ON CONCENTRATION PROPERTIES
OF DISORDERED HAMILTONIANS

ANTONIO AUFFINGER AND WEI-KUO CHEN

(Communicated by Zhen-Qing Chen)

ABSTRACT. We present an elementary approach to concentration of disordered Hamiltonians. Assuming differentiability of the limiting free energy F with respect to the inverse temperature β, we show that the Hamiltonian concentrates around the energy level $F'(\beta)$ under the free energy and Gibbs average.

1. INTRODUCTION AND MAIN RESULTS

The aim of this short note is to establish a general principle on the concentration of disordered Hamiltonians, arising from various contexts of statistical mechanics models, assuming the differentiability of the limiting free energy. To begin with, for each $N \geq 1$, let $(\Sigma_N, \mathcal{F}_N)$ be a measurable space and ν_N be a random probability measure on this space. We call Σ_N a configuration space. A Hamiltonian H_N is a stochastic process indexed by Σ_N with

$$(1) \qquad \mathbb{E} \int_{\Sigma_N} \exp\bigl(\beta|H_N(\sigma)|\bigr)\nu_N(d\sigma) < \infty, \ \forall \beta > 0,$$

where $\mathbb{E}$ is the expectation with respect to the randomness of H_N and ν_N. For a given (inverse) temperature $\beta \geq 0$, the free energy and Gibbs measure associated to H_N are defined respectively as

$$F_N(\beta) = \frac{1}{N} \log Z_N(\beta)$$

and

$$G_{N,\beta}(d\sigma) = \frac{\exp\bigl(\beta H_N(\sigma)\bigr)\nu_N(d\sigma)}{Z_N(\beta)},$$

where

$$Z_N(\beta) := \int_{\Sigma_N} \exp\bigl(\beta H_N(\sigma)\bigr)\nu_N(d\sigma)$$

is called the partition function. Here the assumption (1) justifies the definiteness of F_N and $G_{N,\beta}$. We emphasize that F_N and $G_{N,\beta}$ are random objects depending on H_N and ν_N. In particular, from Hölder's inequality, F_N is convex in β. Denote by $\langle \cdot \rangle_\beta$ the expectation (Gibbs average) with respect to the Gibbs measure $G_{N,\beta}$, that is, for an integrable function $\Psi : \Sigma_N \to \mathbb{R}$, write

$$\langle \Psi \rangle_\beta = \frac{1}{Z_N(\beta)} \int_{\Sigma_N} \Psi(\sigma) \exp\bigl(\beta H_N(\sigma)\bigr)\nu_N(d\sigma).$$

We assume that the following condition is in force throughout the remainder of the paper. Let $\delta > 0$ and $\beta > 0$ be fixed. Suppose that there exists a nonrandom function $F : (\beta - \delta, \beta + \delta) \to \mathbb{R}$ such that for any $\beta' \in (\beta - \delta, \beta + \delta)$,

$$\lim_{N \to \infty} F_N(\beta') = F(\beta'), \ a.s. \tag{2}$$

The assumption of F being nonrandom appears in many examples of disordered systems as a consequence of concentration of measure, where in most cases

$$F(\beta) = \lim_{N \to \infty} \mathbb{E} F_N(\beta);$$

see Example 1 below. First we show that the Hamiltonian is concentrated around a fixed energy level under the free energy if we assume that F is differentiable at β.

Theorem 1. *Assume that* (2) *holds and F is differentiable at β. For any $\lambda > 0$, define*

$$\delta_N^+(\beta, \lambda) = \frac{F_N(\beta + \lambda) - F_N(\beta)}{\lambda} - F'(\beta),$$

$$\delta_N^-(\beta, \lambda) = \frac{F_N(\beta - \lambda) - F_N(\beta)}{\lambda} + F'(\beta).$$

For any $c, c' > 0$ with $c + c' < 1$, let

$$\varepsilon_N = N^{-c'} + \max\bigl(\delta_N^+(\beta, N^{-c}), \delta_N^-(\beta, N^{-c}), 0\bigr)$$

and set

$$C_N = \left\{ \sigma \in \Sigma_N : \left| \frac{H_N(\sigma)}{N} - F'(\beta) \right| \leq \varepsilon_N \right\}$$

for $N \geq 1$. We have that

(i) $\lim_{N \to \infty} \varepsilon_N = 0$ *a.s.*

(ii) $G_{N,\beta}(C_N^c) \leq 2e^{-N^{1-(c+c')}}$, *where C_N^c is the complement of C_N.*

(iii) *For any $N \geq \frac{\log \log 2}{1-(c+c')}$, the following inequality holds:*

$$\left| \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{C_N}(\sigma) \exp\bigl(\beta H_N(\sigma)\bigr) \nu_N(d\sigma) - F_N(\beta) \right| \leq \frac{2e^{-N^{1-(c+c')}}}{N},$$

where $\mathbf{1}_{C_N}$ is the indicator function on C_N.

Three remarks are in position.

Remark 1. Items (i) and (iii) say that in the computation of the limiting free energy, only the configurations associated to energies near the energy level $F'(\beta)$ with an error estimate ε_N would have nontrivial contributions. As an immediate consequence of (iii), one can compute the entropy of the configurations at the energy level $F'(\beta)$ by

$$\lim_{N \to \infty} \frac{1}{N} \log \nu_N(C_N) = F(\beta) - \beta F'(\beta). \tag{3}$$

Remark 2. From (3) and (ii), we can deduce that the conditional Gibbs measure on C_N at temperature β is equivalent, at exponential scale, to the probability measure

ν_N conditioning on C_N, that is, for any sequence of sets (A_N) with $A_N \subset \Sigma_N$ and $A_N \cap C_N \neq \emptyset$ for all $N \geq 1$,

$$\lim_{N \to \infty} \frac{1}{N} \log \frac{G_{N,\beta}(A_N | C_N)}{\nu_N(A_N | C_N)} = 0 \quad a.s.,$$

where $G_{N,\beta}(A|C)$ and $\nu_N(A|C)$ are the conditional probabilities of event A given C with respect to $G_{N,\beta}$ and ν_N, respectively. Indeed, this can be obtained by

$$\begin{aligned} \frac{1}{N} \log G_{N,\beta}(A_N | C_N) &= \frac{1}{N} \log G_{N,\beta}(A_N \cap C_N) + o(1) \\ &= \frac{1}{N} \log \nu_N(A_N \cap C_N) + \beta F'(\beta) - \frac{1}{N} \log Z_N + o(1) \\ &= \frac{1}{N} \log \nu_N(A_N \cap C_N) + \beta F'(\beta) - F(\beta) + o(1) \\ &= \frac{1}{N} \log \nu_N(A_N \cap C_N) - \frac{1}{N} \log \nu_N(C_N) + o(1) \\ &= \frac{1}{N} \log \nu_N(A_N | C_N) + o(1), \end{aligned}$$

where the first equality used (ii) and $o(1)$ is a function (that may change from line to line) such that $\lim_{N \to \infty} o(1) = 0$.

Remark 3. Naturally, one would wonder what will happen if F is not differentiable at β. In this case, it can be shown, following the argument of Theorem 1, that the Hamiltonian stays inside the interval $(D_- F(\beta) - \varepsilon_N', D_+ F(\beta) + \varepsilon_N')$ under the free energy, where $D_{\pm} F(\beta)$ are the right and left derivatives of F and the quantity ε_N' converges to zero. It would be of great interest to construct examples with their Hamiltonians being supported at more than one point in the interval $(D_- F(\beta), D_+ F(\beta))$.

Note that a classical result in convex analysis [10] states that the derivative of a differentiable convex function on an open interval must be continuous on that interval. Thus, if one further assumes that F is differentiable on an open interval $J \subseteq (\beta - \delta, \beta + \delta)$, then F' is automatically a continuous function on J. Conversely, the following theorem states that if there exists a continuous function E on some open interval $J \subset (\beta - \delta, \beta + \delta)$ such that the Hamiltonian is concentrated around E under the free energy for all temperature in J, then F must be differentiable in J.

Theorem 2. *If there exists an open interval $J \subseteq (\beta - \delta, \beta + \delta)$ and a continuous function $E : J \to \mathbb{R}$ such that for any $\beta \in J$ and $\varepsilon > 0$,*

$$\lim_{N \to \infty} \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{ \left| \frac{H_N(\sigma)}{N} - E(\beta) \right| \leq \varepsilon \right\}}(\sigma) \exp\left(\beta H_N(\beta) \right) \nu_N(d\sigma) = F(\beta), \ a.s.,$$

then F is differentiable on J and $E = F'$ on J.

We now turn our attention to the concentration of the Hamiltonian under the Gibbs average. This property has been derived under various assumptions and plays an essential role in statistical mechanics. Our main result below establishes concentration of the Hamiltonian under the Gibbs measure a.s.

Theorem 3. *Assume that (2) holds and F is differentiable at β. Then*

$$\lim_{N \to \infty} \left\langle \left| \frac{H_N(\sigma)}{N} - F'(\beta) \right| \right\rangle_\beta = 0, \ a.s.$$

A similar result was obtained in Panchenko [7], where he presented an approach, different from the one adopted in this paper, to establish

$$\text{(4)} \qquad \lim_{N\to\infty} \mathbb{E}\Big\langle \Big| \frac{H_N(\sigma)}{N} - \mathbb{E}\Big\langle \frac{H_N(\sigma)}{N} \Big\rangle_\beta \Big| \Big\rangle_\beta = 0$$

under the assumption that $\lim_{N\to\infty} \mathbb{E}|F_N - \mathbb{E}F_N| = 0$, $\lim_{N\to\infty} \mathbb{E}F_N = F$ in $(\beta - \delta, \beta + \delta)$, and F is differentiable at β. Note that Theorem 3 readily implies (4). Although the assumption (2) in Theorem 3 is stronger than Panchenko's setting described above, our argument of Theorem 3 indeed allows us to obtain the same statement as Panchenko [7]. In what follows, we give an example to demonstrate an application of (4) in the study of mean-field spin glasses [6].

Example 1 (Sherrington-Kirkpatrick model). For any $N \geq 1$, consider the hypercube $\Sigma_N = \{-1, +1\}^N$ and let ν_N be the uniform probability measure on Σ_N. The Sherrington-Kirkpatrick (SK) mean-field spin glass model is defined on Σ_N with Hamiltonian,

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij}\sigma_i\sigma_j,$$

where g_{ij}'s are i.i.d. standard normal for all $1 \leq i, j \leq N$. It can be easily computed that the covariance of H_N equals

$$\mathbb{E}H_N(\sigma)H_N(\sigma') = N\Big(\frac{1}{N}\sum_{i=1}^{N}\sigma_i\sigma_i'\Big)^2$$

for any $\sigma, \sigma' \in \Sigma_N$. As the covariance of H_N is of order N, the quantity $1/N$ in the definition of the free energy is the right scaling factor. In Guerra-Toninelli [5], it was known that the limiting free energy of the SK model converges a.s. and, by the virtue of Gaussian concentration of measure, this limit is a nonrandom function F of β. Furthermore, it was later established by Talagrand [11] that F admits a variational principle conjectured by Parisi [9]. One of the important consequences of Parisi's formula guarantees that F is differentiable for all $\beta > 0$ (see, e.g., [8, 12]), so the conclusions of Theorems 1 and 3 as well as (4) hold.

Denote by $(\sigma^\ell)_{\ell \geq 1}$ i.i.d. samplings from the Gibbs measure and by $\langle \cdot \rangle_\beta$ the Gibbs expectation with respect to this sequence. Set $R_{\ell,\ell'} = N^{-1}\sum_{i=1}^{N}\sigma_i^\ell\sigma_i^{\ell'}$ the overlap between $\sigma^\ell, \sigma^{\ell'}$, which measures the degree of similarities between the two configurations. Let $n \geq 2$ be fixed. Assume that ϕ is a bounded function of the overlaps $(R_{\ell,\ell'})_{1 \leq \ell \neq \ell' \leq n}$. Applying (4) yields

$$\text{(5)} \qquad \begin{aligned} &\Big| \mathbb{E}\Big\langle \phi \frac{H_N(\sigma^1)}{N} \Big\rangle_\beta - \mathbb{E}\langle\phi\rangle_\beta \mathbb{E}\Big\langle \frac{H_N(\sigma^1)}{N} \Big\rangle_\beta \Big| \\ &\leq \|\phi\|_\infty \mathbb{E}\Big\langle \Big| \frac{H_N(\sigma^1)}{N} - \mathbb{E}\Big\langle \frac{H_N(\sigma^1)}{N} \Big\rangle_\beta \Big| \Big\rangle_\beta \to 0. \end{aligned}$$

Recall the Gaussian integration by parts states that if z is a standard normal random variable, then $\mathbb{E}zf(z) = \mathbb{E}f'(z)$ for any absolutely continuous function f

with moderate growth. Utilizing this formula, we compute

$$\mathbb{E}\Big\langle \phi \frac{H_N(\sigma^1)}{N} \Big\rangle_\beta = \mathbb{E}\Big\langle \phi \sum_{\ell=1}^{n} R_{1,\ell}^2 \Big\rangle_\beta - n\mathbb{E}\langle \phi R_{1,n+1}^2 \rangle_\beta,$$

$$\mathbb{E}\Big\langle \frac{H_N(\sigma^1)}{N} \Big\rangle_\beta = \mathbb{E}\langle R_{1,1}^2 \rangle_\beta - \mathbb{E}\langle R_{1,2}^2 \rangle_\beta.$$

Plugging these two equations into (5) leads to

$$(6) \qquad \lim_{N\to\infty} \Big(\mathbb{E}\langle \phi R_{1,n+1}^2 \rangle_\beta - \frac{1}{n}\mathbb{E}\langle\phi\rangle_\beta \mathbb{E}\langle R_{1,2}^2 \rangle_\beta - \frac{1}{n}\sum_{\ell=2}^{n} \mathbb{E}\langle \phi R_{1,\ell}^2 \rangle_\beta \Big) = 0.$$

This is called the Ghirlanda-Guerra identity [4] for the SK model. By adding asymptotically vanishing perturbations, one can actually derive the extended Ghirlanda-Guerra identities, where (6) is not only valid for the second moment of the overlap, but also for any higher moments. These identities contain vital information about the Gibbs measure and ultimately connect to the computation of the limiting free energy. See, for instance, Panchenko [8] and Talagrand [12]. We also invite the readers to check some examples of quenched self-averaging of the Hamiltonian in Auffinger-Chen [1], Chatterjee [2], and Chen-Panchenko [3].

2. Proofs of Theorems 1 and 2 and Proposition 3

The idea of our proof for Theorem 1 is motivated by the standard derivation of the large deviation principle. The added technicality comes from the fact that $(F_N)_{N\geq 1}$ is a sequence of random Laplace transforms rather than deterministic ones. This difficulty will be overcome by applying the almost surely pointwise convergence of the sequence $(F_N)_{N\geq 1}$ stated in the following proposition.

Proposition 1. *Let F be a continuous function defined on an interval $I \subseteq \mathbb{R}$. If*

$$\lim_{N\to\infty} F_N(\beta) = F(\beta), \ a.s.$$

for any β in a dense subset of I, then

$$\mathbb{P}\Big(\lim_{N\to\infty} F_N(\beta) = F(\beta), \ \forall \beta \in I \Big) = 1.$$

To prove Proposition 1, we first need a lemma:

Lemma 1. *Let $I \subset \mathbb{R}$ be an open interval. Suppose that $\{f_n\}$ is a sequence of convex functions on I and f is a real-valued function on I. If $\lim_{n\to\infty} f_n(y) = f(y)$ pointwise on a dense subset $D \subset I$ and f is continuous at $y_0 \in \mathbb{R}$, then $\lim_{n\to\infty} f_n(y_0) = f(y_0)$.*

Proof. Let $y_0 \in \mathbb{R}$ be fixed. Suppose that f is continuous at y_0. Choose points $a, b, a', b' \in D$ with $a < b < y_0 < a' < b'$. By the convexity of f_n, for any $b < x < y < a'$,

$$\frac{f_n(b) - f_n(a)}{b - a} \leq \frac{f_n(y) - f_n(x)}{y - x} \leq \frac{f_n(b') - f_n(a')}{b' - a'}.$$

Since $\{f_n\}$ converges to f at a, b, a', b', this inequality means that $\{f_n\}$ is uniform Lipschitz on $[b, a']$ for all $n \geq 1$ with Lipschitz constant $M > 0$. Consequently, for

any $y \in D \cap [b, a']$,

$$|f_n(y_0) - f(y_0)| \leq |f_n(y_0) - f_n(y)| + |f_n(y) - f(y)| + |f(y) - f(y_0)|$$
$$\leq M|y_0 - y| + |f_n(y) - f(y)| + |f(y) - f(y_0)|$$

and passing to limit gives

$$\limsup_{n \to \infty} |f_n(y_0) - f(y_0)| \leq M|y_0 - y| + |f(y) - f(y_0)|.$$

Since this holds for any $y \in D \cap [b, a']$, letting $y \in D \to y_0$ finishes our proof. □

Proof of Proposition 1. Pick a countable dense subset D of I. Denote by $\Omega(\beta)$ the event that $(F_N(\beta))$ converges. Let $\Omega = \bigcap_{\beta \in D} \Omega(\beta)$. Note that $\mathbb{P}(\Omega) = 1$. Therefore, on Ω, $\lim_{N \to \infty} F_N(\beta) = F(\beta)$ for any $\beta \in D$. Since D is dense and F is continuous everywhere, Lemma 1 implies that $\lim_{N \to \infty} F_N = F$ on pointwise on I with probability one. □

Proof of Theorem 1. Let $\lambda_N = N^{-c}$. To verify (i), suppose that we are on the event,

$$\Big\{ \lim_{N \to \infty} F_N(\beta') = F(\beta'), \ \forall \beta' \in (\beta - \delta, \beta + \delta) \Big\}.$$

From the convexity of F_N, for any $0 < \eta < \delta/2$, as long as N is large enough, we have

$$\frac{F_N(\beta) - F_N(\beta - \eta)}{\eta} \leq \frac{F_N(\beta) - F_N(\beta - \lambda_N)}{\lambda_N}$$
$$\leq \frac{F_N(\beta + \lambda_N) - F_N(\beta)}{\lambda_N} \leq \frac{F_N(\beta + \eta) - F_N(\beta)}{\eta}.$$

Here since the left and right sides converge to

$$\frac{F(\beta) - F(\beta - \eta)}{\eta} \quad \text{and} \quad \frac{F(\beta + \eta) - F(\beta)}{\eta},$$

the differentiability of F shows that these two quantities are equal to each other as $\eta \downarrow 0$. To sum up,

$$(7) \qquad \lim_{N \to \infty} \frac{F_N(\beta + \lambda_N) - F_N(\beta)}{\lambda_N} = \lim_{N \to \infty} \frac{F_N(\beta) - F_N(\beta - \lambda_N)}{\lambda_N} = F'(\beta).$$

Note that since F_N is convex, F is continuous on $(\beta - \delta, \beta + \delta)$. From Proposition 1 and (7), (i) follows.

The proof of (ii) and (iii) given below is the main novelty of the paper. For any $\varepsilon > 0$, define

$$B_N^+(\beta, \varepsilon) = \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{A_N^+(\varepsilon)}(\sigma) \exp\big(\beta H_N(\sigma)\big) \nu_N(d\sigma),$$

$$B_N^-(\beta, \varepsilon) = \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{A_N^-(\varepsilon)}(\sigma) \exp\big(\beta H_N(\sigma)\big) \nu_N(d\sigma),$$

where

$$A_N^+(\varepsilon) := \Big\{ \sigma \in \Sigma_N : \frac{H_N(\sigma)}{N} - F'(\beta) > \varepsilon \Big\},$$

$$A_N^-(\varepsilon) := \Big\{ \sigma \in \Sigma_N : \frac{H_N(\sigma)}{N} - F'(\beta) < -\varepsilon \Big\}.$$

Observe that for any $0 < \lambda < \delta$,

$$
\begin{aligned}
B_N^+(\beta, \varepsilon) &\leq \frac{1}{N} \log \int_{\Sigma_N} \exp\big((\beta + \lambda) H_N(\sigma) - \lambda N\big(F'(\beta) + \varepsilon\big)\big) \nu_N(d\sigma) \\
&= F_N(\beta + \lambda) - \lambda(\varepsilon + F'(\beta))
\end{aligned}
$$
(8)

and

$$
\begin{aligned}
B_N^-(\beta, \varepsilon) &\leq \frac{1}{N} \log \int_{\Sigma_N} \exp\big((\beta - \lambda) H_N(\sigma) + \lambda N\big(F'(\beta) - \varepsilon\big)\big) \nu_N(d\sigma) \\
&= F_N(\beta - \lambda) - \lambda(\varepsilon - F'(\beta)).
\end{aligned}
$$
(9)

To control these inequalities, write

$$
\begin{aligned}
F_N(\beta \pm \lambda) \mp \lambda F'(\beta) - \lambda\varepsilon &= F_N(\beta) - \lambda(\varepsilon - \delta_N^\pm(\beta, \lambda)) \\
&\leq F_N(\beta) - \lambda(\varepsilon - \gamma_N(\beta, \lambda)),
\end{aligned}
$$
(10)

where

$$
\gamma_N(\beta, \lambda) := \max\big(\delta_N^+(\beta, \lambda), \delta_N^-(\beta, \lambda), 0\big), \ \forall 0 < \lambda < \delta.
$$

From (8), (9), and (10), it follows that

$$
G_{N,\beta}\big(A_N^\pm(\varepsilon)\big) \leq e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))}.
$$
(11)

Consequently, we have

$$
G_{N,\beta}\big(A_N^+(\varepsilon) \cup A_N^-(\varepsilon)\big) \leq 2e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))}.
$$
(12)

From now on, take $\lambda = \lambda_N$ and $\varepsilon = \varepsilon_N$. Using the bound

$$
e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))} \leq e^{-N \cdot N^{-c} \cdot N^{-c'}} < e^{-N^{1-(c+c')}},
$$

the inequality (12) gives (ii). On the other hand, observe that

$$
\begin{aligned}
G_{N,\beta}(C_N) &= 1 - G_{N,\beta}\big(A_N^+(\varepsilon) \cup A_N^-(\varepsilon)\big) \\
&\geq 1 - 2e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))} \\
&> 1 - 2e^{-N^{1-(c+c')}},
\end{aligned}
$$

from which taking $N^{-1} \log$ on both sides yields that

$$
\begin{aligned}
F_N(\beta) &\geq \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{C_N}(\sigma) \exp\big(\beta H_N(\sigma)\big) \nu_N(d\sigma) \\
&\geq F_N(\beta) + \frac{1}{N} \log\big(1 - 2e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))}\big) \\
&\geq F_N(\beta) + \frac{1}{N} \log\big(1 - 2e^{-N^{1-(c+c')}}\big).
\end{aligned}
$$

This implies (iii) and finishes our proof. $\qquad\square$

Remark 4. From the above proof, under the same assumption and noting $F_N'(\beta) = N^{-1}\langle H_N(\sigma)\rangle_\beta$, one can derive by the same argument to obtain an identical statement as Theorem 1 by replacing every $F'(\beta)$ by $F_N'(\beta)$.

Proof of Theorem 2. Let $\beta \in J$ be fixed. For $\varepsilon > 0$, observe that for any $\beta' < \beta$ and β' being sufficiently close to β, the continuity of E gives

$$\frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta)\right| \leq \varepsilon\right\}}(\sigma) e^{\beta H_N(\sigma)} \nu_N(d\sigma)$$

$$\leq \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta')\right| \leq 2\varepsilon\right\}}(\sigma) e^{\beta H_N(\sigma)} \nu_N(d\sigma)$$

$$\leq \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta')\right| \leq 2\varepsilon\right\}}(\sigma) e^{\beta' H_N(\sigma)} \nu_N(d\sigma) + (\beta - \beta')\big(E(\beta') + 2\varepsilon\big)$$

and

$$\frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta)\right| \leq 2\varepsilon\right\}}(\sigma) e^{\beta H_N(\sigma)} \nu_N(d\sigma)$$

$$\geq \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta')\right| \leq \varepsilon\right\}}(\sigma) e^{\beta H_N(\sigma)} \nu_N(d\sigma)$$

$$\geq \frac{1}{N} \log \int_{\Sigma_N} \mathbf{1}_{\left\{\left|\frac{H_N(\sigma)}{N} - E(\beta')\right| \leq \varepsilon\right\}}(\sigma) e^{\beta' H_N(\sigma)} \nu_N(d\sigma) + (\beta - \beta')\big(E(\beta') - \varepsilon\big).$$

By letting $N \to \infty$, the given assumption leads to

$$(\beta - \beta')\big(E(\beta') - \varepsilon\big) \leq F(\beta) - F(\beta') \leq (\beta - \beta')\big(E(\beta') + 2\varepsilon\big),$$

from this and the continuity of E, F is left differentiable at β. Note that one may actually interchange the role of β and β' in the last inequality to get that for $\beta' > \beta$,

$$(\beta' - \beta)\big(E(\beta) - \varepsilon\big) \leq F(\beta') - F(\beta) \leq (\beta' - \beta)\big(E(\beta) + 2\varepsilon\big).$$

This leads to the right differentiability of F at β. All these together imply that F is differentiable at β and $F'(\beta) = E(\beta)$. This finishes our proof. $\qquad\square$

Proof of Theorem 3. Recall that (11) holds for all $\varepsilon > 0$. Integrating against ε leads to

$$(13) \qquad \left\langle \left(\frac{H_N(\sigma)}{N} - F'(\beta)\right)_{\pm} \right\rangle_{\beta} \leq \int_0^{\infty} e^{-N\lambda(\varepsilon - \gamma_N(\beta, \lambda))} d\varepsilon = \frac{1}{N\lambda} e^{N\lambda \gamma_N(\beta, \lambda)}$$

for all $\lambda \in (0, \delta)$, where $x_+ := \max(x, 0)$ and $x_- := \max(-x, 0)$ for any $x \in \mathbb{R}$. Let $0 < \lambda_0 < \delta$. Observe that by the convexity of F_N, for $0 < \lambda < \lambda_0$, we have

$$\gamma_N(\beta, \lambda) \leq \gamma_N(\beta, \lambda_0).$$

If $\lambda_0 N \gamma_N(\beta, \lambda_0) \leq 1$, we take $\lambda = \lambda_0$. If $\lambda_0 N \gamma_N(\beta, \lambda_0) > 1$, then $1/N\gamma_N(\beta, \lambda_0) \leq \lambda_0$ and we let $\lambda = 1/N\gamma_N(\beta, \lambda_0)$. Thus, we conclude from (13) that

$$\left\langle \left(\frac{H_N(\sigma)}{N} - F'(\beta)\right)_{\pm} \right\rangle_{\beta} \leq e \max\left(\frac{1}{N\lambda_0}, \gamma_N(\beta, \lambda)\right).$$

As a result, for any $0 < \lambda_0 < \delta$,

$$\left\langle \left|\frac{H_N(\sigma)}{N} - F'(\beta)\right| \right\rangle_{\beta} \leq \frac{2e}{N\lambda_0} + 2e\gamma_N(\beta, \lambda_0)$$

and passing to limit gives

$$\limsup_{N \to \infty} \left\langle \left|\frac{H_N(\sigma)}{N} - F'(\beta)\right| \right\rangle_{\beta} \leq 2e \limsup_{N \to \infty} \gamma_N(\beta, \lambda_0).$$

Here the right-hand side tends to zero as $\lambda_0 \downarrow 0$ by a similar argument as (7). This completes our proof. $\qquad\square$

Acknowledgments

References

[1] Antonio Auffinger and Wei-Kuo Chen, *Universality of chaos and ultrametricity in mixed p-spin models*, Comm. Pure Appl. Math. **69** (2016), no. 11, 2107–2130, DOI 10.1002/cpa.21617. MR3552010

[2] Sourav Chatterjee, *Absence of replica symmetry breaking in the random field Ising model*, Comm. Math. Phys. **337** (2015), no. 1, 93–102, DOI 10.1007/s00220-014-2269-5. MR3324156

[3] Wei-Kuo Chen and Dmitry Panchenko, *Some examples of quenched self-averaging in models with Gaussian disorder* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 243–258, DOI 10.1214/15-AIHP715. MR3606741

[4] Stefano Ghirlanda and Francesco Guerra, *General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity*, J. Phys. A **31** (1998), no. 46, 9149–9155, DOI 10.1088/0305-4470/31/46/006. MR1662161

[5] Francesco Guerra and Fabio Lucio Toninelli, *The thermodynamic limit in mean field spin glass models*, Comm. Math. Phys. **230** (2002), no. 1, 71–79, DOI 10.1007/s00220-002-0699-y. MR1930572

[6] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro, *Spin glass theory and beyond*, World Scientific Lecture Notes in Physics, vol. 9, World Scientific Publishing Co., Inc., Teaneck, NJ, 1987. MR1026102

[7] Dmitry Panchenko, *The Ghirlanda-Guerra identities for mixed p-spin model* (English, with English and French summaries), C. R. Math. Acad. Sci. Paris **348** (2010), no. 3-4, 189–192, DOI 10.1016/j.crma.2010.02.004. MR2600075

[8] Dmitry Panchenko, *The Sherrington-Kirkpatrick model*, Springer Monographs in Mathematics, Springer, New York, 2013. MR3052333

[9] G. Parisi, Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, **43**, 1754–1756 (1979)

[10] A. Wayne Roberts and Dale E. Varberg, *Convex functions*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1973. Pure and Applied Mathematics, Vol. 57. MR0442824

[11] Michel Talagrand, *The Parisi formula*, Ann. of Math. (2) **163** (2006), no. 1, 221–263, DOI 10.4007/annals.2006.163.221. MR2195134

[12] M. Talagrand, Mean field models for spin glasses. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, **55**, Springer-Verlag, Berlin (2011)

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, Illinois 60208

Email address: auffing@math.northwestern.edu

School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church Street SE, Minneapolis, Minnesota 55455

Email address: wkchen@umn.edu