathematics*, no. 1364, Springer.

[17] Roberts, A. W., Varberg, D. E. (1973) Convex Functions. Academic Press.

[18] Sherrington, D., Kirkpatrick, S. (1975) Solvable model of a spin glass. *Phys. Rev. Lett.*, **35**, 1792-1796.

[19] Talagrand, M. (2003) Spin Glasses: A Challenge for Mathematicians: Cavity and Mean Field Models. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge A Series of Modern Surveys in Mathematics, Vol. 46. Springer-Verlag.

[20] Talagrand, M. (2006) The Parisi formula. *Ann. of Math.* (2), **163**, no. 1, 221-263.

[21] Talagrand, M. (2006) Parisi measures. *J. Funct. Anal.*, **231**, 269-286.

[22] Talagrand, M. (2006) Free energy of the spherical mean filed model. *Probab. Theory Related Fields*, 134, no. 3, 339-382.

[23] Talagrand, M. (2010) Mean Field Models for Spin Glasses: Volume I: Basic Examples. Ergebnisse der Mathematik und ihrer Gren-zgebiete. 3. Folge A Series of Modern Surveys in Mathematics, Vol. 54. Springer-Verlag.

[24] Talagrand, M. (2011) Mean Field Models for Spin Glasses: Volume II: Advanced Replica-Symmetry and Low Temperature. Ergebnisse der Mathematik und ihrer Gren-zgebiete. 3. Folge A Series of Modern Surveys in Mathematics, Vol. 55. Springer-Verlag.

[25] Toubol, A. (1995) About the original Sherrington-Kirkpatrick model of spin glasses. *C. R. Acad. Sci. Paris Sér. I Math.*, **321**, no. 5, 617-622.