Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2009, Vol. 45, No. 4, 1020–1047
DOI: 10.1214/08-AIHP193

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Cavity method in the spherical SK model[1]

Dmitry Panchenko

Department of Mathematics, Texas A&M University, Mailstop 3386, Room 209, College Station, TX 77843, USA.
E-mail: panchenk@math.tamu.edu

Received 13 February 2008; revised 6 September 2008; accepted 15 September 2008

Abstract. We develop a cavity method for the spherical Sherrington–Kirkpatrick model at high temperature and small external field. As one application we compute the limit of the covariance matrix for fluctuations of the overlap and magnetization.

Résumé. Nous développons la méthode de la cavité pour le modèle sphérique de Sherrington–Kirkpatrick à haute température et champs externe faible. Nous illustrons la méthode par le calcul de la matrice de covariance des fluctuations des recouvrements et de la magnétisation.

MSC: 60K35; 82B44

Keywords: Sherrington–Kirkpatrick model; Cavity method

1. Introduction

The cavity method in the Sherrington–Kirkpatrick model [3], as described for example in [5], is one of the most important tools developed for the analysis of the model at high temperature. Among many applications of the cavity method, the most important include self-averaging and Gaussian fluctuations of the overlap (see [4] or [2]). The present paper was motivated by an attempt to find an analogue of the cavity method for the spherical model and, surprisingly, the task turned out to be more challenging than expected. Of course, one can anticipate technical difficulties due to the fact that uniform measure on the sphere is not a product measure and decoupling one coordinate from the others, which is the idea of the cavity method, is not straightforward. As we shall see, the suggested interpolation has some new features compared to [5].

After we develop the cavity method, as one application we will study the fluctuations of the overlap and magnetization and compute their covariance matrix in the thermodynamic limit. We stop short of proving a central limit theorem since our goal is to give a reasonably simple illustration of the cavity method.

We consider a spherical SK model with Gaussian Hamiltonian (process) $H_N(\boldsymbol{\sigma})$ indexed by *spin configurations* $\boldsymbol{\sigma}$ on the sphere S_N of radius $\sqrt{N}$ in $\mathbb{R}^N$. We will assume that

$$\frac{1}{N}\mathbb{E}H_N\big(\boldsymbol{\sigma}^1\big)H_N\big(\boldsymbol{\sigma}^2\big)=\xi(R_{1,2}), \tag{1.1}$$

[1]Supported in part by NSF grant.

where $R_{1,2}=N^{-1}\sum_{i\le N}\sigma_i^1\sigma_i^2$ is *the overlap* of configurations $\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\in S_N$ and where the function $\xi(x)$ is three times continuously differentiable. A classical example of such Gaussian process, corresponding to the choice of $\xi(x)=x^p$ for integer $p\ge 1$, is the p-spin Hamiltonian

$$H_N(\boldsymbol{\sigma})=\frac{1}{N^{(p-1)/2}}\sum_{1\le i_1,\dots,i_p\le N} g_{i_1,\dots,i_p}\sigma_{i_1}\cdots\sigma_{i_p},$$

where $(g_{i_1,\dots,i_p})$ are i.i.d. standard Gaussian random variables. One could also consider a linear combination of p-spin Hamiltonians.

This model was studied in [1] and mathematically rigorous results were proved in [6]. Under the additional assumptions on ξ,

$$\xi(0)=0,\qquad \xi(x)=\xi(-x),\qquad \xi''(x)>0\quad \text{if } x>0, \tag{1.2}$$

the limit of the free energy

$$F_N=\frac{1}{N}\mathbb{E}\log\int_{S_N}\exp\biggl(\beta H_N(\boldsymbol{\sigma})+h\sum_{i\le N}\sigma_i\biggr)\lambda_N(\boldsymbol{\sigma}) \tag{1.3}$$

was computed in [6] for arbitrary *inverse temperature* $\beta>0$ and *external field* $h\in\mathbb{R}$. Here λ_N denotes the uniform probability measure on S_N. The main results of the present paper will be proved for small enough parameters β and h, i.e. for very high temperature and small external field, and *without* the assumptions in (1.2). For example, our results will apply to p-spin Hamiltonian for both even and odd $p\ge 1$.

To provide some motivation, let us first describe some implications of the results in [6]. For small β and h, they imply that under (1.2) the limit of the free energy takes a particularly simple form:

$$\lim_{N\to\infty}F_N=\inf_{q\in[0,1]}\frac{1}{2}\biggl(h^2(1-q)+\frac{q}{1-q}+\log(1-q)+\beta^2\xi(1)-\beta^2\xi(q)\biggr). \tag{1.4}$$

The critical point equation for the infimum on the right hand side of (1.4) is

$$h^2+\beta^2\xi'(q)=\frac{q}{(1-q)^2}. \tag{1.5}$$

For small enough β the infimum in (1.4) is achieved at $q=0$ if $h=0$ and at the unique solution q of (1.5) if $h\ne 0$. Theorem 1.2 in [6] suggests that the distribution of the overlap $R_{1,2}$ with respect to the Gibbs measure is concentrated near q and by analogy with the classical SK model (see Chapter 2 in [5] or [2]) one expects that the distribution of $\sqrt{N}(R_{1,2}-q)$ is approximately Gaussian. In the setting of the classical SK model this is proved using the cavity method and the main goal of the present paper will be to develop the analogue of the cavity method for the spherical SK model. Many computations will be more involved and, as a result, instead of using the cavity method to prove a central limit theorem we will only compute the covariance matrix of the overlaps and other related quantities. This main application is formulated in Theorem 5 in Section 3.

We note that our results also imply (1.4) without the assumption (1.2). Namely, since we will prove that for small β and h the overlap $R_{1,2}$ is concentrated near the unique solution q of (1.5), it is a simple exercise to show that in this case

$$\lim_{N\to\infty}F_N=\frac{1}{2}\biggl(h^2(1-q)+\frac{q}{1-q}+\log(1-q)+\beta^2\xi(1)-\beta^2\xi(q)\biggr). \tag{1.6}$$

To prove this, one only needs to compare the derivatives of both sides with respect to β. Finally, we suggest to an interested reader not familiar with the original cavity method to learn about it in Chapter 2 in [5], because we believe that one can gain much more by first learning a technically simpler case from which all the main ideas originate.

2. Cavity method

For specificity, from now on we assume that $h \neq 0$ and β is small enough so that q is the *unique* solution of (1.5). The case of $h = 0$ is also significantly simpler because one can use the second moment method as in [5], Section 2.2. All the results below are proved without the assumption (1.2). Given a configuration $\boldsymbol{\sigma} \in S_N$, we will denote $\varepsilon = \sigma_N$ and for $i \leq N-1$ denote

$$\hat{\sigma}_i = \sigma_i \Big/ \sqrt{\frac{N-\varepsilon^2}{N-1}},$$

so that a vector $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1, \ldots, \hat{\sigma}_{N-1}) \in S_{N-1}$, i.e. $|\hat{\boldsymbol{\sigma}}| = \sqrt{N-1}$. We consider a Gaussian Hamiltonian $H_{N-1}(\hat{\boldsymbol{\sigma}})$ independent of $H_N(\boldsymbol{\sigma})$ such that

$$\frac{1}{N-1}\mathbb{E} H_{N-1}\big(\hat{\boldsymbol{\sigma}}^1\big) H_{N-1}\big(\hat{\boldsymbol{\sigma}}^2\big) = \xi(\hat{R}_{1,2}), \tag{2.1}$$

where $\hat{R}_{1,2} = (N-1)^{-1}\sum_{i \leq N-1} \hat{\sigma}_i^1 \hat{\sigma}_i^2$. We define the *interpolating Hamiltonian* by

$$\begin{aligned} H_t(\boldsymbol{\sigma}) = {} & \sqrt{t}\beta H_N(\boldsymbol{\sigma}) + \sqrt{1-t}\beta H_{N-1}(\hat{\boldsymbol{\sigma}}) + h\sum_{i \leq N-1} \hat{\sigma}_i \left(1 + t\left(\sqrt{\frac{N-\varepsilon^2}{N-1}} - 1\right)\right) \\ & + h\varepsilon + \sqrt{1-t}\varepsilon z \beta \sqrt{\xi'(q)} - \frac{1}{2}(1-t)\varepsilon^2 b, \end{aligned} \tag{2.2}$$

where z is a standard Gaussian r.v. independent of H_N and H_{N-1} and

$$b = h^2(1-q) + \beta^2(1-q)\xi'(q). \tag{2.3}$$

Define the *partition function* along the interpolation by

$$Z_t = \int_{S_N} \exp H_t(\boldsymbol{\sigma})\, \mathrm{d}\lambda_N(\boldsymbol{\sigma})$$

and for a function $f : S_N^n \to \mathbb{R}$ define the *Gibbs average* of f with respect to the Hamiltonian (2.2) by

$$\langle f \rangle_t = \frac{1}{Z_t^n} \int_{S_N^n} f \exp \sum_{l \leq n} H_t\big(\boldsymbol{\sigma}^l\big)\, \mathrm{d}\lambda_N^n. \tag{2.4}$$

Let $\nu_t(f) = \mathbb{E}\langle f \rangle_t$ be the annealed Gibbs average.

The main purpose of this interpolation is to devise a way of computing $\nu(f) := \nu_1(f)$ for some particular choices of f – typically, functions of the overlaps – and finding the right expression for H_t is always motivated by the properties that such interpolation should have in order to make these computations work.

First of all, any cavity method aims to decouple the last coordinate ε from the other coordinates to make the computation of $\nu_0(f)$ easier, in some sense. Notice the special form of the Hamiltonian (2.5) at $t = 0$,

$$H_0(\boldsymbol{\sigma}) = \beta H_{N-1}(\hat{\boldsymbol{\sigma}}) + h\sum_{i \leq N-1} \hat{\sigma}_i + \varepsilon a - \frac{1}{2}\varepsilon^2 b, \tag{2.5}$$

where we introduced the notation

$$a = z\beta\sqrt{\xi'(q)} + h. \tag{2.6}$$

The first two terms depend only on $\hat{\boldsymbol{\sigma}} \in S_{N-1}$ and, therefore, the Gibbs average (2.4) at $t = 0$ for functions of the type $f_1(\hat{\boldsymbol{\sigma}}) f_2(\varepsilon)$ will decouple (see (2.15) below). In other words, at $t = 0$ we can think of ε and $\hat{\boldsymbol{\sigma}}$ as independent variables, even though $\hat{\boldsymbol{\sigma}}$ formally depends on ε. This explains a somewhat unusual dependence of H_{N-1} on $(\hat{\sigma}_1,\ldots,\hat{\sigma}_{N-1})$ in the above interpolation instead of the more expected dependence on $(\sigma_1,\ldots,\sigma_{N-1})$. This is probably the main new feature compared to the classical SK model on the hypercube $\{-1,+1\}^N$.

Another desired property of the interpolation is that $\nu_t(f)$ does not change much for $t\in[0,1]$ so that $\nu(f)$ can indeed be approximated by $\nu_0(f)$. The proof of this will have several ingredients and, of course, the appropriate choice of parameter q in (1.5) will play an important role at some point. The main ingredients, however, are the computation and control of the derivative of $\nu_t(f)$ with respect to t. This computation will be a first step toward clarifying the role of all the terms in (2.2). For q in (1.5) we define

$$r=h(1-q). \tag{2.7}$$

Given $(\boldsymbol{\sigma}^l)_{l\geq 1}$ we denote $\varepsilon_l=\sigma_N^l$, $\hat{R}_l=(N-1)^{-1}\sum_{i\leq N-1}\hat{\sigma}_i^l$ and define a_l and $a_{l,l'}$ by

$$a_l=1-\varepsilon_l^2,\qquad 2a_{l,l'}=\xi'(q)-\frac{1}{2}\bigl(\varepsilon_l^2+\varepsilon_{l'}^2\bigr)\bigl(q\xi''(q)+\xi'(q)\bigr)+\varepsilon_l\varepsilon_{l'}\xi''(q). \tag{2.8}$$

Theorem 1. *We have*

$$\begin{aligned}\nu_t'(f)={}&\frac{h}{2}\sum_{l\leq n}\nu_t\bigl(fa_l(\hat{R}_l-r)\bigr)-\frac{h}{2}n\nu_t\bigl(fa_{n+1}(\hat{R}_{n+1}-r)\bigr)\\&+2\beta^2\sum_{1\leq l<l'\leq n}\nu_t\bigl(fa_{l,l'}(\hat{R}_{l,l'}-q)\bigr)-2n\beta^2\sum_{l\leq n}\nu_t\bigl(fa_{l,n+1}(\hat{R}_{l,n+1}-q)\bigr)\\&+n(n+1)\beta^2\nu_t\bigl(fa_{n+1,n+2}(\hat{R}_{n+1,n+2}-q)\bigr)+\nu_t(f\mathcal{R}),\end{aligned} \tag{2.9}$$

where the remainder $\mathcal{R}$ is bounded by

$$|\mathcal{R}|\leq\frac{L}{N}\bigl(\beta^2+h\bigr)\Bigl(1+\sum_{l\leq n+2}\varepsilon_l^4\Bigr)+L\beta^2\sum_{1\leq l\neq l'\leq n+2}\bigl(1+\varepsilon_l^2\bigr)(\hat{R}_{l,l'}-q)^2.$$

This computation will be carried out in Section 4. In order for the interpolation to be useful, the above derivative should be small. Notice that all the terms in (2.9) have factors of four types $a_l, a_{l,l'}$, $(\hat{R}_l-r)$ or $(\hat{R}_{l,l'}-q)$ which is why our goal will be to show that the last coordinate ε is of order one and the overlap $\hat{R}_{l,l'}$ and magnetization $\hat{R}_l$ are concentrated near constants q and r. The corresponding results will be proved in Sections 5 and 6.

Theorem 2. *If β and h are small enough, we can find a constant $L>0$ such that*

$$\nu_t\Bigl(\exp\frac{1}{L}\varepsilon^2\Bigr)\leq L \tag{2.10}$$

for all $t\in[0,1]$.

Theorem 3. *If β and h are small enough then for any $K>0$ we can find $L>0$ such that*

$$\nu_t\biggl(I\biggl(|\hat{R}_{1,2}-q|\geq L\Bigl(\frac{\log N}{N}\Bigr)^{1/4}\biggr)\biggr)\leq\frac{L}{N^K}, \tag{2.11}$$

$$\nu_t\biggl(I\biggl(|\hat{R}_1-r|\geq L\Bigl(\frac{\log N}{N}\Bigr)^{1/4}\biggr)\biggr)\leq\frac{L}{N^K} \tag{2.12}$$

for all $t\in[0,1]$.

A step in the proof of Theorem 3, where the conditions (1.5) and (2.7) that define parameters q and r finally appear, will further clarify the role of all the terms in (2.2). In some sense, this is where all the pieces will fall together.

Finally, let us explain what happens at the end of the interpolation at $t = 0$. We start by writing the integration over S_N as a double integral over ε and $(\sigma_1, \dots, \sigma_{N-1})$. Let λ_N^ρ denote the area measure on the sphere S_N^ρ of radius ρ in $\mathbb{R}^N$, and let $|S_N^\rho|$ denote its area, i.e. $|S_N^\rho| = \lambda_N^\rho(S_N^\rho)$. Then,

$$\begin{aligned}
\int_{S_N} f(\boldsymbol{\sigma})\,\mathrm{d}\lambda_N(\boldsymbol{\sigma}) &= \frac{1}{|S_N^{\sqrt{N}}|}\int_{S_N^{\sqrt{N}}} f(\sigma_1,\dots,\sigma_N)\,\mathrm{d}\lambda_N^{\sqrt{N}}(\sigma_1,\dots,\sigma_N)\\
&= \frac{1}{|S_N^{\sqrt{N}}|}\int_{-\sqrt{N}}^{\sqrt{N}} \frac{\mathrm{d}\varepsilon}{\sqrt{1-\varepsilon^2/N}}\int_{S_{N-1}^{\sqrt{N-\varepsilon^2}}} f(\sigma_1,\dots,\sigma_{N-1},\varepsilon)\,\mathrm{d}\lambda_{N-1}^{\sqrt{N-\varepsilon^2}}(\sigma_1,\dots,\sigma_{N-1})\\
&= \int_{-\sqrt{N}}^{\sqrt{N}} \frac{|S_{N-1}^{\sqrt{N-\varepsilon^2}}|}{|S_N^{\sqrt{N}}|}\frac{\mathrm{d}\varepsilon}{\sqrt{1-\varepsilon^2/N}}\int_{S_{N-1}} f\left(\hat{\sigma}_1\sqrt{\frac{N-\varepsilon^2}{N-1}},\dots,\hat{\sigma}_{N-1}\sqrt{\frac{N-\varepsilon^2}{N-1}},\varepsilon\right)\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}})\\
&= a_N\int_{-\sqrt{N}}^{\sqrt{N}} \mathrm{d}\varepsilon\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2}\int_{S_{N-1}} f\left(\hat{\boldsymbol{\sigma}}\sqrt{\frac{N-\varepsilon^2}{N-1}},\varepsilon\right)\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}}),
\end{aligned}\tag{2.13}$$

where $a_N = |S_{N-1}^1|/(|S_N^1|\sqrt{N}) \to (2\pi)^{-1/2}$ as can be seen by taking $f = 1$. In particular, if

$$f(\boldsymbol{\sigma}) = f_1(\varepsilon) f_2(\hat{\boldsymbol{\sigma}}),$$

then

$$\int_{S_N} f(\boldsymbol{\sigma})\,\mathrm{d}\lambda_N(\boldsymbol{\sigma}) = a_N\int_{-\sqrt{N}}^{\sqrt{N}} f_1(\varepsilon)\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2}\mathrm{d}\varepsilon\int_{S_{N-1}} f_2(\hat{\boldsymbol{\sigma}})\,\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}}).\tag{2.14}$$

Since the Hamiltonian (2.5) decomposed into the sum of terms that depend only on ε or only on $\hat{\boldsymbol{\sigma}}$, (2.14) implies that

$$\langle f\rangle_0 = \langle f_1\rangle_0\langle f_2\rangle_0,\tag{2.15}$$

where

$$\bigl\langle f_1(\varepsilon)\bigr\rangle_0 = \frac{1}{Z_1}\int_{-\sqrt{N}}^{\sqrt{N}} f_1(\varepsilon)\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2}\exp\left(a\varepsilon-\frac{1}{2}b\varepsilon^2\right)\mathrm{d}\varepsilon,\tag{2.16}$$

$$Z_1 = \int_{-\sqrt{N}}^{\sqrt{N}}\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2}\exp\left(a\varepsilon-\frac{1}{2}b\varepsilon^2\right)\mathrm{d}\varepsilon$$

and

$$\bigl\langle f_2(\hat{\boldsymbol{\sigma}})\bigr\rangle_0 = \frac{1}{Z_2}\int_{S_{N-1}} f_2(\hat{\boldsymbol{\sigma}})\exp\left(H_{N-1}(\hat{\boldsymbol{\sigma}}) + h\sum_{i\le N-1}\hat{\sigma}_i\right)\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}}),\tag{2.17}$$

$$Z_2 = \int_{S_{N-1}}\exp\left(H_{N-1}(\hat{\boldsymbol{\sigma}}) + h\sum_{i\le N-1}\hat{\sigma}_i\right)\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}}).$$

As we mentioned above, this decoupling is a crucial feature of the cavity method. Using (2.15), (2.16), we will be able to compute the moments $\nu_0(\varepsilon_1^{k_1}\cdots\varepsilon_n^{k_n})$ for integer $k_i \ge 0$, which is an important part of the second moment computations and of the cavity method in general. The following holds. Let us recall (2.3), (2.6) and define $\gamma_0 = 1$, $\gamma_1 = a/(b+1)$ and, recursively, for $k \ge 2$

$$\gamma_k = \frac{a}{b+1}\gamma_{k-1} + \frac{k-1}{b+1}\gamma_{k-2}.\tag{2.18}$$

Theorem 4. *For small enough $\beta > 0$,*

$$\left|\nu_0\left(\varepsilon_1^{k_1} \cdots \varepsilon_n^{k_n}\right) - \mathbb{E}\gamma_{k_1} \cdots \gamma_{k_n}\right| \le \frac{L}{N}, \tag{2.19}$$

where the constant L is independent of N.

This theorem will be proved in Section 5 below.

3. Second moment computations

We are now ready to demonstrate the promised application of the cavity interpolation, namely, the computation of the covariance matrix for the fluctuations of a certain collection of overlaps and magnetizations. Let us introduce the following seven functions

$$\begin{aligned} &f_1 = (R_{1,2} - q)^2, \qquad f_2 = (R_{1,2} - q)(R_{1,3} - q), \qquad f_3 = (R_{1,2} - q)(R_{3,4} - q), \\ &f_4 = (R_{1,2} - q)(R_1 - r), \qquad f_5 = (R_{1,2} - q)(R_3 - r), \qquad f_6 = (R_1 - r)^2, \\ &f_7 = (R_1 - r)(R_2 - r) \end{aligned} \tag{3.1}$$

and let $\boldsymbol{v}_N = (\nu(f_1), \ldots, \nu(f_7))$. In this section we will compute the vector $N\boldsymbol{v}_N$ up to the terms of order $\mathrm{o}(1)$. As we mentioned above, it is likely that with more effort one can prove the central limit theorem for the joint distribution of

$$\sqrt{N}(R_{1,2} - q), \qquad \sqrt{N}(R_{1,3} - q), \qquad \sqrt{N}(R_{3,4} - q), \qquad \sqrt{N}(R_1 - r), \qquad \sqrt{N}(R_2 - r),$$

so the computation of this section identifies the covariance matrix of the limiting Gaussian distribution. To describe our main result let us first summarize several computations based on Theorem 4. The definition (2.18) implies that

$$\gamma_1 = \frac{a}{b+1}, \qquad \gamma_2 = \left(\frac{a}{b+1}\right)^2 + \frac{1}{b+1}, \qquad \gamma_3 = \left(\frac{a}{b+1}\right)^3 + \frac{3a}{(b+1)^2}. \tag{3.2}$$

The definition (2.3) and (1.5) imply that

$$\frac{1}{b+1} = \frac{1}{1 + (1-q)(\beta^2\xi'(q) + h^2)} = 1 - q.$$

Therefore,

$$\mathbb{E}\frac{a}{b+1} = (1-q)\mathbb{E}a = (1-q)h = r, \tag{3.3}$$

$$\mathbb{E}\left(\frac{a}{b+1}\right)^2 = (1-q)^2\mathbb{E}a^2 = (1-q)^2\left(\beta^2\xi'(q) + h^2\right) = q, \tag{3.4}$$

where we used (1.5) again, and

$$W := \mathbb{E}\left(\frac{a}{b+1}\right)^3 = (1-q)^3\mathbb{E}a^3 = (1-q)^3\left(3\beta^2\xi'(q)h + h^3\right), \tag{3.5}$$

$$U := \mathbb{E}\left(\frac{a}{b+1}\right)^4 = (1-q)^4\mathbb{E}a^4 = (1-q)^4\left(h^4 + 6\beta^2h^2\xi'(q) + 3\beta^4\xi'(q)^2\right). \tag{3.6}$$

For simplicity of notations let us write

$$x \sim y \quad \text{if } x = y + \mathrm{O}\left(N^{-1}\right).$$

Then it is trivial to check that Theorem 4 and (3.2)–(3.6) imply the following relations:

$$\begin{aligned}
&\nu_0(\varepsilon_1)\sim r, \qquad \nu_0(\varepsilon_1\varepsilon_2)\sim q, \qquad \nu_0(\varepsilon_1^2)\sim 1, \qquad \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3)\sim W,\\
&\nu_0(\varepsilon_1\varepsilon_2^2)\sim W+h(1-q)^2, \qquad \nu_0(\varepsilon_1^3)\sim W+3h(1-q)^2,\\
&\nu_0(\varepsilon_1^2\varepsilon_2^2)\sim U+1-q^2, \qquad \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3^2)\sim U+q-q^2,\\
&\nu_0(\varepsilon_1\varepsilon_2^3)\sim U+3q-3q^2, \qquad \nu_0(\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4)\sim U.
\end{aligned} \tag{3.7}$$

Let us recall the definitions a_l and $a_{l,l'}$ in (2.8). Using relations (3.7) it is now straightforward to compute the following nine quantities

$$\begin{aligned}
&\nu_0\big(a_{1,2}(\varepsilon_1\varepsilon_2-q)\big)\sim Y_1, \qquad \nu_0\big(a_{1,3}(\varepsilon_1\varepsilon_2-q)\big)\sim Y_2, \qquad \nu_0\big(a_{3,4}(\varepsilon_1\varepsilon_2-q)\big)\sim Y_3,\\
&\nu_0\big(a_1(\varepsilon_1\varepsilon_2-q)\big)\sim Y_4, \qquad \nu_0\big(a_3(\varepsilon_1\varepsilon_2-q)\big)\sim Y_5, \qquad \nu_0\big(a_{1,2}(\varepsilon_1-r)\big)\sim Y_6,\\
&\nu_0\big(a_{2,3}(\varepsilon_1-r)\big)\sim Y_7, \qquad \nu_0\big(a_1(\varepsilon_1-r)\big)\sim Y_8, \qquad \nu_0\big(a_2(\varepsilon_1-r)\big)\sim Y_9,
\end{aligned} \tag{3.8}$$

where $Y_1,\dots,Y_9$ are functions of q,r,h,U,W. We omit the explicit formulas for Y_j's since they do not serve any particular purpose in the sequel. Let us define a (7×7)-matrix M that consists of four blocks

$$M=\begin{pmatrix} M_1 & O_1\\ O_2 & M_2\end{pmatrix}, \tag{3.9}$$

where O_1 is a (3×2)-matrix and O_2 is a (4×3)-matrix both entirely consisting of zeros,

$$M_1=\begin{pmatrix} 2\beta^2Y_1 & -8\beta^2Y_2 & 6\beta^2Y_3 & hY_4 & -hY_5\\ 2\beta^2Y_2 & 2\beta^2(Y_1-2Y_2-3Y_3) & 6\beta^2(-Y_2+2Y_3) & \frac{h}{2}(Y_4+Y_5) & \frac{h}{2}(Y_4-3Y_5)\\ 2\beta^2Y_3 & 8\beta^2(Y_2-2Y_3) & 2\beta^2(Y_1-8Y_2+10Y_3) & hY_5 & h(Y_4-2Y_5)\end{pmatrix},$$

$$M_2=\begin{pmatrix} 2\beta^2(Y_1-2Y_2) & 2\beta^2(-2Y_2+3Y_3) & (h/2)Y_4 & (h/2)(Y_4-2Y_5)\\ 2\beta^2(2Y_2-3Y_4) & 2\beta^2(Y_1-6Y_2+6Y_3) & (h/2)Y_5 & (h/2)(2Y_4-3Y_5)\\ -2\beta^2Y_6 & 2\beta^2Y_7 & (h/2)Y_8 & -(h/2)Y_9\\ 2\beta^2(Y_6-2Y_7) & 2\beta^2(-2Y_6+3Y_7) & (h/2)Y_9 & (h/2)(Y_8-2Y_9)\end{pmatrix}.$$

Finally, we define a vector $\boldsymbol{v}=(v_1,\dots,v_7)$ by

$$\begin{aligned}
&v_1=(1-q)U+1-4q^2+3q^3, \qquad v_2=(1-q)U+q(1-q)(1-2q),\\
&v_3=(1-q)U-q^2(1-q), \qquad v_4=W-\frac{1}{2}rU+\frac{1}{2}r\big(2-6q+3q^2\big),\\
&v_5=W-\frac{1}{2}rU+\frac{1}{2}r\big(-2q+q^2\big), \qquad v_6=-\frac{1}{2}rW+1+\frac{1}{2}r^2(-4+3q),\\
&v_7=-\frac{1}{2}rW+q+\frac{1}{2}r^2(-2+q).
\end{aligned} \tag{3.10}$$

We are now ready to formulate the main result of this section.

Theorem 5. *For small enough β and h we have*

$$(I-M)\boldsymbol{v}_N^{\mathrm{T}}=\frac{1}{N}\boldsymbol{v}^{\mathrm{T}}+\mathrm{o}\big(N^{-1}\big). \tag{3.11}$$

Here $\boldsymbol{v}^{\mathrm{T}}$ denotes the transpose of vector $\boldsymbol{v}$. In the statement of the Theorem we implicitly assumed that β and h are small enough so that $(I-M)$ is invertible. Notice that each entry in the matrix M has either a factor of β^2 or h and, therefore, for small enough β and h the matrix $(I-M)$ will indeed be invertible. Theorem 5 implies

$$\boldsymbol{v}_N^{\mathrm{T}}=\frac{1}{N}(I-M)^{-1}\boldsymbol{v}^{\mathrm{T}}+\mathrm{o}\big(N^{-1}\big).$$

In the remainder of this section we will prove Theorem 5.

For each function f_l in (3.1), we will define $\hat{f}_l$ by replacing each occurrence of R by $\hat{R}$, i.e. $\hat{f}_1 = (\hat{R}_{1,2} - q)^2$, $\hat{f}_2 = (\hat{R}_{1,2} - q)(\hat{R}_{1,3} - q)$, etc. Next, we introduce functions

$$\begin{aligned} &f_1' = (\varepsilon_1\varepsilon_2 - q)(R_{1,2} - q), \qquad f_2' = (\varepsilon_1\varepsilon_2 - q)(R_{1,3} - q), \qquad f_3' = (\varepsilon_1\varepsilon_2 - q)(R_{3,4} - q), \\ &f_4' = (\varepsilon_1\varepsilon_2 - q)(R_1 - r), \qquad f_5' = (\varepsilon_1\varepsilon_2 - q)(R_3 - r), \qquad f_6' = (\varepsilon_1 - r)(R_1 - r), \\ &f_7' = (\varepsilon_1 - r)(R_2 - r). \end{aligned} \tag{3.12}$$

As in the classical cavity method in [5], we introduce these functions because, first of all, by symmetry,

$$\nu(f_l) = \nu\big(f_l'\big) \tag{3.13}$$

and, second of all, emphasizing the last coordinate in f_l' is perfectly suited for the application of the cavity method. As above, for each function f_l' we will define $\hat{f}_l'$ by replacing each occurrence of R by $\hat{R}$, i.e. $\hat{f}_1' = (\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)$, etc.

To simplify the notations we will write $x \approx y$ whenever

$$|x - y| = \mathrm{o}\left(\frac{1}{N} + \nu_0\big((\hat{R}_{1,2} - q)^2\big) + \nu_0\big((\hat{R}_1 - r)^2\big)\right). \tag{3.14}$$

The proof of Theorem 5 will be based on the following.

Theorem 6. *For small enough β and h, for all $l \le 7$,*

$$\nu_0(\hat{f}_l) \approx \nu_0\big(f_l'\big) + \nu_0'\big(\hat{f}_l'\big). \tag{3.15}$$

We will start with a couple of lemmas.

Lemma 1. *If $f \ge 0$ and $\|f\|_\infty$ is bounded independently of N then for any $K > 0$ we can find $L > 0$ such that*

$$\nu_t(f) \le L\big(N^{-K} + \nu_0(f)\big). \tag{3.16}$$

Proof. The derivative $\nu_t'(f)$ in (2.9) consists of a finite sum of terms of the type $\nu_t(f p_\varepsilon g)$ where p_ε is some polynomial in the last coordinates (ε_l) and g is one of the following:

$$\hat{R}_{l,l'} - q, \qquad \hat{R}_l - r, \qquad (\hat{R}_{l,l'} - q)^2, \qquad N^{-1}. \tag{3.17}$$

Theorem 2 and Chebyshev's inequality imply

$$\nu_t\big(I\big(|\varepsilon_l| \ge \log N\big)\big) \le LN^{-K}$$

and combining this with Theorem 3 yields that for any g in (3.17),

$$\nu_t\big(I\big(|p_\varepsilon g| \ge N^{-1/8}\big)\big) \le LN^{-K}.$$

Therefore, one can control the derivative

$$\big|\nu_t'(f)\big| \le LN^{-K} + LN^{-1/8}\nu_t(f) \le L\big(N^{-K} + \nu_t(f)\big) \tag{3.18}$$

and (3.16) follows by integration. □

Lemma 2. *For small enough β and h and all $l \le 7$ we have*

$$\nu(f_l) \approx \nu_0(\hat{f}_l) \quad \textit{and} \quad \nu_0'\big(f_l'\big) \approx \nu_0'\big(\hat{f}_l'\big). \tag{3.19}$$

Proof. We will only consider the case $l=1,\ f_1=(R_{1,2}-q)^2$, since other cases are similar. We have

$$\begin{aligned}\big|\nu\big((R_{1,2}-q)^2\big)-\nu_0\big((R_{1,2}-q)^2\big)\big| &\le \sup_t\big(\nu_t'\big((R_{1,2}-q)^2\big)\big)\\ &\le \sup_t\big(LN^{-K}+LN^{-1/8}\nu_t\big((R_{1,2}-q)^2\big)\big)\\ &\le \big(LN^{-K}+LN^{-1/8}\nu_0\big((R_{1,2}-q)^2\big)\big), \end{aligned}\tag{3.20}$$

where in the second line we used (3.18) and then (3.16). We will use that

$$R_{1,2}=\hat{R}_{1,2}+s(\varepsilon_1,\varepsilon_2)\hat{R}_{1,2}+N^{-1}\varepsilon_1\varepsilon_2,\tag{3.21}$$

where

$$s(\varepsilon_1,\varepsilon_2)=\sqrt{\left(1-\frac{\varepsilon_1^2}{N}\right)\left(1-\frac{\varepsilon_2^2}{N}\right)}-1.$$

Since

$$\left|\sqrt{1+x}-1-\frac{x}{2}\right|\le Lx^2\quad\text{for } x\in[-1,1]\tag{3.22}$$

we have

$$\left|s(\varepsilon_1,\varepsilon_2)+\frac{1}{2N}\big(\varepsilon_1^2+\varepsilon_2^2\big)\right|\le\frac{L}{N^2}\big(\varepsilon_1^4+\varepsilon_2^4\big).\tag{3.23}$$

Since by (3.21)

$$R_{1,2}-q=(\hat{R}_{1,2}-q)+\left(\sqrt{\left(1-\frac{\varepsilon_1^2}{N}\right)\left(1-\frac{\varepsilon_2^2}{N}\right)}-1\right)\hat{R}_{1,2}+\frac{1}{N}\varepsilon_1\varepsilon_2,\tag{3.24}$$

squaring both sides and using (3.23) yields

$$\big|(R_{1,2}-q)^2-(\hat{R}_{1,2}-q)^2\big|\le\frac{1}{N}p_\varepsilon|\hat{R}_{1,2}-q|+\frac{1}{N^2}p_\varepsilon,$$

where from now on p_ε denotes a quantity such that

$$|p_\varepsilon|\le L\left(1+\sum_l\varepsilon_l^4\right).$$

Therefore,

$$\big|\nu_0\big((R_{1,2}-q)^2\big)-\nu_0\big((\hat{R}_{1,2}-q)^2\big)\big|\le\frac{1}{N}\nu_0\big(p_\varepsilon|\hat{R}_{1,2}-q|\big)+\frac{1}{N^2}\nu_0(p_\varepsilon)=\mathrm{o}\big(N^{-1}\big)$$

by Theorems 2 and 3. Thus, (3.20) implies the first part of (3.19). To prove the second part of (3.19) we notice that

$$\big|f_1'-\hat{f}_1'\big|=\big|(\varepsilon_1\varepsilon_2-q)(R_{1,2}-\hat{R}_{1,2})\big|\le\frac{1}{N}p_\varepsilon$$

by (3.21) and (3.23). Since each term in the derivatives $\nu_0'(f_1')$ and $\nu_0'(\hat{f}_1')$ will contain another factor from the list (3.17), Theorems 2 and 3 imply the result. □

Proof of Theorem 6. We start by writing

$$\left|\nu\left(f_l'\right)-\nu_0\left(f_l'\right)-\nu_0'\left(f_l'\right)\right|\leq\sup_t\left|\nu_t''\left(f_l'\right)\right|.$$

If we can show that

$$\sup_t\left|\nu_t''\left(f_l'\right)\right|\approx 0 \tag{3.25}$$

and, thus, $\nu(f_l')\approx\nu_0(f_l')+\nu_0'(f_l')$, then Lemma 2 and (3.13) will imply

$$\nu_0(\hat{f}_l)\approx\nu(f_l)=\nu\left(f_l'\right)\approx\nu_0\left(f_l'\right)+\nu_0'\left(f_l'\right)\approx\nu_0\left(f_l'\right)+\nu_0'\left(\hat{f}_l'\right),$$

which is precisely the statement of Theorem 6. To prove (3.25) we note that by (2.9) the second derivative $\nu_t''(f_l')$ will consist of the finite sum of terms of the type $f_l' p_\varepsilon g_1 g_2$ where g_1, g_2 are from the list (3.17). Clearly,

$$|g_1 g_2|\leq L\left(\frac{1}{N^2}+(\hat{R}_{l,l'}-q)^2+(\hat{R}_{l''}-r)^2\right)$$

and since each f_l' contain another small factor $(R_{l,l'}-q)$ or (R_l-r), Theorems 2 and 3 imply (3.25). □

We are now ready to prove Theorem 5.

Proof of Theorem 5. Let us first note that $\nu_0(\hat{f}_l)$ is defined exactly the same way as $\nu(f_l)$ for $N-1$ instead of N. In other words,

$$\boldsymbol{v}_N^0:=\left(\nu_0(\hat{f}_1),\ldots,\nu_0(\hat{f}_7)\right)=\boldsymbol{v}_{N-1}$$

and, therefore, it is enough to prove that

$$(I-M)\boldsymbol{v}_N^{0\mathrm{T}}=\frac{1}{N}\boldsymbol{v}^{\mathrm{T}}+\mathrm{o}\left(N^{-1}\right). \tag{3.26}$$

Replacing $1/N$ by $1/(N-1)$ on the right hand side is not necessary since the difference is of order N^{-2}. Each equation in the system of Eqs (3.26) will follow from the corresponding equation (3.15). Namely, we will show that

$$\left(\nu_0\left(f_1'\right),\ldots,\nu_0\left(f_7'\right)\right)\approx\frac{1}{N}\boldsymbol{v}\quad\text{and}\quad\left(\nu_0'\left(\hat{f}_1'\right),\ldots,\nu_0'\left(\hat{f}_7'\right)\right)^{\mathrm{T}}\approx M\boldsymbol{v}_N^{0\mathrm{T}}. \tag{3.27}$$

Then (3.15) will imply that $\boldsymbol{v}_N^{0\mathrm{T}}\approx N^{-1}\boldsymbol{v}^{\mathrm{T}}+M\boldsymbol{v}_N^{0\mathrm{T}}$. However, since the definition (3.14) means that the error in each equation is of order $\mathrm{o}(N^{-1}+\nu_0(\hat{f}_1)+\nu_0(\hat{f}_6))$, this system of equation can be rewritten as

$$(I-M-\mathcal{E}_N)\boldsymbol{v}_N^{0\mathrm{T}}=\frac{1}{N}\boldsymbol{v}+\mathrm{o}\left(N^{-1}\right),$$

where the matrix $\mathcal{E}_N$ is such that $\|\mathcal{E}_N\|=\mathrm{o}(1)$. Therefore, whenever the matrix $I-M$ is invertible (for example, for small β and h) we have for N large enough

$$\begin{aligned}\boldsymbol{v}_N^{0\mathrm{T}}&=(I-M-\mathcal{E}_N)^{-1}\left(\frac{1}{N}\boldsymbol{v}^{\mathrm{T}}+\mathrm{o}\left(N^{-1}\right)\right)\\&=\frac{1}{N}(I-M)^{-1}\boldsymbol{v}^{\mathrm{T}}+\mathrm{o}\left(N^{-1}\right).\end{aligned}$$

Hence, to finish the proof we need only to show (3.27). We will only carry out the computations for $l = 1$ since all other cases are similar. Let us start by proving that $\nu_0((\varepsilon_1\varepsilon_2 - q)(R_{1,2} - q)) \approx v_1$. Using (3.24) and (2.15), we write

$$\begin{aligned}\nu_0\big((\varepsilon_1\varepsilon_2 - q)(R_{1,2} - q)\big) &= \frac{1}{N}\nu_0\big(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)\big) + \nu_0(\varepsilon_1\varepsilon_2 - q)\nu_0(\hat{R}_{1,2} - q)\\ &\quad + q\nu_0\left((\varepsilon_1\varepsilon_2 - q)\left(\sqrt{\left(1 - \frac{\varepsilon_1^2}{N}\right)\left(1 - \frac{\varepsilon_2^2}{N}\right)} - 1\right)\right)\\ &\quad + \nu_0\left((\varepsilon_1\varepsilon_2 - q)\left(\sqrt{\left(1 - \frac{\varepsilon_1^2}{N}\right)\left(1 - \frac{\varepsilon_2^2}{N}\right)} - 1\right)\right)\nu_0(\hat{R}_{1,2} - q).\end{aligned}$$

Using (3.23), one can bound the last term by

$$\frac{1}{N}\nu_0(p_\varepsilon)\big|\nu_0(\hat{R}_{1,2} - q)\big| = \mathrm{o}\big(N^{-1}\big)$$

by Theorems 2 and 3. The term

$$\nu_0\big((\varepsilon_1\varepsilon_2 - q)\big)\nu_0\big((\hat{R}_{1,2} - q)\big) = \mathrm{o}\big(N^{-1}\big)$$

by Theorem 3 and the second relation in (3.7), i.e. $\nu_0(\varepsilon_1\varepsilon_2 - q) \sim 0$. Finally, we use

$$\left|\sqrt{\left(1 - \frac{\varepsilon_1^2}{N}\right)\left(1 - \frac{\varepsilon_2^2}{N}\right)} - 1 + \frac{\varepsilon_1^2}{2N} + \frac{\varepsilon_2^2}{2N}\right| \le \frac{1}{N^2}p_\varepsilon$$

to observe that

$$q\nu_0\left((\varepsilon_1\varepsilon_2 - q)\left(\sqrt{\left(1 - \frac{\varepsilon_1^2}{N}\right)\left(1 - \frac{\varepsilon_2^2}{N}\right)} - 1\right)\right) \approx -\frac{1}{2N}q\nu_0\big((\varepsilon_1\varepsilon_2 - q)\big(\varepsilon_1^2 + \varepsilon_2^2\big)\big) = -\frac{1}{N}q\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_1^2\big)$$

by symmetry and, therefore,

$$\begin{aligned}\nu_0\big((\varepsilon_1\varepsilon_2 - q)(R_{1,2} - q)\big) &\approx \frac{1}{N}\big(\nu_0\big(\varepsilon_1\varepsilon_2(\varepsilon_1\varepsilon_2 - q)\big) - q\nu_0\big((\varepsilon_1\varepsilon_2 - q)\varepsilon_1^2\big)\big)\\ &= \frac{1}{N}\big(\nu_0\big(\varepsilon_1^2\varepsilon_2^2\big) - q\nu_0(\varepsilon_1\varepsilon_2) - q\nu_0\big(\varepsilon_1^3\varepsilon_2\big) + q^2\nu_0\big(\varepsilon_1^2\big)\big) \approx v_1\end{aligned}$$

by using (3.7) and comparing with the definition of v_1 in (3.10).

Next, we need to show the second part of (3.27) for $l = 1$, i.e.

$$\nu_0'\big((\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)\big) \approx \big(M\boldsymbol{v}_N^{0\mathrm{T}}\big)_1,$$

where $(\cdot)_1$ denotes the first coordinate of a vector. We use (2.9) for $n = 2$ to write $\nu_0'((\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q))$ as

$$\begin{aligned}&h\nu_0\big(a_1(\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)(\hat{R}_1 - r)\big) - h\nu_0\big(a_3(\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)(\hat{R}_3 - r)\big)\\ &\quad + 2\beta^2h\nu_0\big(a_{1,2}(\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)^2\big) - 8\beta^2h\nu_0\big(a_{1,3}(\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)(\hat{R}_{1,3} - q)\big)\\ &\quad + 6\beta^2h\nu_0\big(a_{3,4}(\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)(\hat{R}_{3,4} - q)\big) + \nu_0\big((\varepsilon_1\varepsilon_2 - q)(\hat{R}_{1,2} - q)\mathcal{R}\big)\\ &\approx 2\beta^2Y_1\nu_0(\hat{f}_1) - 8\beta^2Y_2\nu_0(\hat{f}_2) + 6\beta^2Y_3\nu_0(\hat{f}_3) + hY_4\nu_0(\hat{f}_4) - hY_5\nu_0(\hat{f}_5) + \nu_0\big(\hat{f}_1'\mathcal{R}\big)\\ &= \big(M\boldsymbol{v}_N^{0\mathrm{T}}\big)_1 + \nu_0\big(\hat{f}_1'\mathcal{R}\big),\end{aligned}$$

where in second to last line we used (3.8) and the last line follows by comparison with the definition of M in (3.9). Finally, since clearly $\nu_0(\hat{f}_1'\mathcal{R}) \approx 0$ by Theorems 2 and 3, this finishes the proof of Theorem 5. □

4. Derivative along the interpolation

In this section we prove Theorem 1. We start by writing

$$\nu_t'(f) = \mathbb{E}\left\langle f \sum_{l\le n} \frac{\partial}{\partial t} H_t\big(\boldsymbol{\sigma}^l\big)\right\rangle_t - n\mathbb{E}\left\langle f \frac{\partial}{\partial t} H_t(\boldsymbol{\sigma}^{n+1})\right\rangle_t \tag{4.1}$$

and

$$\begin{aligned}\frac{\partial}{\partial t} H_t(\boldsymbol{\sigma}) = {} & \frac{\beta}{2\sqrt{t}} H_N(\boldsymbol{\sigma}) - \frac{\beta}{2\sqrt{1-t}} H_{N-1}(\hat{\boldsymbol{\sigma}}) + h \sum_{i\le N-1} \hat{\sigma}_i \left(\sqrt{\frac{N-\varepsilon^2}{N-1}} - 1\right) \\ & - \frac{1}{2\sqrt{1-t}} \varepsilon z \beta \sqrt{\xi'(q)} + \frac{1}{2}\varepsilon^2 b.\end{aligned} \tag{4.2}$$

In order to use a Gaussian integration by parts (see, for example, (A.41) in [5]) we first compute the covariance

$$\operatorname{Cov}\left(H_t\big(\boldsymbol{\sigma}^1\big), \frac{\partial}{\partial t} H_t\big(\boldsymbol{\sigma}^2\big)\right) = \frac{\beta^2}{2}\big(N\xi(R_{1,2}) - (N-1)\xi(\hat{R}_{1,2}) - \varepsilon_1\varepsilon_2\xi'(q)\big)$$

by (1.1) and (2.1). We will rewrite this using Taylor's expansion of $\xi(R_{1,2})$ near $\hat{R}_{1,2}$. By assumption, ξ is three times continuously differentiable and (3.21), (3.23) imply

$$\big|\xi(R_{1,2}) - \xi(\hat{R}_{1,2}) - \xi'(\hat{R}_{1,2})(R_{1,2} - \hat{R}_{1,2})\big| \le \frac{L}{N^2}\big(\varepsilon_1^4 + \varepsilon_2^4\big)$$

and

$$\left|\xi(R_{1,2}) - \xi(\hat{R}_{1,2}) + \frac{1}{2N}\big(\varepsilon_1^2 + \varepsilon_2^2\big)\hat{R}_{1,2}\xi'(\hat{R}_{1,2}) - \frac{1}{N}\varepsilon_1\varepsilon_2\xi'(\hat{R}_{1,2})\right| \le \frac{L}{N^2}\big(\varepsilon_1^4 + \varepsilon_2^4\big).$$

Therefore,

$$N\xi(R_{1,2}) - (N-1)\xi(\hat{R}_{1,2}) = \xi(\hat{R}_{1,2}) - \frac{1}{2}\big(\varepsilon_1^2 + \varepsilon_2^2\big)\hat{R}_{1,2}\xi'(\hat{R}_{1,2}) + \varepsilon_1\varepsilon_2\xi'(\hat{R}_{1,2}) + \mathcal{R}_1, \tag{4.3}$$

where from now on $\mathcal{R}_1$ will denote a quantity such that

$$|\mathcal{R}_1| \le \frac{L}{N}\left(1 + \sum_{l\le n+2} \varepsilon_l^4\right).$$

Since ξ is three times continuously differentiable,

$$\begin{aligned}&\xi(\hat{R}_{1,2}) - \xi(q) = \xi'(q)(\hat{R}_{1,2} - q) + \mathcal{R}_2, \qquad \xi'(\hat{R}_{1,2}) - \xi'(q) = \xi''(q)(\hat{R}_{1,2} - q) + \mathcal{R}_2,\\ &\hat{R}_{1,2}\xi'(\hat{R}_{1,2}) - q\xi'(q) = \big(\xi'(q) + q\xi''(q)\big)(\hat{R}_{1,2} - q) + \mathcal{R}_2,\end{aligned}$$

where $\mathcal{R}_2$ denotes a quantity such that

$$|\mathcal{R}_2| \le L(\hat{R}_{1,2} - q)^2.$$

Using this in (4.3) and recalling the definition of $a_{l,l'}$ in (2.8) we get

$$\operatorname{Cov}\left(H_t\big(\boldsymbol{\sigma}^l\big), \frac{\partial}{\partial t} H_t\big(\boldsymbol{\sigma}^{l'}\big)\right) = \frac{\beta^2}{2}\left(2a_{l,l'}(\hat{R}_{l,l'} - q) - \frac{1}{2}\big(\varepsilon_l^2 + \varepsilon_{l'}^2\big)q\xi'(q) + \xi(q)\right) + \beta^2\mathcal{R}_3, \tag{4.4}$$

where

$$|\mathcal{R}_3| \le \frac{L}{N}\Big(1+\sum_{l\le n+2}\varepsilon_l^4\Big)+L\sum_{l\ne l'\le n+2}\big(1+\varepsilon_l^2\big)(\hat{R}_{l,l'}-q)^2.$$

On the other hand, when $l=l'$ we get directly

$$\mathrm{Cov}\Big(H_t\big(\boldsymbol{\sigma}^l\big),\frac{\partial}{\partial t}H_t\big(\boldsymbol{\sigma}^l\big)\Big)=\frac{\beta^2}{2}\big(\xi(1)-\varepsilon_l^2\xi'(q)\big). \tag{4.5}$$

Next, we simplify the third term on the right hand side of (4.2). Equation (3.22) implies

$$\left|\sqrt{\frac{N-\varepsilon^2}{N-1}}-\Big(1+\frac{1-\varepsilon^2}{2(N-1)}\Big)\right|\le L\frac{(1-\varepsilon^2)^2}{(N-1)^2}$$

and, therefore,

$$(N-1)\Big(\sqrt{\frac{N-\varepsilon^2}{N-1}}-1\Big)-\frac{1-\varepsilon^2}{2}=\mathcal{R}_1.$$

We can write

$$\begin{aligned} h\sum_{i\le N-1}\hat{\sigma}_i^l\Big(\sqrt{\frac{N-\varepsilon_l^2}{N-1}}-1\Big) &= \frac{h}{2}\hat{R}_l\big(1-\varepsilon_l^2\big)+h\mathcal{R}_1 \\ &= \frac{h}{2}a_l(\hat{R}_l-r)+\frac{hr}{2}\big(1-\varepsilon_l^2\big)+h\mathcal{R}_1, \end{aligned} \tag{4.6}$$

where in the last line we used the definition of a_l in (2.8). Finally, using (4.4)–(4.6), Gaussian integration by parts in (4.1) gives

$$\nu_t'(f)=\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}+\mathrm{V}+\mathrm{VI}+\nu_t(f\mathcal{R}),$$

where I is created by the first term in (4.6):

$$\mathrm{I}=\frac{h}{2}\sum_{l\le n}\nu_t\big(fa_l(\hat{R}_l-r)\big)-n\frac{h}{2}\nu_t\big(fa_{n+1}(\hat{R}_{n+1}-r)\big),$$

II is created by the first term in (4.4):

$$\begin{aligned} \mathrm{II}=\beta^2&\sum_{1\le l\ne l'\le n}\nu_t\big(fa_{l,l'}(\hat{R}_{l,l'}-q)\big)-2n\beta^2\sum_{l\le n}\nu_t\big(fa_{l,n+1}(\hat{R}_{l,n+1}-q)\big) \\ &+n(n+1)\beta^2\nu_t\big(fa_{n+1,n+2}(\hat{R}_{n+1,n+2}-q)\big), \end{aligned}$$

III is created by the second term in (4.6):

$$\mathrm{III}=-\frac{hr}{2}\Big(\sum_{l\le n}\nu_t\big(f\varepsilon_l^2\big)-n\nu_t\big(f\varepsilon_{n+1}^2\big)\Big),$$

IV is created by the second term in (4.4):

$$\mathrm{IV}=-\frac{\beta^2}{4}q\xi'(q)\Big(\sum_{1\le l\ne l'\le n}\nu_t\big(f\big(\varepsilon_l^2+\varepsilon_{l'}^2\big)\big)-2n\sum_{l\le n}\nu_t\big(f\big(\varepsilon_l^2+\varepsilon_{n+1}^2\big)\big)+n(n+1)\nu_t\big(f\big(\varepsilon_{n+1}^2+\varepsilon_{n+2}^2\big)\big)\Big),$$

V is created by (4.5):

$$\mathrm{V} = -\frac{\beta^2}{2}\xi'(q)\biggl(\sum_{l\le n}\nu_t\bigl(f\varepsilon_l^2\bigr) - n\nu_t\bigl(f\varepsilon_{n+1}^2\bigr)\biggr)$$

and VI is created by the last term in (4.2):

$$\mathrm{VI} = \frac{1}{2}b\biggl(\sum_{l\le n}\nu_t\bigl(f\varepsilon_l^2\bigr) - n\nu_t\bigl(f\varepsilon_{n+1}^2\bigr)\biggr).$$

Using that by symmetry, $\nu_t(f\varepsilon_{n+1}^2) = \nu_t(f\varepsilon_{n+2}^2)$, and counting terms in IV it is easy to see that

$$\mathrm{IV} = \frac{\beta^2}{2}q\xi'(q)\biggl(\sum_{l\le n}\nu_t\bigl(f\varepsilon_l^2\bigr) - n\nu_t\bigl(f\varepsilon_{n+1}^2\bigr)\biggr).$$

Since, by definition, $b = hr + \beta^2(1-q)\xi'(q)$, we have $\mathrm{III}+\mathrm{IV}+\mathrm{V}+\mathrm{VI} = 0$. This finishes the proof of Theorem 1.

5. Control of the last coordinate

In this section we will prove Theorems 2 and 4. We start with the following.

Lemma 3. *If $c_0 < 1$ then for β small enough*,

$$\nu_0\bigl(\exp c_0\varepsilon^2\bigr) \le L.$$

Proof. By (2.15) and using $1 - x \le \exp(-x)$,

$$\begin{aligned}\bigl\langle\exp c_0\varepsilon^2\bigr\rangle_0 &= \frac{1}{Z_1}\int_{-\sqrt{N}}^{\sqrt{N}}\biggl(1-\frac{\varepsilon^2}{N}\biggr)^{(N-3)/2}\exp\biggl(a\varepsilon - \frac{1}{2}(b-c_0)\varepsilon^2\biggr)\,\mathrm{d}\varepsilon,\\ &\le \frac{1}{Z_1}\int_{-\sqrt{N}}^{\sqrt{N}}\exp\biggl(a\varepsilon - \frac{1}{2}\bigl(b-c_0+1-3N^{-1}\bigr)\varepsilon^2\biggr)\,\mathrm{d}\varepsilon \le \frac{1}{Z_1}L\exp\bigl(La^2\bigr)\end{aligned} \tag{5.1}$$

since for $c_0 < 1$ we have $b + 1 - 3N^{-1} - c_0 > 0$ for large enough N. On the other hand, one can show that

$$Z_1 \ge \frac{1}{L}\exp\bigl(-La^2\bigr). \tag{5.2}$$

Indeed, using that $1 - x \ge \exp(-Lx)$ for $x \le 1/2$,

$$\begin{aligned}Z_1 &= \int_{-\sqrt{N}}^{\sqrt{N}}\biggl(1-\frac{\varepsilon^2}{N}\biggr)^{(N-3)/2}\exp\biggl(a\varepsilon-\frac{1}{2}b\varepsilon^2\biggr)\,\mathrm{d}\varepsilon \ge \int_{-\sqrt{N}/2}^{\sqrt{N}/2}\exp\biggl(a\varepsilon-\frac{1}{2}L\varepsilon^2\biggr)\,\mathrm{d}\varepsilon\\ &= \frac{1}{\sqrt{L}}\exp\biggl(\frac{a^2}{2L}\biggr)\int_{-\sqrt{LN}/2-a}^{\sqrt{LN}/2-a}\exp\biggl(-\frac{x^2}{2}\biggr)\,\mathrm{d}x \ge \frac{1}{L}\int_{-L\sqrt{N}-a}^{L\sqrt{N}-a}\exp\biggl(-\frac{x^2}{2}\biggr)\,\mathrm{d}x.\end{aligned}$$

When $|a| \le L\sqrt{N}+1$, this implies that $Z_1 \ge 1/L$. Otherwise, say, when $a \ge L\sqrt{N}+1$, we can use the well known estimates for the Gaussian tail to write

$$\begin{aligned}\int_{-L\sqrt{N}-a}^{L\sqrt{N}-a}\exp\biggl(-\frac{x^2}{2}\biggr)\,\mathrm{d}x &\ge \frac{1}{L(a-L\sqrt{N})}\exp\biggl(-\frac{1}{2}\bigl(a-L\sqrt{N}\bigr)^2\biggr)\\ &\quad - L\exp\biggl(-\frac{1}{2}\bigl(a+L\sqrt{N}\bigr)^2\biggr) \ge \frac{1}{L}\exp\bigl(-La^2\bigr)\end{aligned}$$

which proves (5.2). Finally, (5.1) and (5.2) imply that

$$\nu_0\big(\exp c_0\varepsilon^2\big) \le L\mathbb{E}\exp\big(La^2\big) = L\mathbb{E}\exp\big(L\big(z\beta\sqrt{\xi'(q)}+h\big)^2\big) \le L,$$

if β is small enough, $L\beta^2\xi'(q) < 1/2$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Let us apply (2.9) to $f = \varepsilon^{2k}$ for integer $k \ge 1$. Since factors a_l and $a_{l,l'}$ are second degree polynomials in the last coordinates and $|\hat{R}_{l,l'} - q| \le L$, $|\hat{R}_l - r| \le L$ we can bound the derivative by

$$\big|\nu_t'\big(\varepsilon^{2k}\big)\big| \le L\big(\beta^2+h\big)\nu_t\big(\big(1+\varepsilon_1^2+\varepsilon_2^2\big)\varepsilon_1^{2k}\big) + \nu_t\big(\varepsilon^{2k}|\mathcal{R}|\big) \le L\big(\beta^2+h\big)\nu_t\big(\big(1+\varepsilon^2\big)\varepsilon^{2k}\big) + \nu_t\big(\varepsilon^{2k}|\mathcal{R}|\big).$$

Since $\varepsilon_l^2 \le N$, for a polynomial $p(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ of the fourth degree we have

$$\frac{1}{N}p(\varepsilon_1,\varepsilon_2,\varepsilon_3) \le \frac{L}{N}\sum_{l\le 3}\big(1+\varepsilon_l^4\big) \le L\bigg(1+\sum_{l\le 3}\varepsilon_l^2\bigg).$$

Therefore,

$$|\mathcal{R}| \le L\big(\beta^2+h\big)L\bigg(1+\sum_{l\le 3}\varepsilon_l^2\bigg)$$

and

$$\big|\nu_t'\big(\varepsilon^{2k}\big)\big| \le L\big(\beta^2+h\big)\big(\nu_t\big(\big(1+\varepsilon^2\big)\varepsilon^{2k}\big)\big).$$

Using this, we can write

$$\begin{aligned}\nu_t'\big(\exp c\varepsilon^2\big) &= \sum_{k\ge 1}\frac{c^k}{k!}\nu_t'\big(\varepsilon^{2k}\big) \le L\big(\beta^2+h\big)\sum_{k\ge 1}\frac{c^k}{k!}\nu_t\big(\big(1+\varepsilon^2\big)\varepsilon^{2k}\big)\\ &\le L\big(\beta^2+h\big)\nu_t\big(\big(1+\varepsilon^2\big)\exp c\varepsilon^2\big).\end{aligned}$$

If we take $c_0 < 1$ and let $c(t) = (c_0 - L(\beta^2+h)t)$ then

$$\begin{aligned}\nu_t'\big(\exp c(t)\varepsilon^2\big) &\le L\big(\beta^2+h\big)\nu_t\big(\big(1+\varepsilon^2\big)\exp c(t)\varepsilon^2\big) - L\big(\beta^2+h\big)\nu_t\big(\varepsilon^2\exp c(t)\varepsilon^2\big)\\ &= L\big(\beta^2+h\big)\nu_t\big(\exp c(t)\varepsilon^2\big).\end{aligned}$$

Integrating this over t yields

$$\nu_t\big(\exp c(t)\varepsilon^2\big) \le \exp\big(L\big(\beta^2\big)t\big)\nu_0\big(\exp c_0\varepsilon^2\big) \le L \tag{5.3}$$

for small enough β, by Lemma 3. If $\beta^2 + h$ is small enough then $c(t) > c_0/2$ and this finishes the proof of Theorem 2. □

Proof of Theorem 4. Let us denote

$$f(\varepsilon) = \bigg(1-\frac{\varepsilon^2}{N}\bigg)^{(N-3)/2}\exp\bigg(a\varepsilon - \frac{1}{2}b\varepsilon^2\bigg).$$

Then, using (2.15) as in (5.1), we can write

$$Z_1\langle\varepsilon^k\rangle_0 = \int_{-\sqrt{N}}^{\sqrt{N}} \varepsilon^k f(\varepsilon)\,\mathrm{d}\varepsilon = -\frac{1}{b}\int_{-\sqrt{N}}^{\sqrt{N}} \varepsilon^{k-1}\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2} \exp(a\varepsilon)\,\mathrm{d}\exp\left(-\frac{1}{2}b\varepsilon^2\right)$$
$$= \frac{1}{b}\int_{-\sqrt{N}}^{\sqrt{N}}\big((k-1)\varepsilon^{k-2}+a\varepsilon^{k-1}\big) f(\varepsilon)\,\mathrm{d}\varepsilon - \frac{1}{b}\frac{N-3}{N}\int_{-\sqrt{N}}^{\sqrt{N}} \varepsilon^k\left(1-\frac{\varepsilon^2}{N}\right)^{-1} f(\varepsilon)\,\mathrm{d}\varepsilon$$

by integration by parts. Moving the last integral to the left hand side of the equation,

$$\int_{-\sqrt{N}}^{\sqrt{N}}\left(1+\frac{1}{b}\frac{N-3}{N-\varepsilon^2}\right)\varepsilon^k f(\varepsilon)\,\mathrm{d}\varepsilon = \frac{1}{b}\int_{-\sqrt{N}}^{\sqrt{N}}\big((k-1)\varepsilon^{k-2}+a\varepsilon^{k-1}\big) f(\varepsilon)\,\mathrm{d}\varepsilon. \tag{5.4}$$

If we rewrite

$$1+\frac{1}{b}\frac{N-3}{N-\varepsilon^2} = \frac{b+1}{b}\left(1+\frac{\varepsilon^2-3}{(b+1)(N-\varepsilon^2)}\right)$$

then (5.4) implies

$$\int_{-\sqrt{N}}^{\sqrt{N}} \varepsilon^k f(\varepsilon)\,\mathrm{d}\varepsilon = \frac{1}{b+1}\int_{-\sqrt{N}}^{\sqrt{N}}\big((k-1)\varepsilon^{k-2}+a\varepsilon^{k-1}\big) f(\varepsilon)\,\mathrm{d}\varepsilon$$
$$+ \frac{1}{N(b+1)}\int_{\sqrt{N}}^{\sqrt{N}} \varepsilon^k\big(3-\varepsilon^2\big)\left(1-\frac{\varepsilon^2}{N}\right)^{-1} f(\varepsilon)\,\mathrm{d}\varepsilon.$$

Dividing both sides by Z_1 gives

$$S_k = \frac{a}{b+1}S_{k-1} + \frac{k-1}{b+1}S_{k-2} + r_k, \tag{5.5}$$

where we denoted $S_k = \langle\varepsilon^k\rangle_0$, and where

$$r_k = \frac{1}{N(b+1)}\left\langle \varepsilon^k\big(3-\varepsilon^2\big)\left(1-\frac{\varepsilon^2}{N}\right)^{-1}\right\rangle_0.$$

Comparing (5.5) with (2.18), it should be obvious that $S_k = \gamma_k + \hat{r}_k$, where $\hat{r}_k$ is a polynomial in a and $(r_l)_{l\le k}$ where each term contains a least one factor r_l. Therefore,

$$S_{k_1}\cdots S_{k_n} = \gamma_{k_1}\cdots\gamma_{k_n} + r,$$

where r is a polynomial in a and $(r_l)_{l\le k_0}$ for $k_0 = \max(k_1,\ldots,k_n)$ and each term contains at least one factor r_l. Therefore, each term in r will have at least one factor $1/N$ and if we can show that for any $k, m > 0$

$$\mathbb{E}\left\langle\left(\varepsilon^k\big(3-\varepsilon^2\big)\left(1-\frac{\varepsilon^2}{N}\right)^{-1}\right)^m\right\rangle_0 \le L \tag{5.6}$$

then, by Hölder's inequality, $\mathbb{E}|r| \le L/N$ and this finishes the proof of Theorem 4. To prove (5.6), we write that for any polynomial $p(\varepsilon)$, by (2.15),

$$\mathbb{E}\left\langle p(\varepsilon)\left(1-\frac{\varepsilon^2}{N}\right)^{-m}\right\rangle_0 = \mathbb{E}\frac{1}{Z_1}\int_{-\sqrt{N}}^{\sqrt{N}} p(\varepsilon)\left(1-\frac{\varepsilon^2}{N}\right)^{(N-3-2m)/2}\exp\left(a\varepsilon-\frac{1}{2}b\varepsilon^2\right)\mathrm{d}\varepsilon.$$

Repeating the argument of Lemma 3 one can show that for small enough $\beta > 0$ the right-hand side is bounded by some $L > 0$ which proves (5.6). $\square$

6. Control of the overlap and magnetization

In this section we will prove Theorem 3 using, among other things, an argument of R. Latala described in detail in [7]. The proof will also clarify the origin of the definitions of parameters q, r in (1.5) and (2.7). We will start with some auxiliary results. Given a set $A \subseteq S_{N-1}^n$, let us denote

$$I_A = I\big((\hat{\boldsymbol{\sigma}}^1, \dots, \hat{\boldsymbol{\sigma}}^n) \in A\big).$$

Then the following lemma holds.

Lemma 4. *If $A \subseteq S_{N-1}^n$ is symmetric with respect to permutations of the coordinates, then for small enough β and h,*

$$\left|\frac{1}{N}\mathbb{E}\log\langle I_A\rangle_t - \frac{1}{N}\mathbb{E}\log\langle I_A\rangle_0\right| \le \frac{L}{N}. \tag{6.1}$$

We will apply (6.1) to sets A of the type

$$\big\{\hat{\boldsymbol{\sigma}}^1\colon |\hat{R}_1 - r| \ge x\big\} \quad \text{or} \quad \big\{(\hat{\boldsymbol{\sigma}}^1, \hat{\boldsymbol{\sigma}}^2)\colon |\hat{R}_{1,2} - q| \ge x\big\} \tag{6.2}$$

to say that their Gibbs' measure does not change much along the interpolation (2.2).

Proof of Lemma 4. For a set $A \subseteq S_{N-1}^n$, let us consider

$$\phi_A(t) = \frac{1}{N}\mathbb{E}\log\int_A \exp\sum_{l\le n} H_t\big(\boldsymbol{\sigma}^l\big)\,\mathrm{d}\lambda_N^n.$$

Then

$$\frac{1}{N}\mathbb{E}\log\langle I_A\rangle_t = \phi_A(t) - \phi_{S_{N-1}^n}(t)$$

and Lemma 4 follows from the following. □

Lemma 5. *For small enough β and h we have*

$$\big|\phi_A'(t)\big| \le \frac{L}{N}. \tag{6.3}$$

Proof. Given a function $f = f(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n)$, we define

$$\langle f\rangle_{t,A} = \frac{\langle f I_A\rangle_t}{\langle I_A\rangle_t} = \int_A f\exp\sum_{l\le n} H_t\big(\boldsymbol{\sigma}^l\big)\,\mathrm{d}\lambda_N^n \Big/ \int_A \exp\sum_{l\le n} H_t\big(\boldsymbol{\sigma}^l\big)\,\mathrm{d}\lambda_N^n. \tag{6.4}$$

Then

$$N\phi_A'(t) = \mathbb{E}\bigg\langle\sum_{l\le n}\frac{\partial H_t(\boldsymbol{\sigma}^l)}{\partial t}\bigg\rangle_{t,A}.$$

If we denote

$$S_{l,l'} = N\xi(R_{l,l'}) - (N-1)\xi(\hat{R}_{l,l'}) - \varepsilon_l\varepsilon_{l'}\xi'(q),$$

then integration by parts as in Theorem 1 gives

$$N\phi_A'(t) = \sum_{l\le n} \mathbb{E}\bigg\langle h \sum_{i\le N-1} \hat{\sigma}_i^l \bigg(\sqrt{\frac{N-\varepsilon_l^2}{N-1}} - 1\bigg) + \frac{1}{2}\varepsilon_l^2 b\bigg\rangle_{t,A} + \frac{\beta^2}{2}\sum_{l,l'\le n} \mathbb{E}\langle S_{l,l'}\rangle_{t,A} - \frac{\beta^2}{2}\sum_{l\le n}\sum_{l'=n+1}^{2n} \mathbb{E}\langle S_{l,l'}\rangle_{t,A}. \tag{6.5}$$

The Gibbs average in the last term is defined on two copies $(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n)$ and $(\boldsymbol{\sigma}^{n+1},\dots,\boldsymbol{\sigma}^{2n})$. Since

$$\bigg|(N-1)\bigg(\sqrt{\frac{N-\varepsilon_l^2}{N-1}} - 1\bigg)\bigg| \le L\big(1+\varepsilon_l^2\big)$$

and $|S_{l,l'}| \le L(1+\varepsilon_l^2+\varepsilon_{l'}^2)$, (6.5) implies that

$$\big|N\phi_A'(t)\big| \le L\bigg(1+\sum_{l\le n}\mathbb{E}\big\langle\varepsilon_l^2\big\rangle_{t,A}\bigg) \le L\big(1+\mathbb{E}\big\langle\varepsilon_1^2\big\rangle_{t,A}\big), \tag{6.6}$$

where in the last inequality we used the fact that $\mathbb{E}\langle\varepsilon_l^2\rangle_{t,A}$ does not depend on l due to the symmetry of A. One can now repeat the proof of Theorem 2 to obtain the analogue of (5.3):

$$\mathbb{E}\big\langle\exp c(t)\varepsilon_1^2\big\rangle_{t,A} \le \exp\big(L\big(\beta^2+h\big)t\big)\mathbb{E}\big\langle\exp c_0\varepsilon_1^2\big\rangle_{0,A},$$

where $c(t) = c_0 - L(\beta^2+h)t > c_0/2$ for small enough β, h. Using (6.4) and (2.15), we can write

$$\mathbb{E}\big\langle\exp c_0\varepsilon_1^2\big\rangle_{0,A} = \mathbb{E}\frac{\langle I_A \exp c_0\varepsilon_1^2\rangle_0}{\langle I_A\rangle_0} = \mathbb{E}\frac{\langle I_A\rangle_0\langle\exp c_0\varepsilon_1^2\rangle_0}{\langle I_A\rangle_0} = \mathbb{E}\big\langle\exp c_0\varepsilon_1^2\big\rangle_0 \le L$$

for $c_0 < 1$ and small enough β, by Lemma 3. Hence, $\mathbb{E}\langle\varepsilon_1^2\rangle_{t,A} \le L$ and (6.6) finishes the proof of Lemma 5. □

To apply Lemma 4 to the sets of the type (6.2), we need to control $N^{-1}\mathbb{E}\log\langle I_A\rangle_0$. Let us notice that $\langle I_A\rangle_0$ for the sets in (6.2) is defined exactly in the same way as $\langle I_A\rangle$ (i.e. for $t=1$) for the sets of the type

$$\big\{\boldsymbol{\sigma}^1\colon |R_1 - r| \ge x\big\} \quad \text{or} \quad \big\{\big(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\big)\colon |R_{1,2} - q| \ge x\big\} \tag{6.7}$$

only for $N-1$ instead of N. The terms $\varepsilon a - \varepsilon^2 b/2$ in H_0 decouple in both numerator and denominator of the Gibbs average $\langle I_A\rangle_0$ and cancel after integration. Therefore, for simplicity of notations, we will show how to control $N^{-1}\mathbb{E}\log\langle I_A\rangle$ for A in (6.7) and then apply this result to $N^{-1}\mathbb{E}\log\langle I_A\rangle_0$ for A in (6.2).

For $\bar{q}\in[0,1]$ consider the Hamiltonian

$$h_t(\boldsymbol{\sigma}) = \sqrt{t}H_N(\boldsymbol{\sigma}) + \sum_{i\le N}\sigma_i\big(\sqrt{1-t}z_i\beta\sqrt{\xi'(\bar{q})}+h\big). \tag{6.8}$$

Let $\langle\cdot\rangle_t^-$ define the Gibbs average with respect to the Hamiltonian (6.8). Let us define $\bar{q}$ as any solution of the equation

$$\bar{q} = \mathbb{E}\langle R_{1,2}\rangle_0^-, \tag{6.9}$$

where the right-hand side depends on $\bar{q}$ through (6.8). We will show that there exists a solution close to q. Given $\bar{q}$ that satisfies (6.9) we define

$$\bar{r} = \mathbb{E}\langle R_1\rangle_0^-. \tag{6.10}$$

Lemma 6. *For small enough β we can find small enough $\alpha > 0$ such that*

$$\mathbb{E}\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle \le L \quad \textit{and} \quad \mathbb{E}\big\langle \exp N\alpha(R_1-\bar{r})^2\big\rangle \le L. \tag{6.11}$$

This lemma is all we really need if we use $\bar{q}$ instead of q in our main interpolation (2.2) and in Theorem 3, as well as $\bar{r}$ instead of r. In the next lemma we will show that $\bar{q},\bar{r}$ are actually close to q,r and the proof will explain how (6.9) and (6.10) give rise to the more explicit definitions in (1.5) and (2.7). As an observation, note that (6.11) implies that all solutions $\bar{q}$ of (6.9) are close to $\nu(R_{1,2})$ and if we could show that $\nu(R_{1,2})$ converges as $N\to\infty$, Lemma 7 would imply the uniqueness of q that satisfies (1.5). Of course, since our analysis is limited to small values of β, the uniqueness of q follows directly from the definition as well.

Lemma 7. *For small enough β and h there exists a solution of* (6.9) *such that*

$$|\bar{q}-q| \le \frac{L\log^2 N}{N}, \qquad |\bar{r}-r| \le \frac{L\log^2 N}{N}.$$

Before we prove Lemmas 7 and 6, let us first show how they imply Theorem 3.

Proof of Theorem 3. Lemma 6 implies that

$$\begin{aligned}\mathbb{E}\log\big\langle I\big(|R_{1,2}-\bar{q}|\ge x\big)\big\rangle &\le \mathbb{E}\log\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle - N\alpha x^2\\ &\le \log\mathbb{E}\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle - N\alpha x^2 \le L - N\alpha x^2.\end{aligned}$$

Using this for $N-1$ instead of N yields

$$\frac{1}{N}\mathbb{E}\log\big\langle I\big(|\hat{R}_{1,2}-\bar{q}|\ge x\big)\big\rangle_0 \le \frac{L}{N} - \alpha x^2$$

and by Lemma 4

$$\frac{1}{N}\mathbb{E}\log\big\langle I\big(|\hat{R}_{1,2}-\bar{q}|\ge x\big)\big\rangle_t \le \frac{L}{N} - \alpha x^2.$$

For $x = L(\log N/N)^{1/4}$ we get

$$\frac{1}{N}\mathbb{E}\log\left\langle I\left(|\hat{R}_{1,2}-\bar{q}|\ge L\left(\frac{\log N}{N}\right)^{1/4}\right)\right\rangle_t \le -L\left(\frac{\log N}{N}\right)^{1/2} =: \delta.$$

Gaussian concentration of measure (as in Corollary 2.2.5 in [5]) implies that

$$\frac{1}{N}\log\left\langle I\left(|\hat{R}_{1,2}-\bar{q}|\ge L\left(\frac{\log N}{N}\right)^{1/4}\right)\right\rangle_t \le -L\left(\frac{\log N}{N}\right)^{1/2}$$

with probability at least $1 - L\exp(-N\delta^2/L) \ge 1 - LN^{-K}$ for any $K>0$, by choosing L in the definition of x sufficiently large. Therefore, with probability at least $1-LN^{-K}$,

$$\left\langle I\left(|\hat{R}_{1,2}-\bar{q}|\ge L\left(\frac{\log N}{N}\right)^{1/4}\right)\right\rangle_t \le \exp\big(-L(N\log N)^{1/2}\big) \le LN^{-K}$$

and, thus,

$$\mathbb{E}\left\langle I\left(|\hat{R}_{1,2}-\bar{q}|\ge L\left(\frac{\log N}{N}\right)^{1/4}\right)\right\rangle_t \le LN^{-K}.$$

Lemma 7 implies

$$\mathbb{E}\Big\langle I\Big(|\hat{R}_{1,2}-q|\geq L\Big(\frac{\log N}{N}\Big)^{1/4}\Big)\Big\rangle_t \leq LN^{-K}$$

and this proves the first part of Theorem 3. The second part is proved similarly. □

Let us prove Lemma 7 first.

Proof of Lemma 7. If we denote

$$\boldsymbol{v}=\big(z_1\beta\sqrt{\xi'(\bar{q})}+h,\dots,z_N\beta\sqrt{\xi'(\bar{q})}+h\big)$$

Then

$$\langle R_{1,2}\rangle_0^-=\frac{1}{Z^2}\int_{S_N^2}\frac{1}{N}\big(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\big)\exp\big(\big(\boldsymbol{\sigma}^1,\boldsymbol{v}\big)+\big(\boldsymbol{\sigma}^2,\boldsymbol{v}\big)\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^1\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^2\big),$$

where $Z=\int_{S_N}\exp((\boldsymbol{\sigma},\boldsymbol{v}))\,\mathrm{d}\lambda_N(\boldsymbol{\sigma})$. If O is an orthogonal transformation such that $O\boldsymbol{v}=(0,\dots,0,|\boldsymbol{v}|)$ then making a change of variables $\boldsymbol{\sigma}^l\to O^{-1}\boldsymbol{\sigma}^l$ we get

$$\langle R_{1,2}\rangle_0^-=\frac{1}{Z^2}\int_{S_N^2}\frac{1}{N}\big(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\big)\exp\big(\varepsilon_1|\boldsymbol{v}|+\varepsilon_2|\boldsymbol{v}|\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^1\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^2\big)$$

and $Z=\int_{S_N}\exp(\varepsilon|\boldsymbol{v}|)\,\mathrm{d}\lambda_N(\boldsymbol{\sigma})$. By (3.21)

$$\frac{1}{N}\big(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2\big)=R_{1,2}=\sqrt{\Big(1-\frac{\varepsilon_1^2}{N}\Big)\Big(1-\frac{\varepsilon_2^2}{N}\Big)}\hat{R}_{1,2}+\frac{1}{N}\varepsilon_1\varepsilon_2$$

and by (2.14)

$$\begin{aligned}&\int_{S_N^2}\sqrt{\Big(1-\frac{\varepsilon_1^2}{N}\Big)\Big(1-\frac{\varepsilon_2^2}{N}\Big)}\hat{R}_{1,2}\exp\big(\varepsilon_1|\boldsymbol{v}|+\varepsilon_2|\boldsymbol{v}|\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^1\big)\,\mathrm{d}\lambda_N\big(\boldsymbol{\sigma}^2\big)\\&\quad=a_N^2\bigg(\int_{-\sqrt{N}}^{\sqrt{N}}\Big(1-\frac{\varepsilon^2}{N}\Big)^{(N-2)/2}\exp\big(\varepsilon|\boldsymbol{v}|\big)\,\mathrm{d}\varepsilon\bigg)^2\int_{S_{N-1}^2}\hat{R}_{1,2}\,\mathrm{d}\lambda_{N-1}\big(\hat{\boldsymbol{\sigma}}^1\big)\,\mathrm{d}\lambda_{N-1}\big(\hat{\boldsymbol{\sigma}}^2\big)=0\end{aligned}$$

since the last integral is equal to zero by symmetry. Therefore,

$$\langle R_{1,2}\rangle_0^-=\frac{1}{N}\langle\varepsilon_1\varepsilon_2\rangle_0^-=\big\langle N^{-1/2}\varepsilon\big\rangle_0^{-2} \tag{6.12}$$

and using (2.14) again

$$\big\langle N^{-1/2}\varepsilon\big\rangle_0^-=\int_{-\sqrt{N}}^{\sqrt{N}}\frac{\varepsilon}{\sqrt{N}}\Big(1-\frac{\varepsilon^2}{N}\Big)^{(N-3)/2}\exp\big(\varepsilon|\boldsymbol{v}|\big)\,\mathrm{d}\varepsilon\Big/\int_{-\sqrt{N}}^{\sqrt{N}}\Big(1-\frac{\varepsilon^2}{N}\Big)^{(N-3)/2}\exp\big(\varepsilon|\boldsymbol{v}|\big)\,\mathrm{d}\varepsilon.$$

By making a change of variable $\varepsilon=\sqrt{N}x$ we can rewrite the right hand side as

$$\big\langle N^{-1/2}\varepsilon\big\rangle_0^-=\int_{-1}^{1}x\exp N\varphi(x)\,\mathrm{d}x\Big/\int_{-1}^{1}\exp N\varphi(x)\,\mathrm{d}x, \tag{6.13}$$

where

$$\varphi(x)=cx+\frac{N-3}{2N}\log\big(1-x^2\big) \tag{6.14}$$

and

$$c = N^{-1/2}|\boldsymbol{v}| = \left(\frac{1}{N}\sum_{i\le N}\bigl(z_i\beta\sqrt{\xi(\bar{q})}+h\bigr)^2\right)^{1/2}. \tag{6.15}$$

Let x_0 denotes the point where $\varphi(x)$ achieves its maximum which satisfies

$$\varphi'(x_0) = 0 \quad \Longrightarrow \quad c = \frac{N-3}{N}\frac{x_0}{1-x_0^2}. \tag{6.16}$$

Since $|\varepsilon|/\sqrt{N} \le 1$ and $|x_0| \le 1$,

$$\bigl|\mathbb{E}\langle N^{-1/2}\varepsilon\rangle_0^{-2} - \mathbb{E}x_0^2\bigr| \le 2\mathbb{E}\left|\int_{-1}^{1}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x\right| \bigg/ \int_{-1}^{1}\exp N\varphi(x)\,\mathrm{d}x. \tag{6.17}$$

For c in (6.15) and $c' > 2h^2$,

$$\begin{aligned}\mathbb{P}\bigl(c \ge c'\bigr) &= \mathbb{P}\left(\sum_{i\le N}\bigl(z_i\beta\sqrt{\xi'(\bar{q})}+h\bigr)^2 \ge Nc'^2\right)\\ &\le \mathbb{P}\left(2\beta^2\xi'(\bar{q})\sum_{i\le N}z_i^2 \ge N\bigl(c'^2-2h^2\bigr)\right) = \mathbb{P}\left(\sum_{i\le N}z_i^2 \ge Nc''\right) \le \exp(-LN),\end{aligned} \tag{6.18}$$

where L can be made arbitrarily large by increasing c'.

Let us now assume that the event $\{c \le c'\}$ occurs. Then (6.16) implies that $|x_0| \le 1-\delta$ for some $\delta > 0$ that depends on c' only. Let us define

$$\Omega = \left\{x \in [-1,1]\colon\ |x-x_0| \le \omega = \sqrt{\frac{L\log N}{N}}\right\}$$

for L large enough and write $\int_{-1}^{1}\exp N\varphi(x)\,\mathrm{d}x = \mathrm{I}+\mathrm{II}$, where

$$\mathrm{I} = \int_{\Omega}\exp N\varphi(x)\,\mathrm{d}x \quad \text{and} \quad \mathrm{II} = \int_{\Omega^c}\exp N\varphi(x)\,\mathrm{d}x.$$

We have

$$\varphi''(x) = -\frac{1+x^2}{(1-x^2)^2} \le -1 \tag{6.19}$$

and for $|x| \le 1-\delta/2$, clearly, $-L \le \varphi''(x)$ and $|\varphi'''(x)| \le L$. Since $\varphi'(x_0) = 0$, we have $\varphi(x) \ge \varphi(x_0) - L(x-x_0)^2$ for $x \in \Omega$ and, therefore,

$$\begin{aligned}\mathrm{I} &\ge \exp N\varphi(x_0)\int_{\Omega}\exp\bigl(-LN(x-x_0)^2\bigr)\\ &= \exp\bigl(N\varphi(x_0)\bigr)\frac{1}{\sqrt{N}}\int_{|y|\le(L\log N)^{1/2}}\exp\bigl(-Ly^2\bigr)\,\mathrm{d}y \ge \frac{1}{L\sqrt{N}}\exp N\varphi(x_0).\end{aligned} \tag{6.20}$$

On the other hand, by (6.19), $\varphi(x) \le \varphi(x_0) - (x-x_0)^2/2$ and, thus,

$$\mathrm{II} \le \exp\bigl(N\varphi(x_0)\bigr)\frac{1}{\sqrt{N}}\int_{|y|\ge(L\log N)^{1/2}}\exp\bigl(-Ly^2\bigr)\,\mathrm{d}y \le \frac{L}{N^K}\exp N\varphi(x_0),$$

where K can be made arbitrarily large by a proper choice of L in the definition of Ω. The denominator in (6.17) can be bounded from below by

$$\int_{-1}^{1} \exp N\varphi(x)\,\mathrm{d}x \geq \mathrm{I} \geq \frac{1}{L\sqrt{N}} \exp N\varphi(x_0). \tag{6.21}$$

Next, we write $\int_{-1}^{1}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x = \mathrm{III}+\mathrm{IV}$, where

$$\mathrm{III} = \int_{\Omega}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x \quad \text{and} \quad \mathrm{IV} = \int_{\Omega^c}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x.$$

We control IV by

$$|\mathrm{IV}| \leq 2|\mathrm{II}| \leq \frac{L}{N^K}\exp N\varphi(x_0). \tag{6.22}$$

To control III we use that for $x \in \Omega$

$$\left|\varphi(x)-\varphi(x_0)-\frac{1}{2}\varphi''(x_0)(x-x_0)^2\right| \leq L\left(\frac{\log N}{N}\right)^{3/2} =: \Delta.$$

We have

$$\begin{aligned}
\mathrm{III} &= \int_{x_0-\omega}^{x_0}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x + \int_{x_0}^{x_0+\omega}(x-x_0)\exp N\varphi(x)\,\mathrm{d}x \\
&\leq \int_{x_0-\omega}^{x_0}(x-x_0)\exp N\left(\varphi(x_0)+\frac{1}{2}\varphi''(x_0)(x-x_0)^2-\Delta\right)\mathrm{d}x \\
&\quad + \int_{x_0}^{x_0+\omega}(x-x_0)\exp N\left(\varphi(x_0)+\frac{1}{2}\varphi''(x_0)(x-x_0)^2+\Delta\right)\mathrm{d}x \\
&= \left(\mathrm{e}^{N\Delta}-\mathrm{e}^{-N\Delta}\right)\int_{x_0}^{x_0+\omega}(x-x_0)\exp N\left(\varphi(x_0)+\frac{1}{2}\varphi''(x_0)(x-x_0)^2\right)\mathrm{d}x \\
&\leq LN\Delta\omega\exp N\varphi(x_0)\int_{x_0}^{x_0+\omega}\exp\left(-\frac{1}{2}N(x-x_0)^2\right)\mathrm{d}x \\
&\leq LN^{1/2}\Delta\omega\exp N\varphi(x_0) \leq \frac{L\log^2 N}{N^{3/2}}\exp N\varphi(x_0).
\end{aligned}$$

The lower bound can be carried out similarly and, thus,

$$|\mathrm{III}| \leq \frac{L\log^2 N}{N^{3/2}}\exp N\varphi(x_0).$$

Combining this with (6.17), (6.18), (6.20) and (6.22) proves

$$\begin{aligned}
\left|\mathbb{E}\langle N^{-1/2}\varepsilon\rangle_0^{-2}-\mathbb{E}x_0^2\right| &\leq \exp(-LN)+\frac{L}{N^K}+\mathbb{E}\frac{L\log^2 N\exp N\varphi(x_0)}{N^{3/2}}\Bigg/\frac{\exp N\varphi(x_0)}{L\sqrt{N}} \\
&\leq \frac{L\log^2 N}{N}.
\end{aligned}$$

By (6.12), we proved that

$$\left|\mathbb{E}\langle R_{1,2}\rangle_0^{-}-\mathbb{E}x_0^2\right| \leq \frac{L\log^2 N}{N}. \tag{6.23}$$

If we denote

$$c_N = \frac{N}{N-3} c = \frac{N}{N-3} \frac{|\boldsymbol{v}|}{\sqrt{N}},$$

then solving (6.16) for x_0 gives

$$x_0 = \frac{2c_N}{1+\sqrt{1+4c_N^2}} \quad \text{and} \quad x_0^2 = 1 - \frac{2}{1+\sqrt{1+2c_N^2}}. \tag{6.24}$$

It is easy to check that the first two derivatives of $y(x) = 1/(1+\sqrt{1+4x})$ are bounded by an absolute constant for $x \geq 0$ and, therefore,

$$\left|y\left(c_N^2\right) - y\left(\mathbb{E}c_N^2\right) - y'\left(\mathbb{E}c_N^2\right)\left(c_N^2 - \mathbb{E}c_N^2\right)\right| \leq L\left(c_N^2 - \mathbb{E}c_N^2\right)^2.$$

Taking expectations proves that

$$\left|\mathbb{E}x_0^2 - \left(1 - \frac{2}{1+\sqrt{1+4\mathbb{E}c_N^2}}\right)\right| \leq L\mathbb{E}\left(c_N^2 - \mathbb{E}c_N^2\right)^2 \leq \frac{L}{N} \tag{6.25}$$

since

$$c_N^2 = \left(\frac{N}{N-3}\right)^2 \frac{1}{N} \sum_{i \leq N} \left(z_i \beta \sqrt{\xi'(\bar{q})} + h\right)^2.$$

If we denote

$$\delta = \mathbb{E}\langle R_{1,2}\rangle_0^- - \left(1 - \frac{2}{1+\sqrt{1+4\mathbb{E}c_N^2}}\right),$$

then (6.23) and (6.25) imply that $|\delta| \leq L \log^2 N/N$. By (6.9), $\mathbb{E}\langle R_{1,2}\rangle_0^- = \bar{q}$ and, therefore,

$$\bar{q} - \delta = 1 - \frac{2}{1+\sqrt{1+4\mathbb{E}c_N^2}}$$

or, equivalently,

$$\mathbb{E}c_N^2 = \frac{\bar{q}-\delta}{(1-\bar{q}+\delta)^2} = \left(\frac{N}{N-3}\right)^2 \left(\beta^2 \xi'(\bar{q}) + h^2\right).$$

Comparing with (1.5), it is now a simple exercise to show that

$$|\bar{q} - q| \leq \frac{L \log^2 N}{N}$$

and this proves the first part of Lemma 7. The computation of $\bar{r}$ is slightly different. If $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^N$ then

$$\langle R_1\rangle_0^- = \frac{1}{Z} \int_{S_N} \frac{1}{N} (\boldsymbol{\sigma}, \mathbf{1}) \exp(\boldsymbol{\sigma}, \boldsymbol{v})\, \mathrm{d}\lambda_N(\boldsymbol{\sigma}) = \frac{1}{Z} \int_{S_N} \frac{1}{N} \left(O^{\mathrm{T}} \boldsymbol{\sigma}, \mathbf{1}\right) \exp \varepsilon |v|\, \mathrm{d}\lambda_N(\boldsymbol{\sigma}),$$

where O is the orthogonal transformation as above. Note that the last row of O is $\boldsymbol{v}/|\boldsymbol{v}|$. Next, we use (2.13) to write $\int_{S_N}(O^{\mathrm{T}}\boldsymbol{\sigma}, \mathbf{1}) \exp\varepsilon|v| \,\mathrm{d}\lambda_N(\boldsymbol{\sigma})$ as

$$a_N \int_{-\sqrt{N}}^{\sqrt{N}} \mathrm{d}\varepsilon \exp\varepsilon|\boldsymbol{v}| \left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2} \int_{S_{N-1}} \left(O^{\mathrm{T}}\left(\sqrt{\frac{N-\varepsilon^2}{N-1}}\hat{\boldsymbol{\sigma}}, \varepsilon\right), \mathbf{1}\right) \mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}})$$
$$= a_N \int_{-\sqrt{N}}^{\sqrt{N}} \mathrm{d}\varepsilon \exp\varepsilon|\boldsymbol{v}| \left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2} \int_{S_{N-1}} \big(O^{\mathrm{T}}(0,\dots,0,\varepsilon), \mathbf{1}\big) \,\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}})$$

by symmetry $\hat{\boldsymbol{\sigma}} \to -\hat{\boldsymbol{\sigma}}$. Since the last column of O^{T} is $\boldsymbol{v}/|\boldsymbol{v}|$

$$\big(O^{\mathrm{T}}(0,\dots,0,\varepsilon), \mathbf{1}\big) = \frac{1}{|\boldsymbol{v}|}\varepsilon \sum_{i\le N} v_i$$

and, therefore,

$$\int_{S_N} \big(O^{\mathrm{T}}\boldsymbol{\sigma}, \mathbf{1}\big) \exp\varepsilon|v| \,\mathrm{d}\lambda_N(\boldsymbol{\sigma}) = a_N \frac{1}{|\boldsymbol{v}|} \sum_{i\le N} v_i \int_{-\sqrt{N}}^{\sqrt{N}} \varepsilon \exp\varepsilon|\boldsymbol{v}| \left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2} \mathrm{d}\varepsilon.$$

Similarly

$$Z = a_N \int_{-\sqrt{N}}^{\sqrt{N}} \exp\varepsilon|\boldsymbol{v}| \left(1-\frac{\varepsilon^2}{N}\right)^{(N-3)/2} \mathrm{d}\varepsilon$$

and making the change of variable $\varepsilon = \sqrt{N}x$ we get

$$\langle R_1\rangle_0^- = \frac{1}{\sqrt{N}} \frac{1}{|\boldsymbol{v}|} \sum_{i\le N} v_i \int_{-1}^{1} x \exp N\varphi(x)\,\mathrm{d}x \Big/ \int_{-1}^{1} \exp N\varphi(x)\,\mathrm{d}x.$$

Repeating the argument leading to (6.23) one can now show that

$$\left|\mathbb{E}\langle R_1\rangle_0^- - \mathbb{E}\frac{1}{\sqrt{N}}\frac{1}{|\boldsymbol{v}|}\sum_{i\le N} v_i x_0\right| \le \frac{L\log^2 N}{N}. \tag{6.26}$$

By (6.24),

$$\frac{1}{\sqrt{N}}\frac{1}{|\boldsymbol{v}|}\sum_{i\le N} v_i x_0 = \frac{1}{N-3}\sum_{i\le N} v_i \frac{2}{1+\sqrt{1+4c_N^2}}.$$

Since c_N^2 is concentrated near $\mathbb{E}(z_1\beta\sqrt{\xi'(\bar{q})}+h)^2 = \beta^2\xi'(\bar{q})+h^2$ and $\mathbb{E}v_i = h$, it is a simple exercise to show that

$$\left|\mathbb{E}\frac{1}{N-3}\sum_{i\le N} v_i \frac{2}{1+\sqrt{1+4c_N^2}} - \frac{2h}{1+\sqrt{1+4(\beta^2\xi'(\bar{q})+h^2)}}\right| \le \frac{L}{N}.$$

Since $|\bar{q}-q| \le L\log^2 N/N$, we get

$$\left|\mathbb{E}\langle R_1\rangle_0^- - \frac{2h}{1+\sqrt{1+4(\beta^2\xi'(q)+h^2)}}\right| \le \frac{L\log^2 N}{N}$$

and since by (1.5)

$$\frac{2h}{1+\sqrt{1+4(\beta^2\xi'(q)+h^2)}} = h(1-q) = r$$

we proved that $|\bar{r}-r| \le L\log^2 N/N$. This finishes the proof of Lemma 7. □

Proof of Lemma 6. We will use that $\langle\cdot\rangle = \langle\cdot\rangle_1^-$ and proceed by interpolation in (6.8). If is easy to show similarly to Theorem 1 that for a function $f = f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)$,

$$\frac{\partial}{\partial t}\mathbb{E}\langle f\rangle_t^- = N\beta^2 \sum_{1\le l<l'\le n} \mathbb{E}\big\langle f\,\Delta(R_{l,l'})\big\rangle_t^- - N\beta^2 n\sum_{l\le n}\mathbb{E}\big\langle f\,\Delta(R_{l,n+1})\big\rangle_t^-$$
$$+ N\beta^2\frac{n(n+1)}{2}\mathbb{E}\big\langle f\,\Delta(R_{n+1,n+2})\big\rangle_t^-,$$

where

$$\Delta(R_{l,l'}) = \xi(R_{l,l'}) - R_{l,l'}\xi'(\bar{q}) + \theta(\bar{q})$$

and $\theta(x) = x\xi'(x)-\xi(x)$. Since ξ is three times continuously differentiable we have

$$\big|\Delta(R_{l,l'})\big| \le L(R_{l,l'}-\bar{q})^2.$$

For $n=2$ and for any $k\ge 1$ this implies, by Hölder's inequality,

$$\frac{\partial}{\partial t}\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k}\big\rangle_t^- \le LN\beta^2\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k+2}\big\rangle_t^-.$$

Next, we use an argument due to R. Latala. We can write,

$$\frac{\partial}{\partial t}\mathbb{E}\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_t^- \le \sum_{k\ge 1} LN\beta^2\frac{N^k\alpha^k}{k!}\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k+2}\big\rangle_t^-$$
$$\le LN\beta^2\mathbb{E}\big\langle (R_{1,2}-\bar{q})^2\exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_t^-.$$

For $\alpha(t) = \alpha - L\beta^2 t$ this implies

$$\frac{\partial}{\partial t}\mathbb{E}\big\langle \exp N\alpha(t)(R_{1,2}-\bar{q})^2\big\rangle_t^- \le 0$$

and, therefore,

$$\mathbb{E}\big\langle \exp N\alpha(t)(R_{1,2}-\bar{q})^2\big\rangle_t^- \le \mathbb{E}\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_0^-.$$

Next, since

$$\mathbb{E}\big\langle (R_1-\bar{r})^{2k}(R_{l,l'}-\bar{q})^2\big\rangle_t^- \le \big(\mathbb{E}\big\langle (R_1-\bar{r})^{2k+2}\big\rangle_t^-\big)^{2k/(2k+2)}\big(\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k+2}\big\rangle_t^-\big)^{2/(2k+2)}$$
$$\le \frac{k}{k+1}\mathbb{E}\big\langle (R_1-\bar{r})^{2k+2}\big\rangle_t^- + \frac{1}{k+1}\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k+2}\big\rangle_t^-$$

we can bound the derivative of $\mathbb{E}\langle \exp N\alpha(R_1-\bar{r})^2\rangle_t^-$ by

$$LN\beta^2\sum_{k\ge 1}\frac{N^k\alpha^k}{k!}\bigg(\frac{k}{k+1}\mathbb{E}\big\langle (R_1-\bar{r})^{2k+2}\big\rangle_t^- + \frac{1}{k+1}\mathbb{E}\big\langle (R_{1,2}-\bar{q})^{2k+2}\big\rangle_t^-\bigg)$$
$$\le LN\beta^2\mathbb{E}\big\langle (R_1-\bar{r})^2\exp N\alpha(R_1-\bar{r})^2\big\rangle_t^- + \frac{L\beta^2}{\alpha}\mathbb{E}\big\langle \exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_t^-.$$

For $\alpha(t)=\alpha-L\beta^2 t$ this implies that

$$\begin{aligned}\frac{\partial}{\partial t}\mathbb{E}\big\langle\exp N\alpha(t)(R_1-\bar{r})^2\big\rangle_t^- &\le \frac{L\beta^2}{\alpha(t)}\mathbb{E}\big\langle\exp N\alpha(t)(R_{1,2}-\bar{q})^2\big\rangle_t^- \\ &\le \frac{L\beta^2}{\alpha-L\beta^2}\mathbb{E}\big\langle\exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_0^-\end{aligned}$$

and, thus,

$$\mathbb{E}\big\langle\exp N\alpha(1)(R_1-\bar{r})^2\big\rangle_1^- \le \mathbb{E}\big\langle\exp N\alpha(R_1-\bar{r})^2\big\rangle_0^- + \frac{L\beta^2}{\alpha-L\beta^2}\mathbb{E}\big\langle\exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_0^-.$$

To finish the proof of Lemma 6 it remains to show that for small enough α,

$$\mathbb{E}\big\langle\exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_0^- \le L \quad\text{and}\quad \mathbb{E}\big\langle\exp N\alpha(R_1-\bar{r})^2\big\rangle^-{}_0 \le L.$$

By (6.9) and Jensen's inequality

$$\begin{aligned}\mathbb{E}\big\langle\exp N\alpha(R_{1,2}-\bar{q})^2\big\rangle_0^- &\le \mathbb{E}\big\langle\exp N\alpha(R_{1,2}-R_{3,4})^2\big\rangle_0^-. \\ &= \mathbb{E}\frac{1}{Z^4}\int_{S_N^4}\exp\Big(N\alpha(R_{1,2}-R_{3,4})^2+\sum_{l\le 4}\big(\boldsymbol{\sigma}^l,\boldsymbol{v}\big)\Big)\,\mathrm{d}\lambda_N^4,\end{aligned}$$

as in the beginning of Lemma 7. For $\boldsymbol{v}$ and O defined in Lemma 7 we have

$$\begin{aligned}&\int_{S_N^4}\exp\Big(N\alpha(R_{1,2}-R_{3,4})^2+\sum_{l\le 4}\big(\boldsymbol{\sigma}^l,\boldsymbol{v}\big)\Big)\,\mathrm{d}\lambda_N^4 \\ &\quad=\int_{S_N^4}\exp\Big(N\alpha(R_{1,2}-R_{3,4})^2+\sum_{l\le 4}\varepsilon_l|\boldsymbol{v}|\Big)\,\mathrm{d}\lambda_N^4.\end{aligned}\tag{6.27}$$

Since

$$\begin{aligned}(R_{1,2}-R_{3,4})^2 &\le 2(\hat{R}_{1,2}-\hat{R}_{3,4})^2+\frac{2}{N^2}(\varepsilon_1\varepsilon_2-\varepsilon_3\varepsilon_4)^2 \\ &\le 4\hat{R}_{1,2}^2+4\hat{R}_{3,4}^2+\frac{2}{N^2}(\varepsilon_1\varepsilon_2-\varepsilon_3\varepsilon_4)^2,\end{aligned}$$

using (2.14), the right hand side of (6.27) is bounded by

$$a_N^4\int_{[-\sqrt{N},\sqrt{N}]^4}\exp\Big(\frac{2\alpha}{N}(\varepsilon_1\varepsilon_2-\varepsilon_3\varepsilon_4)^2+\sum_{l\le 4}\varepsilon_l|\boldsymbol{v}|\Big)\,\mathrm{d}\boldsymbol{\varepsilon}\Big(\int_{S_{N-1}^2}\exp\big(4\alpha N\hat{R}_{1,2}^2\big)\,\mathrm{d}\lambda_{N-1}^2\big(\hat{\boldsymbol{\sigma}}^1,\hat{\boldsymbol{\sigma}}^2\big)\Big)^2,$$

where $\mathrm{d}\boldsymbol{\varepsilon}=\mathrm{d}\varepsilon_1\cdots\mathrm{d}\varepsilon_4$. For a fixed $\hat{\boldsymbol{\sigma}}^2\in S_{N-1}$, let Q be an orthogonal transformation in $\mathbb{R}^{N-1}$ such that

$$Q\hat{\boldsymbol{\sigma}}^2=\big(0,\dots,0,\big|\hat{\boldsymbol{\sigma}}^2\big|\big)=\big(0,\dots,0,\sqrt{N-1}\big).$$

Then

$$\hat{R}_{1,2}=\frac{1}{N-1}\big(Q\hat{\boldsymbol{\sigma}}^1,Q\hat{\boldsymbol{\sigma}}^2\big)=\frac{1}{\sqrt{N-1}}\big(Q\hat{\boldsymbol{\sigma}}^1\big)_{N-1},$$

where $(\cdot)_{N-1}$ denotes the $(N-1)$st coordinate. Therefore, by rotational invariance and then (2.14),

$$\int_{S_{N-1}^2} \exp\bigl(4\alpha N \hat{R}_{1,2}^2\bigr)\,\mathrm{d}\lambda_{N-1}^2\bigl(\hat{\boldsymbol{\sigma}}^1, \hat{\boldsymbol{\sigma}}^2\bigr) = \int_{S_{N-1}} \exp 4\alpha \frac{N}{N-1}\varepsilon^2\,\mathrm{d}\lambda_{N-1}(\hat{\boldsymbol{\sigma}})$$
$$\le a_{N-1}\int_{-\sqrt{N-1}}^{\sqrt{N-1}} \exp\bigl(5\alpha\varepsilon^2\bigr)\biggl(1-\frac{\varepsilon^2}{N-1}\biggr)^{(N-4)/2}\,\mathrm{d}\varepsilon$$
$$\le L\int_{-\infty}^{\infty}\exp\bigl(5\alpha\varepsilon^2 - L\varepsilon^2\bigr)\,\mathrm{d}\varepsilon \le L$$

for small enough α. Therefore, the right hand side of (6.27) is bounded for small α by

$$L\int_{[-\sqrt{N},\sqrt{N}]^4}\exp\biggl(\frac{2\alpha}{N}(\varepsilon_1\varepsilon_2-\varepsilon_3\varepsilon_4)^2+\sum_{l\le 4}\varepsilon_l|\boldsymbol{v}|\biggr)\,\mathrm{d}\boldsymbol{\varepsilon}.$$

Making the change of variables $\varepsilon_l = \sqrt{N}x_l$ (as in (6.13)) proves that $\mathbb{E}\langle \exp N\alpha(R_{1,2}-\bar{q})^2\rangle_0^-$ is bounded up to a constant by

$$\mathbb{E}\int_{[-1,1]^4}\exp N\Phi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\Big/\biggl(\int_{-1}^{1}\exp N\varphi(x)\,\mathrm{d}x\biggr)^4,$$

where

$$\Phi(\boldsymbol{x}) = \Phi(x_1,x_2,x_3,x_4) = 2\alpha(x_1x_2-x_3x_4)^2+\sum_{l\le 4}\varphi(x_l) \tag{6.28}$$

and where $\varphi(x)$ was defined in (6.14). We will use this bound only on the event $\{c\le c'\}$ since by (6.18)

$$\mathbb{E}\bigl\langle \exp N\alpha(R_{1,2}-\bar{q})^2\bigr\rangle_0^- \le \exp(4N\alpha - LN) + \mathbb{E}\bigl\langle \exp N\alpha(R_{1,2}-\bar{q})^2\bigr\rangle_0^- I\bigl(c\le c'\bigr)$$

and L can be made as large as necessary by taking c' sufficiently large. Since by (6.19), $\varphi''(x)\le -1$, for small enough α the function $\Phi(\boldsymbol{x})$ will be strictly concave on $[-1,1]^4$. It is obvious that for $\boldsymbol{x}_0 = (x_0,x_0,x_0,x_0)$

$$\frac{\partial\Phi}{\partial x_l}(\boldsymbol{x}_0) = \varphi'(x_0) = 0$$

which implies that $\boldsymbol{x}_0$ is the unique maximum of Φ. Strict concavity now implies

$$\Phi(\boldsymbol{x}) \le 4\varphi(x_0) - \frac{1}{L}\sum_{l\le 4}(x_l - x_0)^2$$

and, thus,

$$\int_{[-1,1]^4}\exp N\Phi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \exp 4N\varphi(x_0)\biggl(\int_{-1}^{1}\exp\biggl(-\frac{1}{L}N(x-x_0)^2\biggr)\,\mathrm{d}x\biggr)^4 \le \frac{L}{N^2}\exp 4N\varphi(x_0).$$

Combining this with (6.21) finally proves that $\mathbb{E}\langle \exp N\alpha(R_{1,2}-\bar{q})^2\rangle_0^- \le L$. The proof of the corresponding statement for $R_1 - \bar{r}$ is similar. □

Acknowledgment

I would like to thank the referees for a very thorough reading of the paper that lead to a great number of helpful comments and suggestions.

References

[1] Crisanti, A. and Sommers, H. J. The spherical p-spin interaction spin glass model: The statics. *Z. Phys. B. Condensed Matter* **83** (1992) 341–354.

[2] Guerra, F. and Toninelli, F. L. Central limit theorem for fluctuations in the high temperature region of the Sherrington–Kirkpatrick spin glass model. *J. Math. Phys.* **43** (2002) 6224–6237. MR1939641

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

[4] Talagrand, M. Replica symmetry breaking and exponential inequalities for the Sherrington–Kirkpatrick model. *Ann. Probab.* **28** (2000) 1018–1062. MR1797303

[5] Talagrand, M. *Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models*. Springer, Berlin, 2003. MR1993891

[6] Talagrand, M. Free energy of the spherical mean field model. *Probab. Theory Related Fields* **134** (2006) 339–382. MR2226885

[7] Talagrand, M. Large deviations, Guerra's and A.S.S. schemes, and the Parisi hypothesis. *J. Stat. Phys.* **126** (2007) 837–894. MR2311888