Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 2, 1–25.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-1763

Central limit theorems for cavity and local fields of the Sherrington-Kirkpatrick model

Wei-Kuo Chen*

Abstract

One of the remarkable applications of the cavity method in the mean field spin glasses is to prove the validity of the Thouless-Anderson-Palmer (TAP) system of equations in the Sherrington-Kirkpatrick (SK) model in the high temperature regime. This naturally leads us to the study of the limit laws for cavity and local fields. The first quantitative results for both fields were obtained by Chatterjee [1] using Stein's method. In this paper, we approach these problems using the Gaussian interpolation technique and establish central limit theorems for both fields by giving moment estimates of all orders.

Keywords: Sherrington-Kirkpatrick model; Stein's method; TAP equations.
AMS MSC 2010: 60K35; 82B44.
Submitted to EJP on January 26, 2012, final version accepted on January 4, 2013.

1 Introduction and Main Results

1.1 The Sherrington-Kirkpatrick Model and TAP Equations

In the middle 70's, Sherrington and Kirkpatrick [3] introduced a mean field spin glass, now known as the SK model, with the aim of understanding the strange magnetic behaviors of certain alloys. Even for such a seemingly simple formulation, the SK model has already presented very beautiful and intricate structures conjectured by physicists (see [2]) and has been intensively studied in the mathematical community in recent decades. In particular, many results regarding the behavior of the overlap, a quantity that measures the difference between two spin configurations sampled independently from the Gibbs measure, on the high temperature phase are now well-known (see [5, 6] for detailed account). From these, we will present an approach to central limit theorems (CLT) for the cavity and local fields via the Gaussian interpolation technique.

Let us begin with the description of the SK model. For each positive integer N, we consider the configuration space $\Sigma_N := \{-1,+1\}^N$. Every element $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N$ is called a *spin configuration* and its coordinates are called *spins*. Let $\boldsymbol{g} = \{g_{ij}\}_{1\leq i,j\leq N}$

*University of California, Irvine, USA. E-mail: weikuoc@uci.edu

be r.v.s satisfying that $g_{ij} = g_{ji}$ and $\{g_{ij} : 1 \le i \le j \le N\}$ are i.i.d. standard Gaussian. The Hamiltonian H_N for the Sherrington-Kirkpatrick (SK) model is defined on Σ_N through

$$-H_N(\boldsymbol{\sigma}) = \frac{\beta}{\sqrt{N}} \sum_{i<j\le N} g_{ij}\sigma_i\sigma_j + h\sum_{i\le N}\sigma_i. \tag{1.1}$$

Here, $\beta > 0$ is the inverse temperature and $h \in \mathbb{R}$ stands for the strength of the external field. We define the *Gibbs measure* G_N by

$$G_N(\{\boldsymbol{\sigma}\}) = \frac{\exp\left(-H_N(\boldsymbol{\sigma})\right)}{Z_N},$$

where Z_N is the normalizing factor, called the *partition function*.

Throughout the paper, we denote by $(\boldsymbol{\sigma}^\ell)_{\ell\ge 1}$ an i.i.d. sequence of configurations sampled from G_N (with the same given disorder $\boldsymbol{g}$). These r.v.s are also called *replicas* in physics. For each function f defined on Σ_N^n, we use $\langle f\rangle$ to denote the Gibbs average of f corresponding to the Gibbs measure G_N, namely,

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

In the present paper, we will focus on the high temperature (i.e. β small enough) behavior of the SK model. Let us define the *overlap* of the configurations $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$ as

$$R_{1,2} = R_{1,2}(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2) := \frac{1}{N}\sum_{j\le N}\sigma_j^1\sigma_j^2.$$

In the high temperature regime, it is well-known that as N tends to infinity, this quantity under the Gibbs measure converges a.s. to a constant q, which is the unique solution to

$$q = E\tanh^2(\beta z\sqrt{q} + h), \tag{1.2}$$

where z is a standard Gaussian r.v. More precisely, concluding from Theorem 1.4.1 in Talagrand [5], for fixed $\beta_0 < 1/2$, we have the quantitative result

$$E\left\langle (R_{1,2} - q)^{2k}\right\rangle \le \frac{K}{N^k}, \tag{1.3}$$

for every $\beta \le \beta_0$ and h, where K is a constant depending only on k and β_0. Note that for simplicity we will only restrict our discussion to the region $\beta \le \beta_0 < 1/2$. Indeed, it has been shown that the moment estimate (1.3) is valid on a rigorously defined high temperature region in Talagrand [6] and such region contains the case $\beta \le \beta_0 < 1/2$. By a similar argument, our main results can be extended to this high temperature region as well.

An important objective in the SK model is to understand the marginal distributions of the spins. The key observation is that since each spin takes only two values, the marginal distribution at site i can be completely determined by the Gibbs average $\langle\sigma_i\rangle$. This then leads to the computation for $\langle\sigma_1\rangle,\ldots,\langle\sigma_N\rangle$ via the approach of the Thouless-Anderson-Palmer (TAP) system of equations as outlined in [7]:

$$\langle\sigma_i\rangle \approx \tanh\left(\frac{\beta}{\sqrt{N}}\sum_{j\le N, j\ne i} g_{ij}\langle\sigma_j\rangle + h - \beta^2(1-q)\langle\sigma_i\rangle\right), \quad 1\le i\le N.$$

Here $\approx$ means that two quantities are approximately equal with high probability. The first rigorous proof of the validity of TAP equations in the high temperature phase was established by Talagrand (see Theorem 2.4.20 [4]) based on the cavity method. In the later book [5], Corollary 1.7.8 implies that all N equations hold simultaneously with high probability.

The idea of the cavity method is to reduce the N-spin system to a smaller system by removing a spin. More precisely the procedure can be described as follows. Recall the Hamiltonian H_N with inverse temperature β and external field h from (1.1). We define the Hamiltonian H_N^- for a $(N-1)$-spin system by removing the last spin,

$$-H_N^-(\boldsymbol{\rho}) = \frac{\beta_-}{\sqrt{N-1}} \sum_{i<j\leq N-1} g_{ij}\sigma_i\sigma_j + h \sum_{i\leq N-1} \sigma_i \tag{1.4}$$

for $\boldsymbol{\rho} = (\sigma_1, \ldots, \sigma_{N-1}) \in \Sigma_{N-1}$, where $\beta_- = \sqrt{(N-1)/N}\beta$. Directly we can write

$$-H_N(\boldsymbol{\sigma}) = -H_N^-(\boldsymbol{\rho}) + \sigma_N\left(\beta\iota_N + h\right), \quad \boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N),$$

where ι_N is defined as

$$\iota_N = \frac{1}{\sqrt{N}} \sum_{j\leq N-1} g_{Nj}\sigma_j. \tag{1.5}$$

A straightforward computation yields that

$$\begin{aligned}\langle\sigma_N\rangle &= \frac{\sum_{\boldsymbol{\rho}\in\Sigma_{N-1}}\sum_{\sigma_N\in\Sigma_1}\sigma_N\exp\left(-H_N^-(\boldsymbol{\rho}) + \sigma_N\left(\beta\iota_N + h\right)\right)}{\sum_{\boldsymbol{\rho}\in\Sigma_{N-1}}\sum_{\sigma_N\in\Sigma_1}\exp\left(-H_N^-(\boldsymbol{\rho}) + \sigma_N\left(\beta\iota_N + h\right)\right)} \\ &= \frac{\sum_{\boldsymbol{\rho}\in\Sigma_{N-1}}\exp\left(-H_N^-(\boldsymbol{\rho})\right)\cosh\left(\beta\iota_N + h\right)}{\sum_{\boldsymbol{\rho}\in\Sigma_{N-1}}\exp\left(-H_N^-(\boldsymbol{\rho})\right)\sinh\left(\beta\iota_N + h\right)}.\end{aligned}$$

Then the following fundamental identity holds

$$\langle\sigma_N\rangle = \frac{\langle\sinh(\beta\iota_N + h)\rangle_-}{\langle\cosh(\beta\iota_N + h)\rangle_-} = \langle\tanh(\beta\iota_N + h)\rangle, \tag{1.6}$$

where $\langle\cdot\rangle_-$ denotes the Gibbs average corresponding to the Hamiltonian H_N^-. Under $\langle\cdot\rangle_-$, we call ι_N the *cavity* field and under $\langle\cdot\rangle$, we call ι_N the *local* field. Now, to prove the TAP equations, we are led to the study of the limit laws for cavity and local fields. The key observation to establish the limit law for the cavity field is that the disorder $\{g_{Nj}\}_{j\leq N-1}$ in ι_N is independent of the disorder

$$\{g_{ij} : 1 \leq i < j \leq N-1\}$$

in the definition of $\langle\cdot\rangle_-$, which motivates the use of Gaussian interpolation on the cavity field. Consequently, from the CLT of the cavity field, we deduce the limit law for the local field. In both cases, the relevant quantitative moment estimates will be stated in the following section.

1.2 Main Results

For convenience, we first set up some definitions and notations that remain in force throughout this paper. We set ϕ_{μ,σ^2} to be the Gaussian density with mean μ and variance σ^2. Suppose that $U : \mathbb{R}^d \to \mathbb{R}$ is continuous. We say that U is of moderate growth provided $\lim_{|\mathbf{x}|\to\infty} U(\mathbf{x})\exp\left(-a|\mathbf{x}|^2\right) = 0$ for every $a > 0$.

1.2.1 Limit Law for the Cavity Field

Suppose that $g_1, \dots, g_N$ are i.i.d. standard Gaussian r.v.s independent of the randomness of $\{g_{ij}\}_{i<j\leq N}$. Define the *cavity field* ι with respect to the Gibbs average $\langle\cdot\rangle$ by

$$\iota = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} g_j \sigma_j. \tag{1.7}$$

The name "cavity field" is due to the important role of ι played in the cavity method as we have already seen in Section 1.1. Let us emphasize that ι is the cavity field of the N-spin system corresponding to $\langle\cdot\rangle$, while from the definition (1.5), ι_N is the cavity field of the $(N-1)$-spin system corresponding to $\langle\cdot\rangle_-$. If we replace N by $N+1$, then ι_{N+1} is the cavity field of the N-spin system and its form

$$\iota_{N+1} = \frac{1}{\sqrt{N+1}} \sum_{j=1}^{N} g_{(N+1)j}\sigma_j = \sqrt{\frac{N}{N+1}} \cdot \frac{1}{\sqrt{N}} \sum_{j=1}^{N} g_{(N+1)j}\sigma_j$$

is only different from ι in distribution by a factor $\sqrt{N/(N+1)}$ that converges to 1 and that will not affect the asymptotic behavior of ι and ι_{N+1} as N tends to infinity.

Note that the Gibbs average of ι is

$$r = \langle\iota\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} g_j \langle\sigma_j\rangle. \tag{1.8}$$

The limit law of the centered distribution $\iota - r$ under the Gibbs measure was first studied by Talagrand and it is approximately a centered Gaussian distribution with variance $1-q$. The exact quantitative result is stated as follows.

Theorem 1.1 ([4], Theorem 1.7.11)**.** *Let $\beta_0 < 1/2$ and $k \in \mathbb{N}$. Suppose that U is an infinitely differentiable function defined on $\mathbb{R}$ and the derivatives of all orders of U are of moderate growth. Then there exists a constant K depending only on k, U, and β_0 such that for all N, $\beta \leq \beta_0$, and h,*

$$E\left[\langle U(\iota - r)\rangle - \int_{\mathbb{R}} U(x)\phi_{0,1-q}(x)dx\right]^{2k} \leq \frac{K}{N^k}. \tag{1.9}$$

From this theorem, Talagrand then deduced the validity of the TAP equations by letting $U(x) = \exp(\beta x + h)$, see Theorem 2.4.20 in [4] and Theorem 1.7.7 in [5]. However, it seems very difficult to deduce the limit law for the local fields (defined by (1.11) below) from Theorem 1.1. To overcome this difficulty, it would be very helpful if we have good quantitative results for the limit law of ι, which is also one of the research problems proposed by Talagrand ([5], Research Problem 1.7.12). In our study, using Gaussian interpolation technique, we prove that the CLT holds for the cavity field by the following quantitative moment estimates:

Theorem 1.2. *Let $\beta_0 < 1/2$ and $k \in \mathbb{N}$. Suppose that U is an infinitely differentiable function defined on $\mathbb{R}$ and the derivatives of all orders of U are of moderate growth. Then there exists a constant K depending only on k, U, and β_0 such that for all N, $\beta \leq \beta_0$, and h,*

$$E\left[\langle U(\iota)\rangle - \int_{\mathbb{R}} U(x)\phi_{r,1-q}(x)dx\right]^{2k} \leq \frac{K}{N^k}. \tag{1.10}$$

In other words, the limit law of the cavity field is Gaussian with mean r and variance $1-q$. Early study of the CLT for the cavity field was discussed in Talagrand's book, see Section 1.5 [5] and page 87 [4], where he explained informally why the cavity field is Gaussian, but it is by no means obvious to use his arguments to obtain precise quantitative results such as Theorem 1.2. Later, using Stein's method, Chatterjee [1] obtained the first quantitative result: when $k=1$ and U is a bounded measurable function U, the left-hand side of (1.10) has an error bound $c(\beta_0)\|U\|_\infty/\sqrt{N}$, where $c(\beta_0)$ is a constant depending only on β_0. In our case, under more restrictive conditions on U we find moment estimates of all orders, which are going to be crucial in studying the CLT for local fields. In addition, as will be shown in Corollary 1.4 below, these estimates also help proving that the TAP equations hold simultaneously.

1.2.2 Limit Law for the Local Fields

For a fixed site $1 \leq i \leq N$, we define the local field ι_i at site i by

$$\iota_i = \frac{1}{\sqrt{N}} \sum_{j \leq N, j \neq i} g_{ij}\sigma_j. \tag{1.11}$$

Following a beautiful idea of Chatterjee [1], for $1 \leq i \leq N$, we define ν_i to be a random probability measure, whose density is the mixture of two Gaussian densities

$$p_i \phi_{\gamma_i+\beta(1-q),1-q} + (1-p_i)\phi_{\gamma_i-\beta(1-q),1-q}, \tag{1.12}$$

where

$$\gamma_i = \frac{1}{\sqrt{N}} \sum_{j \leq N, j \neq i} g_{ij} \langle \sigma_j \rangle - \beta(1-q) \langle \sigma_i \rangle, \tag{1.13}$$

$$p_i = \frac{e^{\beta\gamma_i + h}}{e^{\beta\gamma_i+h} + e^{-\beta\gamma_i - h}}.$$

Then we prove that the local field ι_i under the Gibbs measure is close to ν_i in the following sense:

Theorem 1.3. *Let $\beta_0 < 1/2$ and $k \in \mathbb{N}$. Suppose that U is an infinitely differentiable function defined on $\mathbb{R}$ and the derivatives of all orders of U are of moderate growth. Then there exists a constant K depending only on k, U, and β_0 such that for all N, $1 \leq i \leq N$, $\beta \leq \beta_0$, and h,*

$$E\left[\langle U(\iota_i)\rangle - \int U(x)\nu_i(dx)\right]^{2k} \leq \frac{K}{N^k}. \tag{1.14}$$

Again, by applying Stein's method, Chatterjee [1] proved the first quantitative result regarding the limit law for the local fields: when $k=1$ and U is a bounded measurable function U, the left-hand side of (1.14) has an error bound $c(\beta_0)\|U\|_\infty/\sqrt{N}$, where $c(\beta_0)$ is a constant depending on β_0 only. In our case, the smoothness of U allows us to obtain moment estimates of all orders. In particular, setting $U(x) = \tanh(\beta x + h)$, we obtain the same quantitative result for the TAP equations as Theorem 1.7.7 and Corollary 1.7.8 in Talagrand [5] :

Corollary 1.4. *Let $\beta_0 < 1/2$ and $k \in \mathbb{N}$. Then there exists a constant K depending only on k, U, and β_0 such that for all N, $1 \leq i \leq N$, $\beta \leq \beta_0$, and h,*

$$E\left[\langle \sigma_i \rangle - \tanh\left(\frac{\beta}{\sqrt{N}} \sum_{j \leq N, j \neq i} g_{ij} \langle \sigma_j \rangle + h - \beta^2(1-q)\langle \sigma_i \rangle\right)\right]^{2k} \leq \frac{K}{N^k}. \tag{1.15}$$

In addition, suppose that $0 < \varepsilon < 1/2$. *Then there is a constant* K' *depending only on* β_0 *and* ε *such that*

$$E \max_{1\le i\le N} \left| \langle \sigma_i \rangle - \tanh\left(\frac{\beta}{\sqrt{N}} \sum_{j\le N, j\ne i} g_{ij} \langle \sigma_j \rangle + h - \beta^2(1-q) \langle \sigma_i \rangle \right) \right| \le \frac{K'}{N^{1/2-\varepsilon}} \tag{1.16}$$

for every N, $\beta \le \beta_0$, *and* h.

Proof. Let us notice a useful formula from Chatterjee $[1, (9)]$: Suppose that X is a r.v., whose density is the mixture of two Gaussian densities: $p\phi_{\mu_1,\sigma^2} + (1-p)\phi_{\mu_2,\sigma^2}$ with $\mu_1 > \mu_2$, $\sigma > 0$, and $0 < p < 1$. Set

$$a = \frac{\mu_1 - \mu_2}{2\sigma^2},\ b = \frac{1}{2}\log\frac{p}{1-p} - \frac{\mu_1^2 - \mu_2^2}{4\sigma^2}.$$

Then we have the identity,

$$E\left[\tanh(aX+b)\right] = \tanh(aE\left[X\right] + b - (2p-1)a^2\sigma^2). \tag{1.17}$$

In particular, consider a r.v. X whose density is given by (1.12). It follows that $E\left[X\right] = \gamma_i$ and from (1.15),

$$\int \tanh(\beta x + h)\nu_i(dx) = E\left[\tanh(\beta X + h)\right] = \tanh\left(\beta\gamma_i + h - \beta^2(1-q)\langle\sigma_i\rangle\right).$$

Applying Theorem 1.3, (1.15) follows. Note that though in this case we apply $U(x) = \tanh(\beta x + h)$, a function depending on h, to Theorem 1.3, the constant K still does not depend on h, which can be verified by going through the proof for Theorem 1.3 and using the uniform boundedness of $U(x) = \tanh(\beta x + h)$. For the proof of (1.16), it follows by using (1.15) and Corollary 1.7.8. in [5]. □

2 Proofs

This section is devoted to proving Theorems 1.2 and 1.3. Before we proceed to prove our main results, let us define some crucial quantities that will be used in our study. Consider the replica $\boldsymbol{\sigma}^\ell$ and define $\dot{\sigma}_j^\ell = \sigma_j^\ell - \langle\sigma_j\rangle$ for $1 \le j \le M$. Define

$$T_\ell = \frac{1}{N}\sum_{j\le N} \dot{\sigma}_j^\ell \langle\sigma_j\rangle, \quad T_{\ell,\ell} = \frac{1}{N}\sum_{j\le N}(\dot{\sigma}_j^\ell)^2 - (1-q), \quad T_{\ell,\ell'} = \frac{1}{N}\sum_{j\le N}\dot{\sigma}_j^\ell\dot{\sigma}_j^{\ell'},\ \ell \ne \ell'. \tag{2.1}$$

By using replicas, these quantities can be controlled through the overlap (1.3) (see Section 1.10 of [5] for details): For fixed $\beta_0 < 1/2$ and $k \in \mathbb{N}$, there exists some K depending only on β_0 and k such that for any $\beta \le \beta_0$ and h,

$$\max_{1\le\ell,\ell'\le N,\ell\ne\ell'} \left\{ E\left\langle |T_\ell|^{2k}\right\rangle,\ E\left\langle |T_{\ell,\ell}|^{2k}\right\rangle,\ E\left\langle |T_{\ell,\ell'}|^{2k}\right\rangle \right\} \le \frac{K}{N^k}. \tag{2.2}$$

Note that in the following we use E_ζ to denote the expectation with respect to the randomness of ζ when ζ is a random variable.

2.1 Proof of Theorem 1.2

Using replicas, we set for $1 \le \ell \le 2k$,

$$\iota^\ell = \frac{1}{\sqrt{N}}\sum_{j=1}^N g_j\sigma_j^\ell. \tag{2.3}$$

Suppose that $\xi, \xi^1, \ldots, \xi^{2k}$ are i.i.d. Gaussian r.v.s with mean zero and variance $1-q$ and they are independent of $\{g_j\}_{j\le N}$ and $\{g_{ij}\}_{i<j\le N}$. Recalling definitions (1.7), (1.8), and (2.3), we consider the Gaussian interpolations,

$$\begin{aligned} u(t) &= \sqrt{t}(\iota - r) + \sqrt{1-t}\xi, \\ u_\ell(t) &= \sqrt{t}(\iota^\ell - r) + \sqrt{1-t}\xi^\ell, \ 1 \le \ell \le 2k. \end{aligned} \tag{2.4}$$

Suppose that U is a smooth function defined on $\mathbb{R}$ and its derivatives of all orders are of moderate growth. Define

$$V(x,y) = U(x+y) - E_\xi \left[U(\xi + y)\right] \tag{2.5}$$

and

$$\psi(t) = E\left\langle \prod_{\ell \le 2k} V(u_\ell(t), r)\right\rangle. \tag{2.6}$$

Notice that $\psi(0) = 0$ and $\psi(1)$ equals the left-hand side of (1.10). Let us explain the major difficulty that we will encounter in studying the CLT for the cavity field. Recall that Talagrand's main idea to prove the CLT for the centered cavity field in Theorem 1.1 is to study the following function:

$$\hat{\psi}(t) := E\left\langle \prod_{\ell \le 2k} \hat{V}\left(u_\ell(t)\right)\right\rangle, \quad t \in [0,1], \tag{2.7}$$

where $\hat{V}(x) := U(x) - E_\xi U(\xi)$. Notice that from (2.6) and (2.7), ψ and $\hat{\psi}$ both use the same Gaussian interpolations u_ℓ, but the functions V and $\hat{V}$ in ψ and $\hat{\psi}$ are defined differently. By definition, $\hat{\psi}(1)$ gives the left-hand side of (1.9) and $\hat{\psi}(0) = 0$. Using Gaussian integration by parts, the derivative of $\hat{\psi}$ can be computed up to any order and further, $\hat{\psi}^{(n)}(0) = 0$ for $0 \le n < 2k$ and $\sup_{0\le t\le 1} |\hat{\psi}^{(2k)}(t)| \le K/N^k$. Consequently, the mean value theorem yields (1.9) since

$$\hat{\psi}(1) = \sum_{n=0}^{2k-1} \frac{1}{n!}\hat{\psi}^{(n)}(0) + \frac{1}{(2k)!}\hat{\psi}^{(2k)}(c) = \frac{1}{(2k)!}\hat{\psi}^{(2k)}(c) \le \frac{K}{N^k}$$

for some $c \in (0,1)$. For more details, one may refer to the proof of Theorem 1.7.11 in [5]. In particular, the first derivative of $\hat{\psi}$ is given by

$$\begin{aligned} \hat{\psi}'(t) &= \frac{1}{2}\sum_{\ell\le 2k} E\left\langle T_{\ell,\ell}\frac{\partial^2 \hat{V}}{\partial x^2}(u_\ell(t)) \prod_{\ell'\le 2k, \ell'\ne \ell} \hat{V}(u_{\ell'}(t))\right\rangle \\ &\quad + \frac{1}{2}\sum_{\ell,\ell'\le 2k, \ell\ne\ell'} E\left\langle T_{\ell,\ell'}\frac{\partial \hat{V}}{\partial x}(u_\ell(t))\frac{\partial \hat{V}}{\partial x}(u_{\ell'}(t)) \prod_{\ell''\le 2k, \ell''\ne \ell,\ell'} \hat{V}(u_{\ell''}(t))\right\rangle. \end{aligned}$$

In our case, this situation is more complicated since the first derivative of ψ, as will be shown in Lemma 2.4 below, is given by

$$\psi'(t) = \frac{1}{2}\psi_0(t) + \frac{1}{2\sqrt{t}}\psi_1(t), \tag{2.8}$$

where

$$\begin{aligned} \psi_0(t) &= \sum_{\ell\le 2k} E\left\langle T_{\ell,\ell}\frac{\partial^2 V}{\partial x^2}(u_\ell(t), r) \prod_{\ell'\le 2k, \ell'\ne \ell} V(u_{\ell'}(t), r)\right\rangle \\ &\quad + \sum_{\ell,\ell'\le 2k, \ell\ne\ell'} E\left\langle T_{\ell,\ell'}\frac{\partial V}{\partial x}(u_\ell(t), r)\frac{\partial V}{\partial x}(u_{\ell'}(t), r) \prod_{\ell''\le 2k, \ell''\ne \ell,\ell'} V(u_{\ell''}(t), r)\right\rangle \end{aligned} \tag{2.9}$$

and

$$\psi_1(t) = \sum_{\ell \le 2k} E\left\langle T_\ell \frac{\partial^2 V}{\partial x \partial y}(u_\ell(t), r) \prod_{\ell' \le 2k, \ell' \ne \ell} V(u_{\ell'}(t), r) \right\rangle + \sum_{\ell, \ell' \le 2k, \ell \ne \ell'} E\left\langle T_\ell \frac{\partial V}{\partial x}(u_\ell(t), r) \frac{\partial V}{\partial y}(u_{\ell'}(t), r) \prod_{\ell'' \le 2k, \ell'' \ne \ell, \ell'} V(u_{\ell''}(t), r) \right\rangle. \tag{2.10}$$

One may find that from (2.9), the first part of the right-hand side of (2.8) is very similar to $\hat{\psi}'(t)$ and can be controlled by Talagrand's argument. Nonetheless, the main obstacles come from the second part of the right-hand side of (2.8). First, it makes the higher order derivatives of ψ even more complicated than that of $\hat{\psi}$. Second, it is not continuous at $t = 0$, which means that a direct application of mean value theorem to control $\psi(1)$ seems not applicable. To overcome these difficulties, instead, we will use Lemma 2.7 below.

We now state two calculus lemmas, whose proofs are defered to the appendix. The first lemma says that mild composition and integration of functions of moderate growth is still of moderate growth.

Lemma 2.1. *Suppose that $U, V : \mathbb{R} \to \mathbb{R}$ are of moderate growth. Define $U_0(x) = U(x) + V(x)$, $U_1(x) = U(x)^k$, $U_2(x) = U(rx)$, $U_3(x) = U(x + r)$, $U_4(x, y) = U(x + y)$, and $U_5(x) = E_\xi\left[U(\xi + x)\right]$ for $k \in \mathbb{N}$, $r \in \mathbb{R}$, and ξ a centered Gaussian r.v. with variance σ^2. Then all of the functions defined above are of moderate growth.*

The second lemma establishes the differentiability properties of ψ.

Lemma 2.2. *Let $U : \mathbb{R} \to \mathbb{R}$ be continuously differentiable. Suppose that U and its derivative U' are of moderate growth. Define $\psi(x) = E_\xi\left[U(\xi + x)\right]$ for ξ a centered Gaussian r.v. with variance σ^2. Then ψ is differentiable and $\psi'(x) = E_\xi\left[U'(\xi + x)\right]$.*

Recall that the constants K in the statements of Theorems 1.1, 1.2, and 1.3 are independent of N and $\beta \le \beta_0$. The main reason is because, when conditioning on the randomness of $\{g_{ij}\}_{i<j\le N}$, the cavity field and its Gibbs average are centered Gaussian distributions, whose variances are bounded above by some constants that are independent of $\{g_{ij}\}_{i<j\le N}$, N, and β. Therefore, we still have good control on the moment estimates of the cavity field and its Gibbs average. This observation will be used repeatedly and for convenience, we formulate it as Lemma 2.3.

Lemma 2.3. *Let $I, J \subset [0, \infty)$ be two bounded intervals. Suppose that K_1, K_2, K_3 are positive constants and that the following assumptions hold.*

i) *Let $z, g_1^N, \ldots, g_N^N$ be i.i.d. standard Gaussian r.v.s for $N \in \mathbb{N}$.*

ii) *For $N \in \mathbb{N}$, suppose that*

$$\left\{X_{j,\beta}^N : 1 \le j \le N, \beta \in J\right\}$$

is a family of random variables such that $|X_{j,\beta}^N| \le K_1$ for $1 \le j \le N$, and $\beta \in J$.

iii) *Let $f_1, f_2 : \mathbb{N} \times I \times J \to \mathbb{R}$ be measurable functions such that*

$$|f_1(N, t, \beta)| \le K_2/\sqrt{N}, \quad |f_2(N, t, \beta)| \le K_3,$$

for $(N, t, \beta) \in \mathbb{N} \times I \times J$.

iv) *Let $U : \mathbb{R} \to \mathbb{R}$ be a continuous function. Suppose that there are some $A > 0$ and some a with $0 < a < \min\{(4K_1^2K_2^2)^{-1}, (4K_3^2)^{-1}\}$ such that $|U(x)| \le A\exp(ax^2)$ for all $x \in \mathbb{R}$.*

Then there is a constant $K > 0$ such that

$$\sup_{N\in\mathbb{N},\beta\in J} E_0 \left[\sup_{t\in I}\left|U\left(f_1(N,t,\beta)\sum_{j\leq N} g_j^N X_{j,\beta}^N + f_2(N,t,\beta)z\right)\right|\right] \leq K, \tag{2.11}$$

where E_0 means the expectation with respect to $\{g_j^N : j \leq N,\, N \in \mathbb{N}\}$ and z.

Proof. From the given conditions, we obtain

$$\begin{aligned}
&\left(f_1(N,t,\beta)\sum_{j\leq N} g_j^N X_{j,\beta}^N + f_2(N,t,\beta)z\right)^2 \\
&\leq 2f_1(N,t,\beta)^2\left(\sum_{j\leq N} g_j^N X_{j,\beta}^N\right)^2 + 2f_2(N,t,\beta)^2 z^2 \\
&\leq 2K_2^2\left(\frac{1}{\sqrt{N}}\sum_{j\leq N} g_j^N X_{j,\beta}^N\right)^2 + 2K_3^2 z^2
\end{aligned}$$

and so

$$\begin{aligned}
&\sup_{t\in I}\left|U\left(f_1(N,t,\beta)\sum_{j\leq N} g_j^N X_{j,\beta}^N + f_2(N,t,\beta)z\right)\right| \\
&\leq A\exp\left(2aK_2^2\left(\frac{1}{\sqrt{N}}\sum_{j\leq N} g_j^N X_{j,\beta}^N\right)^2 + 2aK_3^2 z^2\right).
\end{aligned}$$

Since $N^{-1/2}\sum_{j\leq N} g_j^N X_{j,\beta}^N$ is a centered Gaussian r.v. with variance $N^{-1}\sum_{i\leq N}(X_{j,\beta}^N)^2 \leq K_1^2$, from the assumption iv), the left-hand side of (2.11) is then bounded above by

$$E_0\left[\exp\left(2aK_2^2 g^2\right)\right] E_0\left[\exp\left(2aK_3^2 z^2\right)\right] < \infty,$$

where g is a centered Gaussian r.v. with variance K_1^2. This completes our proof. □

In the sequel, we use E_0 to denote the expectation with respect to the randomness of $\{g_j\}_{j\leq N}$ and $\{\xi^\ell\}_{\ell\leq 2k}$. Recall formulas (1.8) and (2.4) for r and u_ℓ, respectively. The following lemma, as an application of Gaussian integration by parts, is our main equation to control the derivatives of all orders of ψ.

Lemma 2.4. *Let $k \in \mathbb{N}$. Suppose that $V_1, V_2, \ldots, V_{2k} : \mathbb{R}^2 \to \mathbb{R}$ are twice continuously differentiable functions and their first and second order partial derivatives are of moderate growth. Define*

$$\varphi(t) = E_0\left[\prod_{\ell\leq 2k} V_\ell(u_\ell(t), r)\right], \quad 0 \leq t \leq 1.$$

Then φ is differentiable on $(0,1)$ and

$$\varphi'(t) = \frac{1}{2}\varphi_0(t) + \frac{1}{2\sqrt{t}}\varphi_1(t), \tag{2.12}$$

where

$$\varphi_0(t) = \sum_{\ell \le 2k} T_{\ell,\ell} E_0 \left[\frac{\partial^2 V_i}{\partial x^2}(u_\ell(t), r) \prod_{\ell' \le 2k, \ell' \ne \ell} V_{\ell'}(u_{\ell'}(t), r) \right] + \sum_{\ell,\ell' \le 2k, \ell \ne \ell'} T_{\ell,\ell'} E_0 \left[\frac{\partial V_\ell}{\partial x}(u_\ell(t), r) \frac{\partial V_{\ell'}}{\partial x}(u_{\ell'}(t), r) \prod_{\ell'' \le 2k, \ell'' \ne \ell, \ell'} V_{\ell''}(u_{\ell''}(t), r) \right] \tag{2.13}$$

and

$$\varphi_1(t) = \sum_{\ell \le 2k} T_\ell E_0 \left[\frac{\partial^2 V_\ell}{\partial x \partial y}(u_\ell(t), r) \prod_{\ell' \le 2k, \ell' \ne \ell} V_{\ell'}(u_{\ell'}(t), r) \right] + \sum_{\ell,\ell' \le 2k, \ell \ne \ell'} T_\ell E_0 \left[\frac{\partial V_\ell}{\partial x}(u_\ell(t), r) \frac{\partial V_{\ell'}}{\partial y}(u_{\ell'}(t), r) \prod_{\ell'' \le 2k, \ell'' \ne \ell, \ell'} V_{\ell''}(u_{\ell''}(t), r) \right]. \tag{2.14}$$

Here, $\frac{\partial^{a+b} V_\ell}{\partial^a x \partial^b y}$ means that we differentiate V_ℓ with respect to the first variable a times and with respect to the second variable b times.

An immediate observation from (2.13) and (2.14) leads to the following remarks, that will be very useful when we control the derivatives of all order of ψ :

Remark 2.5. *In* (2.13), *the partial derivatives are only with respect to the x variable; the partial derivative with respect to the y variable occurs in each term of* (2.14).

Remark 2.6. *Formula* (2.12) *implies that our computation on the derivative of φ can be completely determined by $T_{\ell,\ell}$, $T_{\ell,\ell'}$, and T_ℓ in the following manner. Each $T_{\ell,\ell}$ is associated with $\frac{\partial^2 V_\ell}{\partial x^2}(u_\ell(t), r)$; each $T_{\ell,\ell'}$ is associated with $\frac{\partial V_\ell}{\partial x}(u_\ell(t), r)\frac{\partial V_{\ell'}}{\partial x}(u_{\ell'}(t), r)$. As for each T_ℓ, it is associated with $\frac{\partial^2 V_\ell}{\partial x \partial y}(u_\ell(t), r)$ and $\frac{\partial V_\ell}{\partial x}(u_i(t), r)\frac{\partial V_{\ell'}}{\partial y}(u_{\ell'}(t), r)$ for $1 \le \ell' \le 2k$ with $\ell' \ne \ell$.*

Proof of Lemma 2.4. To prove the differentiability of φ, it suffices to prove, with the help of the mean value theorem and the dominated convergence theorem, that for $0 < \delta < 1/2$ and $1 \le \ell \le 2k$,

$$E_0 \left[\sup_{\delta \le t \le 1-\delta} |u'_\ell(t)| \sup_{0 \le t \le 1} \left| \frac{\partial V_\ell}{\partial x}(u_\ell(t), r) \right| \prod_{\ell' \le 2k, \ell' \ne \ell} \sup_{0 \le t \le 1} |V_{\ell'}(u_{\ell'}(t), r)| \right] \le K, \tag{2.15}$$

for some constant K. Note that

$$u'_\ell(t) = \frac{1}{2\sqrt{Nt}} \sum_{j \le N} g_j \dot{\sigma}_j^\ell - \frac{1}{2\sqrt{1-t}} \xi^\ell.$$

Since $\frac{\partial V_\ell}{\partial x}$ and V_ℓ are of moderate growth, for any $a > 0$, there exists some $A > 0$ such that

$$\left| \frac{V_\ell}{\partial x}(x, y) \right|, |V_{\ell'}(x, y)| \le A \exp(a(x^2 + y^2)), \quad (x, y) \in \mathbb{R}^2. \tag{2.16}$$

Set $\left(z, g_j^N\right) = \left((1-q)^{-1}\xi^\ell, g_j\right)$ and also let $\left(X_{j,\beta}^N, f_1(N, t, \beta), f_2(N, t, \beta), U(x)\right)$ be any one of the following vectors

$$\left(\dot{\sigma}_j^\ell, \frac{1}{2\sqrt{Nt}}, -\frac{1-q}{2\sqrt{1-t}}, x^{2k+1}\right),$$
$$\left(\dot{\sigma}_j^\ell, \frac{\sqrt{t}}{\sqrt{N}}, (1-q)\sqrt{1-t}, \exp((4k+2)ax^2)\right),$$
$$\left(\langle\sigma_j^\ell\rangle, \frac{1}{\sqrt{N}}, 0, \exp((4k+2)ax^2)\right).$$

Then by choosing a small enough and applying Lemma 2.3, there exists a constant K independent of β and N such that

$$E_0\left[\sup_{\delta\le t\le 1-\delta}|u_\ell'(t)|^{2k+1}\right] \le K,$$
$$E_0\left[\sup_{0\le t\le 1}\exp\left((4k+2)au_\ell(t)^2\right)\right] \le K,$$
$$E_0\left[\sup_{0\le t\le 1}\exp\left((4k+2)ar^2\right)\right] \le K.$$

Therefore, from (2.16) and Cauchy-Schwarz inequality,

$$E_0\left[\sup_{0\le t\le 1}\left|\frac{\partial V_\ell}{\partial x}(u_\ell(t),r)\right|^{2k+1}\right],\ E_0\left[\sup_{0\le t\le 1}|V_{\ell'}(u_{\ell'}(t),r)|^{2k+1}\right] \le K$$

and by Hölder's inequality, (2.15) holds.

To prove (2.12), we use Gaussian integration by parts,

$$\begin{aligned}
&\varphi'(t)\\
&= \sum_{\ell\le 2k} E_0\left[u_\ell'(t)\frac{\partial V_\ell}{\partial x}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V_{\ell'}(u_{\ell'}(t),r)\right]\\
&= \sum_{\ell\le 2k} E_0\left[u_\ell'(t)u_\ell(t)\right]E_0\left[\frac{\partial^2 V_\ell}{\partial x^2}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V_{\ell'}(u_{\ell'}(t),r)\right]\\
&+ \sum_{\ell,\ell'\le 2k,\ell\ne\ell'} E_0\left[u_\ell'(t)u_{\ell'}(t)\right]E_0\left[\frac{\partial V_\ell}{\partial x}(u_\ell(t),r)\frac{\partial V_{\ell'}}{\partial x}(u_{\ell'}(t),r)\prod_{\ell''\le 2k,\ell''\ne\ell,\ell'}V_h(u_{\ell''}(t),r)\right]\\
&+ \sum_{\ell\le 2k} E_0\left[u_\ell'(t)r\right]E_0\left[\frac{\partial^2 V_\ell}{\partial x\partial y}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V_{\ell'}(u_{\ell'}(t),r)\right]\\
&+ \sum_{\ell,\ell'\le 2k,\ell\ne\ell'} E_0\left[u_\ell'(t)r\right]E_0\left[\frac{\partial V_\ell}{\partial x}(u_\ell(t),r)\frac{\partial V_{\ell'}}{\partial y}(u_{\ell'}(t),r)\prod_{\ell''\le 2k,\ell''\ne\ell,\ell'}V_{\ell''}(u_{\ell''}(t),r)\right].
\end{aligned} \tag{2.17}$$

Recalling definition (2.1), a straightforward computation yields

$$\begin{aligned}
E_0\left[u_\ell'(t)u_\ell(t)\right] &= T_{\ell,\ell}/2,\\
E_0\left[u_\ell'(t)u_{\ell'}(t)\right] &= T_{\ell,\ell'}/2, \quad \text{for } \ell\ne\ell',\\
E_0\left[u_\ell'(t)r\right] &= T_\ell/2\sqrt{t}.
\end{aligned} \tag{2.18}$$

Combining (2.17) and (2.18) gives (2.12). □

Lemma 2.4 will be used iteratively up to some optimal order. Since on each iteration, equation (2.12) brings us many terms, we will finally obtain a huge number of summations. Therefore, in order to make our argument clearer, we formulate the following lemma.

Lemma 2.7. *Fix an integer $m > 0$ and let ψ and $\psi_{\mathbf{s}_n}$ be real-valued smooth functions defined on $[0,1]$ for every $\mathbf{s}_n = (s_n(1), \dots, s_n(n)) \in \{0,1\}^n$ with $1 \le n \le m+1$. Suppose that $\psi(0) = 0$ and $\psi_{\mathbf{s}_n}(0) = 0$ for every $\mathbf{s}_n \in \{0,1\}^n$ with $1 \le n < m$. If*

$$\psi'(t) = \frac{1}{2}\psi_{(0)}(t) + \frac{1}{2\sqrt{t}}\psi_{(1)}(t) \tag{2.19}$$

and

$$\psi'_{\mathbf{s}_n}(t) = \frac{1}{2}\psi_{(\mathbf{s}_n,0)}(t) + \frac{1}{2\sqrt{t}}\psi_{(\mathbf{s}_n,1)}(t), \tag{2.20}$$

for every $\mathbf{s}_n \in \{0,1\}^n$ with $1 \le n \le m$, then

$$\begin{aligned} \psi(t) = & \frac{1}{2^m}\sum_{\mathbf{s}_m}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-1}}\frac{1}{\prod_{\ell=1}^m t_\ell^{s_m(\ell)/2}}dt_m\dots dt_2dt_1\psi_{\mathbf{s}_m}(0) \\ & + \frac{1}{2^{m+1}}\sum_{\mathbf{s}_{m+1}}\int_0^t\int_0^{t_1}\cdots\int_0^{t_m}\frac{1}{\prod_{\ell=1}^{m+1} t_\ell^{s_{m+1}(\ell)/2}}\psi_{\mathbf{s}_{m+1}}(t_{m+1})dt_{m+1}\dots dt_2dt_1. \end{aligned} \tag{2.21}$$

Proof. It suffices to prove that

$$\psi(t) = \frac{1}{2^m}\sum_{\mathbf{s}_m}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-1}}\frac{1}{\prod_{\ell=1}^m t_\ell^{s_m(\ell)/2}}\psi_{\mathbf{s}_m}(t_m)dt_m\dots dt_2dt_1. \tag{2.22}$$

Indeed, if (2.22) holds, then (2.21) can be deduced by applying

$$\psi_{\mathbf{s}_m}(t_m) = \int_0^{t_m}\psi'_{\mathbf{s}_m}(t_{m+1})dt_{m+1} + \psi_{\mathbf{s}_m}(0)$$

and (2.20) to (2.22).

Let us prove (2.22) by induction on m. If $m = 1$, from (2.19), (2.22) holds clearly by

$$\psi(t) = \int_0^t \psi'(t_1)dt_1 + \psi(0) = \frac{1}{2}\int_0^t\left(\psi_{(0)}(t_1) + \frac{1}{\sqrt{t_1}}\psi_{(1)}(t_1)\right)dt_1.$$

Suppose that the announced result is true for $m-1 \ge 1$. Let ψ and $\psi_{\mathbf{s}_n}$ be real-valued smooth functions for every $\mathbf{s}_n \in \{0,1\}^n$ with $n \le m+1$ satisfying the assumptions of this lemma. Notice that from (2.20), we obtain

$$\begin{aligned} \psi_{\mathbf{s}_{m-1}}(t_{m-1}) &= \int_0^{t_{m-1}}\psi'_{\mathbf{s}_{m-1}}(t_m)dt_m + \psi_{\mathbf{s}_{m-1}}(0) \\ &= \frac{1}{2}\int_0^{t_{m-1}}\left(\psi_{(\mathbf{s}_{m-1},0)}(t_m) + \frac{1}{\sqrt{t_m}}\psi_{(\mathbf{s}_{m-1},1)}(t_m)\right)dt_m \end{aligned}$$

and also by induction hypothesis,

$$\psi(t) = \frac{1}{2^{m-1}}\sum_{\mathbf{s}_{m-1}}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-2}}\frac{1}{\prod_{\ell=1}^{m-1} t_\ell^{s_{m-1}(\ell)/2}}\psi_{\mathbf{s}_{m-1}}(t_{m-1})dt_{m-1}\dots dt_2dt_1.$$

Now (2.22) follows by combining last two equations together

$$\begin{aligned}\psi(t) &= \frac{1}{2^{m-1}}\sum_{\mathbf{s}_{m-1}}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-2}}\int_0^{t_{m-1}}\frac{1}{\prod_{\ell=1}^{m-1}t_\ell^{s_{m-1}(\ell)/2}}\\ &\qquad\times\left(\frac{1}{2}\psi_{(\mathbf{s}_{m-1},0)}(t_m)+\frac{1}{2\sqrt{t_m}}\psi_{(\mathbf{s}_{m-1},1)}(t_m)\right)dt_m dt_{m-1}\dots dt_2 dt_1\\ &= \frac{1}{2^m}\sum_{\mathbf{s}_m}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-1}}\frac{1}{\prod_{\ell=1}^{m}t_\ell^{s_m(\ell)/2}}\psi_{\mathbf{s}_m}(t_m)dt_{m-1}\dots dt_2 dt_1.\end{aligned}$$

□

Proof of Theorem 1.2 for $k=1$. Recall formulas (2.5) and (2.6) for V and ψ, respectively. From Lemmas 2.1 and 2.2, V is an infinitely differentiable function and the partial derivatives of all orders of V are of moderate growth. We also note that ψ is infinitely differentiable by applying the same argument as Lemma 2.4. Recall the definition of E_0 and use Fubini's theorem, we can write

$$\psi(t) = E\left\langle E_0\left[V(u_1(t),r)V(u_2(t),r)\right]\right\rangle.$$

Lemma 2.4 implies that

$$\psi'(t) = \frac{1}{2}\psi_{(0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(1)}(t),$$

where

$$\begin{aligned}\psi_{(0)}(t) &= E\left\langle T_{1,1}E_0\left[\frac{\partial^2 V}{\partial x^2}(u_1(t),r)V(u_2(t),r)\right]+T_{2,2}E_0\left[\frac{\partial^2 V}{\partial x^2}(u_2(t),r)V(u_1(t),r)\right]\right\rangle\\ &\quad + E\left\langle T_{1,2}E_0\left[\frac{\partial V}{\partial x}(u_1(t),r)\frac{\partial V}{\partial x}(u_2(t),r)\right]+T_{1,2}E_0\left[\frac{\partial V}{\partial x}(u_1(t),r)\frac{\partial V}{\partial x}(u_2(t),r)\right]\right\rangle\end{aligned}$$

and

$$\begin{aligned}\psi_{(1)}(t) &= E\left\langle T_1E_0\left[\frac{\partial^2 V}{\partial x\partial y}(u_1(t),r)V(u_2(t),r)\right]+T_2E_0\left[\frac{\partial^2 V}{\partial x\partial y}(u_2(t),r)V(u_1(t),r)\right]\right\rangle\\ &\quad + E\left\langle T_1E_0\left[\frac{\partial V}{\partial x}(u_1(t),r)\frac{\partial V}{\partial y}(u_2(t),r)\right]+T_2E_0\left[\frac{\partial V}{\partial x}(u_2(t),r)\frac{\partial V}{\partial y}(u_1(t),r)\right]\right\rangle.\end{aligned}$$

Since

$$E_0\left[\frac{\partial^2 V}{\partial x^2}(u_\ell(0),r)V(u_{\ell'}(0),r)\right] = E_0\left[U''(\xi^\ell+r)\right]E_0\left[U(\xi^{\ell'}+r)-E_\xi\left[U(\xi+r)\right]\right]=0$$

for $1\le \ell,\ell'\le 2$ with $\ell\neq\ell'$ and $\langle T_{1,2}\rangle = 0$, we obtain $\psi_{(0)}(0)=0$. On the other hand, since

$$\begin{aligned}E_0\left[\frac{\partial^2 V}{\partial x\partial y}(u_\ell(0),r)V(u_{\ell'}(0),r)\right] &= E_0\left[U''(\xi^\ell+r)\right]E_0\left[U(\xi^{\ell'}+r)-E_\xi\left[U(\xi+r)\right]\right]=0\\ E_0\left[\frac{\partial V}{\partial x}(u_\ell(0),r)\frac{\partial V}{\partial y}(u_{\ell'}(0),r)\right] &= E_0\left[U'(\xi^\ell+r)\right]E_0\left[U'(\xi^{\ell'}+r)-E_\xi\left[U'(\xi+r)\right]\right]=0\end{aligned}$$

for $1\le\ell,\ell'\le 2$ with $\ell\neq\ell'$, this implies $\psi_{(1)}(0)=0$. Applying Lemma 2.4 again to $\psi_{(0)}$ and $\psi_{(1)}$, we may write

$$\begin{aligned}\psi'_{(0)}(t) &= \frac{1}{2}\psi_{(0,0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(0,1)}(t)\\ \psi'_{(1)}(t) &= \frac{1}{2}\psi_{(1,0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(1,1)}(t)\end{aligned}\tag{2.23}$$

for four smooth functions $\psi_{(0,0)}, \psi_{(0,1)}, \psi_{(1,0)}, \psi_{(1,1)}$ on $[0,1]$. Here come the crucial observations: First, from Remark 2.5, the partial derivatives in the expression of $\psi_{(0)}$ are only respect to the x variable and the partial derivative with respect to the y variable occurs in every term in the expression of $\psi_{(1)}$. Second, from Remark 2.6,

(i) each $T_{\ell,\ell}$ is associated with $\frac{\partial^2}{\partial x^2}\Big|_{(u_\ell(t),r)}$ on V;

(ii) each $T_{1,2}$ is associated with $\frac{\partial}{\partial x}\Big|_{(u_1(t),r)} \cdot \frac{\partial}{\partial x}\Big|_{(u_2(t),r)}$ on V;

(iii) each T_ℓ is associated with $\frac{\partial^2}{\partial x \partial y}\Big|_{(u_\ell(t),r)}$ and $\frac{\partial}{\partial x}\Big|_{(u_\ell(t),r)} \cdot \frac{\partial}{\partial x}\Big|_{(u_{\ell'}(t),r)}$ on V for $\ell' \neq \ell$.

Based on these observations and our experience in obtaining $\psi_{(0)}$ and $\psi_{(1)}$, it should be clear that for each $\mathbf{s}_2 = (s_2(1), s_2(2)) \in \{0,1\}^2$, $\psi_{\mathbf{s}_2}$ is given by the summation of the terms that are of the form

$$E\left\langle T_{1,1}^{k_1(1)} T_{2,2}^{k_1(2)} T_{1,2}^{k_2(1,2)} T_{2,1}^{k_2(2,1)} T_1^{k_3(1)} T_2^{k_3(2)} \right. \\ \left. \cdot E_0\left[\frac{\partial^{k_4(1)+k_5(1)} V}{\partial x^{k_4(1)} \partial y^{k_5(1)}}(u_1(t), r) \frac{\partial^{k_4(2)+k_5(2)} V}{\partial x^{k_4(2)} \partial y^{k_5(2)}}(u_2(t), r)\right]\right\rangle \tag{2.24}$$

with

$$\begin{aligned} &k_1(1) + k_1(2) + k_2(1,2) + k_2(2,1) + k_3(1) + k_3(2) = 2 \\ &2k_1(\ell) + k_{1,2}(\ell) + k_{2,1}(\ell) + k_3(\ell) = k_4(\ell), \quad \ell = 1,2 \\ &k_3(1) + k_3(2) = s_2(1) + s_2(2) = k_5(1) + k_5(2) \end{aligned} \tag{2.25}$$

for some $k_1(\ell), k_2(\ell,\ell'), k_3(\ell)$, $k_4(\ell)$, and $k_5(\ell)$ nonnegative integers for $1 \leq \ell, \ell' \leq 2$ with $\ell \neq \ell'$.

In the following we claim that $\psi_{\mathbf{s}_2}(0) = 0$ for every $\mathbf{s}_2 \in \{0,1\}^2$ except for $\mathbf{s}_2 = (0,0)$ and

$$\psi_{(0,0)}(0) = E\left\langle \left(2T_{1,1}T_{2,2} + 4T_{1,2}^2\right) E_0\left[U''(\xi^1 + r)U''(\xi^2 + r)\right]\right\rangle. \tag{2.26}$$

To this end, let us first prove some properties for k_4. Assume that

$$\left\langle T_{1,1}^{k_1(1)} T_{2,2}^{k_1(2)} T_{1,2}^{k_2(1,2)} T_{2,1}^{k_2(2,1)} T_1^{k_3(1)} T_2^{k_3(2)} \right\rangle \neq 0. \tag{2.27}$$

If $k_4(1) = 1$, that is, the index 1 appears only once in the subscript of

$$T_{1,1}^{k_1(1)} T_{2,2}^{k_1(2)} T_{1,2}^{k_2(1,2)} T_{2,1}^{k_2(2,1)} T_1^{k_3(1)} T_2^{k_3(2)}.$$

From (2.25), $k_1(1) = 0$ and using the independence of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$, this yields that

1. if $(k_2(1,2), k_2(2,1), k_3(1)) = (1,0,0)$ or $(0,1,0)$, then the left-hand side of (2.27) equals

$$\left\langle T_{1,2} T_{2,2}^{k_1(2)} T_2^{k_3(2)} \right\rangle = \frac{1}{N}\sum_{j' \leq N} \left\langle \dot{\sigma}_{j'}^1 \right\rangle \left\langle \dot{\sigma}_{j'}^2 T_{2,2}^{k_1(2)} T_2^{k_3(2)} \right\rangle = 0;$$

2. if $(k_2(1,2), k_2(2,1), k_3(1)) = (0,0,1)$, then the left-hand side of (2.27) equals

$$\left\langle T_1 T_{2,2}^{k_1(2)} T_2^{k_3(2)} \right\rangle = \left(\frac{1}{N}\sum_{j' \leq N} \left\langle \dot{\sigma}_{j'}^1 \right\rangle \left\langle \sigma_{j'}^2 \right\rangle\right) \left\langle T_{2,2}^{k_1(2)} T_2^{k_3(2)} \right\rangle = 0.$$

Thus, if (2.27) occurs, then $k_4(1) \geq 2$ or $= 0$ and so is $k_4(2)$. Next, suppose that $t = 0$ and

$$E_0\left[\frac{\partial^{k_4(1)+k_5(1)}V}{\partial x^{k_4(1)}\partial y^{k_5(1)}}(u_1(0),r)\frac{\partial^{k_4(2)+k_5(2)}V}{\partial x^{k_4(2)}\partial y^{k_5(2)}}(u_2(0),r)\right] \neq 0. \tag{2.28}$$

We check that this will yield $k_4(1), k_4(2) \geq 1$. Indeed, if $k_4(1) = 0$, then the left-hand side of (2.28) becomes

$$\begin{aligned}
&E_0\left[\frac{\partial^{k_5(1)}V}{\partial y^{k_5(1)}}(u_1(0),r)\frac{\partial^{k_4(2)+k_5(2)}V}{\partial x^{k_4(2)}\partial y^{k_5(2)}}(u_2(0),r)\right]\\
&= E_0\left[U^{(k_5(1))}(\xi^1+r) - E_\xi\left[U^{(k_5(1))}(\xi+r)\right]\right]E_0\left[\frac{\partial^{k_4(2)+k_5(2)}V}{\partial x^{k_4(2)}\partial y^{k_5(2)}}(u_2(t),r)\right] = 0,
\end{aligned}$$

which contradicts to (2.28). So we conclude that $k_4(1) \geq 1$ and the same argument also implies $k_4(2) \geq 1$. To sum up, if (2.24) does not vanish, then both (2.27) and (2.28) must occur and hence, $k_4(1), k_4(2) \geq 2$. From (2.25), we then have

$$\begin{aligned}
4 &= 2\left(k_1(1) + k_1(2) + k_2(1,2) + k_2(2,1) + k_3(1) + k_3(2)\right)\\
&= k_4(1) + k_4(2) + k_3(1) + k_3(1)\\
&\geq 2 + 2 + k_3(1) + k_3(2)\\
&\geq 4
\end{aligned}$$

and this implies that $k_4(1) = k_4(2) = 2$ and $k_3(1) = k_3(2) = 0 = s_2(1) = s_2(2)$. Thus, we conclude that $\psi_{\mathbf{s}_2}(0) = 0$ for all $\mathbf{s}_2 \in \{0,1\}^2$ except for $\mathbf{s}_2 = (0,0)$ and $\psi_{(0,0)}(0)$ is the summation of the terms

$$E\left\langle T_{\ell_1,\ell_1'}T_{\ell_2,\ell_2'}E_0\left[U''(\xi^1+r)U''(\xi^2+r)\right]\right\rangle,$$

where the summation is over all $(\ell_1, \ell_1', \ell_2, \ell_2') \in \{1,2\}^4$ satisfying that each of the indices 1 and 2 appears exactly twice in the list $(\ell_1, \ell_1', \ell_2, \ell_2')$. Consequently, a simple computation yields (2.26) and this completes the proof of our claim.

Finally, again we may use Lemma 2.7 to write

$$\psi_{\mathbf{s}_2}'(t) = \frac{1}{2}\psi_{(s_2(1),s_2(1),0)}(t) + \frac{1}{2\sqrt{t}}\psi_{(s_2(1),s_2(1),1)}(t)$$

for $\mathbf{s}_2 \in \{0,1\}^2$, where $\{\psi_{\mathbf{s}_3} : \mathbf{s}_3 \in \{0,1\}^3\}$ are smooth functions. Then ψ and $\psi_{\mathbf{s}_n}$ for $\mathbf{s}_n \in \{0,1\}^n$ and $n = 1,2,3$ satisfy the assumption of Lemma 2.7 and it follows that

$$\begin{aligned}
\psi(t) &= \frac{t^2}{2^3}E\left\langle\left(2T_{1,1}T_{2,2} + 4T_{1,2}^2\right)E_0\left[U''(\xi^1+r)U''(\xi^2+r)\right]\right\rangle\\
&\quad + \frac{1}{2^3}\sum_{\mathbf{s}_3}\int_0^t\int_0^{t_1}\int_0^{t_2}\frac{1}{t_1^{s_3(1)/2}t_2^{s_3(2)/2}t_3^{s_3(3)/2}}\psi_{\mathbf{s}_3}(t_3)dt_3dt_2dt_1.
\end{aligned} \tag{2.29}$$

Since each term in the summation of $\psi_{\mathbf{s}_3}$ is of the form (2.24) with

$$k_1(1) + k_1(2) + k_2(1,2) + k_2(2,1) + k_3(1) + k_3(2) = 3,$$

by using (2.2), Hölder's inequality, and Lemma 2.3, it yields that $\sup_{0\leq t\leq 1}|\psi_{\mathbf{s}_3}(t)| \leq K/N^{3/2}$ for each $\mathbf{s}_3 \in \{0,1\}^3$ and from (2.29), we are done. □

Proof of Theorem 1.2 for the general value of k. As in the case $k = 1$, we write

$$\psi(t) = E\left\langle E_0\left[\prod_{\ell\leq 2k}V(u_\ell(t),r)\right]\right\rangle.$$

Now applying Lemma 2.4 to

$$E_0\left[\prod_{\ell\le 2k} V(u_\ell(t),r)\right]$$

and then taking expectation $E\langle\cdot\rangle$ on (2.12), we obtain

$$\psi'(t)=\frac{1}{2}\psi_{(0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(1)}(t),$$

where

$$\begin{aligned}\psi_{(0)}(t)=&\sum_{\ell\le 2k}E\left\langle T_{i,i}E_0\left[\frac{\partial^2 V}{\partial x^2}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V(u_{\ell'}(t),r)\right]\right\rangle\\&+\sum_{\ell,\ell'\le 2k,\ell\ne\ell'}E\left\langle T_{\ell,\ell'}E_0\left[\frac{\partial V}{\partial x}(u_\ell(t),r)\frac{\partial V}{\partial x}(u_{\ell'}(t),r)\prod_{\ell''\le 2k,\ell''\ne\ell,\ell''}V(u_{\ell''}(t),r)\right]\right\rangle\end{aligned}$$

and

$$\begin{aligned}\psi_{(1)}(t)=&\sum_{\ell\le 2k}E\left\langle T_{\ell}E_0\left[\frac{\partial^2 V}{\partial x\partial y}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V(u_{\ell'}(t),r)\right]\right\rangle\\&+\sum_{\ell,\ell'\le 2k,\ell\ne\ell'}E\left\langle T_{\ell}E_0\left[\frac{\partial V}{\partial x}(u_\ell(t),r)\frac{\partial V}{\partial y}(u_{\ell'}(t),r)\prod_{\ell''\le 2k,\ell''\ne\ell,\ell'}V(u_{\ell''}(t),r)\right]\right\rangle.\end{aligned}$$

Next, to compute the derivative of $\psi_{(0)}$, let us apply Lemma 2.4 again to

$$E_0\left[\frac{\partial^2 V}{\partial x^2}(u_\ell(t),r)\prod_{\ell'\le 2k,\ell'\ne\ell}V(u_{\ell'}(t),r)\right],\quad \ell\le 2k,$$

and

$$E_0\left[\frac{\partial V}{\partial x}(u_\ell(t),r)\frac{\partial V}{\partial x}(u_{\ell'}(t),r)\prod_{\ell''\le 2k,\ell''\ne\ell,\ell'}V(u_{\ell''}(t),r)\right],\quad \ell,\ell'\le 2k,\ \ell\ne\ell',$$

and then take expectation $E\langle\cdot\rangle$. Then we obtain two smooth functions $\psi_{(0,0)}$ and $\psi_{(0,1)}$ defined on $[0,1]$ such that

$$\psi'_{(0)}(t)=\frac{1}{2}\psi_{(0,0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(0,1)}(t).$$

Similarly, the derivative of $\psi'_{(1)}$ can also be computed in the same way, which leads to two smooth functions $\psi_{(1,0)}$ and $\psi_{(1,1)}$ defined on $[0,1]$ such that

$$\psi'_{(1)}(t)=\frac{1}{2}\psi_{(1,0)}(t)+\frac{1}{2\sqrt{t}}\psi_{(1,1)}(t).$$

Continuing this process, we get $\{\psi_{\mathbf{s}_n}:s_n\in\{0,1\}^n,n\le 2k+1\}$ for which (2.19) and (2.20) hold.

We claim that $\psi_{\mathbf{s}_n}(0)=0$ for every $\mathbf{s}_n=(s_n(1),\dots,s_n(n))\in\{0,1\}^n$ with $n<2k$. To see this, let us observe that from Remarks 2.5 and 2.6, a typical term in those complicated summations in the expression of $\psi_{\mathbf{s}_n}$ is of the form

$$E\left\langle\prod_{\ell\le 2k}T_{\ell,\ell}^{k_1(\ell)}\prod_{\ell,\ell'\le 2k,\ell\ne\ell'}T_{\ell,\ell'}^{k_2(\ell,\ell')}\prod_{\ell\le 2k}T_{\ell}^{k_3(\ell)}E_0\left[\prod_{\ell\le 2k}\frac{\partial^{k_4(\ell)+k_5(\ell)}V}{\partial x^{k_4(\ell)}\partial y^{k_5(\ell)}}(u_\ell(t),r)\right]\right\rangle,\quad (2.30)$$

with

$$\sum_{\ell\le 2k} k_1(\ell) + \sum_{\ell,\ell'\le 2k,\ell\neq\ell'} k_2(\ell,\ell') + \sum_{\ell\le 2k} k_3(\ell) = n,$$
$$2k_1(\ell) + \sum_{\ell'\le 2k,\ell'\neq\ell} (k_2(\ell,\ell') + k_2(\ell',\ell)) + k_3(\ell) = k_4(\ell),\ \ \ell\le 2k,$$
$$\sum_{\ell\le 2k} k_3(\ell) = \sum_{\ell\le n} s_n(\ell) = \sum_{\ell\le 2k} k_5(\ell),$$

where $k_1(\ell), k_2(\ell,\ell'), k_3(\ell), k_4(\ell)$, and $k_5(\ell)$ are nonnegative integers for $\ell,\ell'\le 2k$ with $\ell\neq\ell'$.

First of all, notice that if

$$\left\langle \prod_{\ell\le 2k} T_{\ell,\ell}^{k_1(\ell)} \prod_{\ell,\ell'\le 2k,\ell\neq\ell'} T_{\ell,\ell'}^{k_2(\ell,\ell')} \prod_{\ell\le 2k} T_{\ell}^{k_3(\ell)} \right\rangle \neq 0,$$

then for $\ell\le 2k$, either $k_4(\ell)\ge 2$ or it is equal to zero. That is, if ℓ occurs in one of the subscripts of $T_{\ell,\ell}$, $T_{\ell,\ell'}$ or T_ℓ, it must occur more than once and we suppose that this is the case. Second, if

$$E_0\left[\prod_{\ell\le 2k} \frac{\partial^{k_4(\ell)+k_5(\ell)} V}{\partial x^{k_4(\ell)}\partial y^{k_5(\ell)}}(u_\ell(0),r)\right] \neq 0,$$

then $k_4(\ell)\ge 1$ for every $\ell\le 2k$ since $E_{\xi_\ell}\left[\frac{\partial^a V}{\partial y^a}(u_\ell(0),r)\right] = 0$ for all $a\ge 0$ and $\ell\le 2k$. Therefore, we conclude that $k_4(\ell)\ge 2$ for every $\ell\le 2k$ and it implies

$$\begin{aligned}
2n &= 2\sum_{\ell\le 2k} k_1(\ell) + 2\sum_{\ell,\ell'\le 2k,\ell\neq\ell'} k_2(\ell,\ell') + 2\sum_{\ell\le 2k} k_3(\ell)\\
&= \sum_{\ell\le 2k}\left[2k_1(\ell) + \sum_{\ell'\le 2k,\ell'\neq\ell}(k_2(\ell,\ell')+k_2(\ell',\ell)) + k_3(\ell)\right] + \sum_{\ell\le 2k} k_3(\ell)\\
&= \sum_{\ell\le 2k} k_4(\ell) + \sum_{\ell\le 2k} k_3(\ell)\\
&\ge 2k\cdot 2 + \sum_{\ell\le 2k} k_3(\ell)\\
&\ge 2k\cdot 2\\
&= 4k.
\end{aligned}$$

Hence, if (2.30) is not equal to zero, then $n\ge 2k$. So $\psi_{\mathbf{s}_n}(0)=0$ for every $\mathbf{s}_n\in\{0,1\}^n$ with $n<2k$, which completes the proof of our claim. In addition, we can conclude more from above that if $n=2k$ and (2.30) does not vanish, since $\sum_{\ell\le 2k} k_3(\ell) = \sum_{\ell\le n} s_n(\ell)$ and $k_4(\ell)\ge 2$ for every $\ell\le 2k$, it implies that $s_n(\ell)=0$ and $k_4(\ell)=2$ for every $\ell\le 2k$. Consequently, this means that $\psi_{\mathbf{s}_{2k}}(0)=0$ for every $\mathbf{s}_{2k}\in\{0,1\}^{2k}$ unless $\mathbf{s}_{2k}=\mathbf{0}_{2k} := (0,\ldots,0)$ and in this case,

$$\psi_{\mathbf{0}_{2k}}(0) = \sum_{\ell_1,\ell_1',\ldots,\ell_{2k},\ell_{2k}'} E\left\langle T_{\ell_1,\ell_1'}\cdots T_{\ell_{2k},\ell_{2k}'} E_0\left[\prod_{\ell\le 2k} U''(\xi^\ell + r)\right]\right\rangle,$$

where $(\ell_1,\ell_1',\ldots,\ell_{2k},\ell_{2k}')\in\{1,\ldots,2k\}^{4k}$ satisfies that each number $\ell\le 2k$ occurs exactly twice in this vector.

Now, concluding from (2.21) in Lemma 2.7 and using

$$\int_0^t \int_0^{t_1} \cdots \int_0^{t_{2k-1}} dt_{2k} \dots dt_2 dt_1 = \frac{t^{2k}}{(2k)!},$$

we obtain

$$\psi(t) = \frac{t^{2k}}{2^{2k}(2k)!} \sum_{\ell_1,\ell_1',\dots,\ell_{2k},\ell_{2k}'} E\left\langle T_{\ell_1,\ell_1'} \cdots T_{\ell_{2k},\ell_{2k}'} E_0\left[\prod_{\ell\le 2k} U''(\xi^\ell + r)\right]\right\rangle$$
$$+ \frac{1}{2^{2k+1}} \sum_{\mathbf{s}_{2k+1}} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{2k}} \frac{1}{\prod_{\ell=1}^{2k+1} t_\ell^{s_{2k+1}(\ell)/2}} \psi_{\mathbf{s}_{2k+1}}(t_{2k+1}) dt_{2k+1} \dots dt_2 dt_1.$$

Finally, since each term in those summations in the expression of $\psi_{\mathbf{s}_{2k+1}}$ is given by formula (2.30) with $n = 2k+1$, by applying Lemma 2.3, the known result (2.2), and Hölder's inequality, we obtain some $K > 0$ depending on β_0, k, and U only such that $\sup_{0\le t\le 1} |\psi_{\mathbf{s}_{2k+1}}(t)| \le K/N^{k+1/2}$ for every $\beta \le \beta_0$ and h. Similarly we also have

$$\left| E\left\langle T_{\ell_1,\ell_1'} \cdots T_{\ell_{2k},\ell_{2k}'} E_0\left[\prod_{\ell\le 2k} U''(\xi^\ell + r)\right]\right\rangle\right| \le \frac{K}{N^k}.$$

Therefore, $\psi(1) \le K/N^k$ and we are done. □

2.2 Proof of Theorem 1.3

The following proposition is the key to proving Theorem 1.3.

Proposition 2.8. *Let $\beta_0 < 1/2$ and $k \in \mathbb{N}$. Suppose that U is an infinitely differentiable function defined on $\mathbb{R}$ and the derivatives of all orders of U are of moderate growth. Recall ι and r as defined by* (1.7) *and* (1.8)*. Then for any $\beta \le \beta_0$ and h,*

$$E\left[\frac{\langle U(\iota)\cosh(\beta\iota + h)\rangle}{\langle\cosh(\beta\iota + h)\rangle} - \frac{E_\xi\left[U(\xi + r)\cosh(\beta(\xi + r) + h)\right]}{\exp\left(\frac{\beta^2}{2}(1-q)\right)\cosh(\beta r + h)}\right]^{2k} \le \frac{K}{N^k},$$

where ξ is a centered Gaussian distribution with variance $1-q$ and K depends on β_0, k and U only.

Proof. Define for $\varepsilon = \pm 1$,

$$A(\varepsilon) = \langle U(\iota)\exp(\varepsilon\beta\iota)\rangle - E_\xi\left[U(\xi + r)\exp(\varepsilon\beta(\xi + r))\right]$$

and also

$$B(\varepsilon) = \langle\exp(\varepsilon\beta\iota)\rangle - E_\xi\left[\exp(\varepsilon\beta(\xi + r))\right].$$

Notice that both $U(x)\exp\varepsilon\beta x$ and $\exp\varepsilon\beta x$ are infinitely differentiable and their derivatives of all orders are of moderate growth. Applying Theorem 1.2 to these two functions, we obtain that $EA(\varepsilon)^{4k} \le K/N^{2k}$ and $EB(\varepsilon)^{8k} \le K/N^{4k}$. Now using Jensen's inequality, Hölder's inequality, and Lemma 2.3, implies that

$$E\left[\frac{1}{\langle\exp(\varepsilon\beta\iota)\rangle^{4k}}\right] \le E\left[\exp(-4k\varepsilon\beta r)\right] \le K, \tag{2.31}$$

$$E\left[\frac{1}{\exp(\varepsilon\beta r)^{8k}}\right] = E\left[\exp(-8k\varepsilon\beta r)\right] \le K, \tag{2.32}$$

and

$$E\left[U(\xi+r)^{4k}\frac{\exp(4k\varepsilon\beta(\xi+r))}{\langle\exp(\varepsilon\beta\iota)\rangle^{4k}}\right]\le E\left[U(\xi+r)^{4k}\exp(4k\varepsilon\beta\xi)\right]\le K. \tag{2.33}$$

For convenience, we set

$$\begin{aligned}
A &= \langle U(\iota)\cosh(\beta\iota+h)\rangle\,,\\
B &= \langle\cosh(\beta\iota+h)\rangle\,,\\
A' &= E_\xi\left[U(\xi+r)\cosh(\beta(\xi+r)+h)\right],\\
B' &= \exp\left(\frac{\beta^2}{2}(1-q)\right)\cosh(\beta r+h).
\end{aligned}$$

Consequently, by using Hölder's inequality, (2.31), (2.32), and (2.33), the following three inequalities hold

$$\begin{aligned}
E\left[\frac{A-A'}{B}\right]^{2k} &= E\left[\frac{A(1)e^h+A(-1)e^{-h}}{\langle\exp(\beta\iota)\rangle e^h+\langle\exp(-\beta\iota)\rangle e^{-h}}\right]^{2k}\\
&\le 2^{2k}\left(E\left[\frac{A(1)}{\langle\exp(\beta\iota)\rangle}\right]^{2k}+E\left[\frac{A(-1)}{\langle\exp(-\beta\iota)\rangle}\right]^{2k}\right)\\
&\le 2^{2k}\Bigg(\left(EA(1)^{4k}\right)^{1/2}\left(E\left[\frac{1}{\langle\exp(\beta\iota)\rangle^{4k}}\right]\right)^{1/2}\\
&\quad+\left(EA(-1)^{4k}\right)^{1/2}\left(E\left[\frac{1}{\langle\exp(-\beta\iota)\rangle^{4k}}\right]\right)^{1/2}\Bigg)\\
&\le\frac{K}{N^k},
\end{aligned} \tag{2.34}$$

$$\begin{aligned}
E\left[\frac{B'-B}{B'}\right]^{4k} &\le E\left[\frac{B(1)e^h+B(-1)e^{-h}}{\exp(\beta r)e^h+\exp(-\beta r)e^{-h}}\right]^{4k}\\
&\le 2^{4k}\left(E\left[\frac{B(1)}{\exp(\beta r)}\right]^{4k}+E\left[\frac{B(-1)}{\exp(-\beta r)}\right]^{4k}\right)\\
&\le 2^{4k}\Bigg(\left(EB(1)^{8k}\right)^{1/2}\left(E\left[\frac{1}{\exp(8k\beta r)}\right]\right)^{1/2}\\
&\quad+\left(EB(-1)^{8k}\right)^{1/2}\left(E\left[\frac{1}{\exp(-8k\beta r)}\right]\right)^{1/2}\Bigg)\\
&\le\frac{K}{N^{2k}},
\end{aligned} \tag{2.35}$$

and

$$\begin{aligned}
E\left[\frac{A'}{B}\right]^{4k} &= E\left[\frac{E_\xi\left[U(\xi+r)\left(\exp(\beta(\xi+r))e^h+\exp(-\beta(\xi+r))e^{-h}\right)\right]}{\langle\exp(\beta\iota)e^h+\exp(-\beta\iota)e^{-h}\rangle}\right]^{4k}\\
&\le 2^{4k}\left(E\left[U(\xi+r)\frac{\exp(\beta(\xi+r))}{\langle\exp(\beta\iota)\rangle}\right]^{4k}+E\left[U(\xi+r)\frac{\exp(-\beta(\xi+r))}{\langle\exp(-\beta\iota)\rangle}\right]^{4k}\right)\\
&\le K.
\end{aligned} \tag{2.36}$$

Finally, by applying (2.34), (2.35), and (2.36) to the following inequality

$$\left|\frac{A}{B}-\frac{A'}{B'}\right|^{2k} \le 2^{2k}\left(\left|\frac{A-A'}{B}\right|^{2k}+\left|\frac{A'}{B}\right|^{2k}\left|\frac{B'-B}{B'}\right|^{2k}\right),$$

and using Hölder's inequality, the announced result follows. □

Recall that q is defined by (1.2). We also define q_- as the unique solution of $q_- = E\tanh^2(\beta_- z\sqrt{q_-}+h)$, where $\beta_- = \sqrt{(N-1)/N}\beta$ and z is a standard Gaussian distribution. Notice that the existence and uniqueness of q_- are always guaranteed since we only consider the high temperature region, that is, $\beta_- < 1/2$. Also, recall the quantity γ_i from (1.13).

Lemma 2.9. *There is a constant $L>0$ so that*

$$|q-q_-| \le \frac{L}{N} \tag{2.37}$$

for every $\beta<1/2$, h, and N. Let $\beta_0<1/2$ be fixed. Then for every $1\le i\le N$, $\beta\le\beta_0$, and h,

$$E\left[\gamma_i - \frac{1}{\sqrt{N-1}}\sum_{j\le N, j\ne i} g_{ij}\left\langle\sigma_j\right\rangle_-\right]^{2k} \le \frac{K}{N^k} \tag{2.38}$$

where K is a constant depending only on β_0 and k.

Proof. The inequality (2.37) is from Lemma 1.7.5. [5], while (2.38) follows from the inequalities on page 86 of [5]. □

Proof of Theorem 1.3. By symmetry among the sites, it suffices to prove (1.14) is true when $i=N$. Recall ι_N and γ_N from (1.11) and (1.13). We set

$$\iota_N^- = \frac{1}{\sqrt{N-1}}\sum_{j\le N-1} g_{Nj}\sigma_j, \quad r_N^- = \left\langle \iota_N^-\right\rangle_-,$$

where $\langle\cdot\rangle_-$ is the Gibbs measure with Hamiltonian (1.4) and inverse temperature $\beta_- = \sqrt{(N-1)/N}\beta$. Since ι_N^- is a cavity field in $\langle\cdot\rangle_-$, from Proposition 2.8, we know

$$E\left[\frac{\left\langle U(\iota_N^-)\cosh(\beta_-\iota_N^-+h)\right\rangle_-}{\left\langle\cosh(\beta_-\iota_N^-+h)\right\rangle_-} - \frac{E_\xi\left[U(\xi+r_N^-)\cosh(\beta_-(\xi+r_N^-)+h)\right]}{\exp\left(\frac{\beta_-^2}{2}(1-q_-)\right)\cosh(\beta_- r_N^-+h)}\right]^{2k} \le \frac{K}{N^k}, \tag{2.39}$$

where K is a constant depending on β_0, k, and U only. The goal of the proof is then to prove that (1.14) can be related to (2.39). We perform our estimates in several steps.

Step 1. Similar to (1.6), from the Gibbs measure, the following identity holds

$$\langle U(\iota_N)\rangle = \frac{\langle U(\iota_N)\cosh(\beta\iota_N+h)\rangle_-}{\langle\cosh(\beta\iota_N+h)\rangle_-}.$$

Note that $\beta_-\iota_N^- = \beta\iota_N$. Therefore,

$$\frac{\left\langle U(\iota_N^-)\cosh(\beta_-\iota_N^-+h)\right\rangle_-}{\left\langle\cosh(\beta_-\iota_N^-+h)\right\rangle_-} = \frac{\left\langle U(\iota_N^-)\cosh(\beta\iota_N+h)\right\rangle_-}{\left\langle\cosh(\beta\iota_N+h)\right\rangle_-},$$

and this quantity is very close to $\langle U(\iota_N)\rangle$ in the sense that

$$
\begin{aligned}
&E\left[\frac{\left\langle U(\iota_N^-)\cosh(\beta_-\iota_N^-+h)\right\rangle_-}{\left\langle\cosh(\beta_-\iota_N^-+h)\right\rangle_-}-\langle U(\iota_N)\rangle\right]^{2k}\\
&=E\left[\frac{\left\langle\left(U(\iota_N^-)-U(\iota_N)\right)\cosh(\beta\iota_N+h)\right\rangle_-}{\langle\cosh(\beta\iota_N+h)\rangle_-}\right]^{2k}\\
&\le E\left[U(\iota_N^-)-U(\iota_N)\right]^{2k}\\
&\le\frac{K}{N^k}.
\end{aligned}
\tag{2.40}
$$

Indeed, the first inequality is true by the use of Jensen's inequality. The second inequality holds by using the mean value theorem and $\sqrt{N}/\sqrt{N-1}-1\le 2/\sqrt{N}$ together to obtain

$$
\begin{aligned}
|U(\iota_N^-)-U(\iota_N)|&\le|\iota_N^--\iota_N|\sup_{0\le t\le 1}|U'(t\iota_N^-+(1-t)\iota_N)|\\
&\le\frac{\sqrt{2}}{\sqrt{N}}|\iota_N|\sup_{0\le t\le 1}|U'(t\iota_N^-+(1-t)\iota_N)|
\end{aligned}
$$

and then applying Lemma 2.3.

Step 2. Since for any $a,b,c\in\mathbb{R}$,

$$
\frac{\cosh(c(\xi+a)+h)}{\cosh(ca+h)}=\frac{e^{c\xi}}{1+e^{-2ca-2h}}+\frac{e^{-c\xi}}{1+e^{2ca+2h}}
\tag{2.41}
$$

and

$$
\frac{d}{da}\left(\frac{1}{1+e^{\pm 2(ca+h)}}\right)=\frac{\mp 2c}{(e^{ca+h}+e^{-ca-h})^2},
$$

we have

$$
\left|\frac{\cosh(c(\xi+a)+h)}{\cosh(ca+h)}-\frac{\cosh(c(\xi+b)+h)}{\cosh(cb+h)}\right|\le|c||a-b|\cosh(c\xi).
$$

Therefore, by Hölder's inequality, Lemmas 2.3 and 2.9,

$$
\begin{aligned}
&E\left[E_\xi\left[U(\xi+r_N^-)\left(\frac{\cosh(\beta_-(\xi+r_N^-)+h)}{\cosh\left(\beta_-r_N^-+h\right)}-\frac{\cosh(\beta_-(\xi+\gamma_N)+h)}{\cosh\left(\beta_-\gamma_N+h\right)}\right)\right]\right]^{2k}\\
&\le\beta_-^{2k}E\left[|U(\xi+r_N^-)|\cosh(\beta_-\xi)|r_N^--\gamma_N|\right]^{2k}\\
&\le\beta_-^{2k}\left(E\left[|U(\xi+r_N^-)|\cosh(\beta_-\xi)\right]^{4k}\right)^{1/2}\left(E\left[|r_N^--\gamma_N|\right]^{4k}\right)^{1/2}\\
&\le\frac{K}{N^k}.
\end{aligned}
\tag{2.42}
$$

Step 3. Similar to the first step, by the mean value theorem, Lemmas 2.3 and 2.9, we have

$$
E\left[E_\xi\left[\left(U(\xi+r_N^-)-U(\xi+\gamma_N)\right)\frac{\cosh(\beta_-(\xi+\gamma_N)+h)}{\cosh\left(\beta_-\gamma_N+h\right)}\right]\right]^{2k}\le\frac{K}{N^k}.
\tag{2.43}
$$

Step 4. Let us apply the same trick as in the proof for Proposition 2.8 to obtain

$$\begin{aligned}
&E\left[U(\xi+\gamma_N)\frac{\cosh(\beta_-(\xi+\gamma_N)+h)}{\cosh(\beta_-\gamma_N+h)}\right]^{2k}\\
&\le E\left[U(\xi+\gamma_N)^{2k}\left(\frac{\exp(\beta_-(\xi+\gamma_N))e^h+\exp(-\beta_-(\xi+\gamma_N))e^{-h}}{\exp(\beta_-\gamma_N)e^h+\exp(-\beta_-\gamma_N)e^{-h}}\right)^{2k}\right]\\
&\le 2^{2k}\left(E\left[U(\xi+\gamma_N)^{2k}\left(\exp(2k\beta_-\xi)+\exp(-2k\beta_-\xi)\right)\right]\right)\\
&\le K.
\end{aligned}$$

On the other hand, a straightforward computation gives

$$\begin{aligned}
&\frac{\partial^2}{\partial(\beta^2)\partial q}\exp\left(-\frac{\beta^2}{2}(1-q)\right)\\
&=\frac{1}{2}\exp\left(-\frac{\beta^2}{2}(1-q)\right)-\frac{(1-q)\beta^2}{4}\exp\left(-\frac{\beta^2}{2}(1-q)\right).
\end{aligned}$$

Thus, from Lemma 2.9,

$$\begin{aligned}
&\left|\exp\left(-\frac{\beta_-^2}{2}(1-q_-)\right)-\exp\left(-\frac{\beta^2}{2}(1-q)\right)\right|\\
&\le\left(\frac{1}{2}+\frac{1}{4}\beta^2\right)|\beta^2-\beta_-^2||q-q_-|\\
&\le\frac{K}{N^2}
\end{aligned}$$

and so by Jensen's inequality

$$\begin{aligned}
E\Bigg[\Bigg(&\exp\left(-\frac{\beta_-^2}{2}(1-q_-)\right)-\exp\left(-\frac{\beta^2}{2}(1-q)\right)\Bigg)\\
&\times E_\xi\left[U(\xi+\gamma_N)\frac{\cosh(\beta_-(\xi+\gamma_N)+h)}{\cosh(\beta_-\gamma_N+h)}\right]\Bigg]^{2k}\le\frac{K}{N^k}.
\end{aligned}\tag{2.44}$$

Step 5. Notice that from (2.41), we obtain

$$\begin{aligned}
\frac{d}{dc}\frac{\cosh(c(\xi+a)+h)}{\cosh(ca+h)}&=\frac{\xi e^{c\xi}}{1+e^{-2(ca+h)}}+\frac{-\xi e^{-c\xi}}{1+e^{2(ca+h)}}\\
&\quad+\frac{-2ae^{c\xi-2(ca+h)}}{(1+e^{-2(ca+h)})^2}+\frac{2ae^{-c\xi+2(ca+h)}}{(1+e^{2(ca+h)})^2}.
\end{aligned}$$

Since

$$\left|\frac{2ae^{\pm c\xi\mp2(ca+h)}}{(1+e^{\mp2(ca+h)})^2}\right|=\left|\frac{e^{\mp2(ca+h)}}{1+e^{\mp2(ca+h)}}\right|\left|\frac{2ae^{\pm c\xi}}{1+e^{\mp2(ca+h)}}\right|\le2|a|e^{\pm c\xi},$$

it follows that

$$\left|\frac{d}{dc}\frac{\cosh(c(\xi+a)+h)}{\cosh(ca+h)}\right|\le2\left(|\xi|+2|a|\right)\cosh(c\xi)$$

and for $c'<c$,

$$\left|\frac{\cosh(c(\xi+a)+h)}{\cosh(ca+h)}-\frac{\cosh(c'(\xi+a)+h)}{\cosh(c'a+h)}\right|\le2\left(|\xi|+2|a|\right)\int_{c'}^{c}\cosh(t\xi)dt.\tag{2.45}$$

Since from Lemma 2.3, we know

$$E\left[|U(\xi+\gamma_N)|(|\xi|+2|\gamma_N|)\sup_{0\le t\le\beta_0}\cosh(t\xi)\right]^{2k}\le K,$$

it follows, by (2.45) and Jensen's inequality, that

$$E\left[E_\xi\left[U(\xi+\gamma_N)\left(\frac{\cosh(\beta_-(\xi+\gamma_N)+h)}{\cosh(\beta_-\gamma_N+h)}-\frac{\cosh(\beta(\xi+\gamma_N)+h)}{\cosh(\beta\gamma_N+h)}\right)\right]\right]^{2k}\le\frac{K}{N^k}. \quad (2.46)$$

Step 6. Combining (2.39), (2.40), (2.42), (2.43), (2.44), and (2.46), we finally obtain

$$E\left[\langle U(\iota_N)\rangle-\frac{E_\xi\left[U(\xi+\gamma_N)\cosh(\beta(\xi+\gamma_N)+h)\right]}{\exp\left(\frac{\beta^2}{2}(1-q)\right)\cosh(\beta\gamma_N+h)}\right]^{2k}\le\frac{K}{N^k}. \quad (2.47)$$

Substitute the identity

$$\frac{(x-\gamma_N)^2}{2(1-q)}\mp(\beta x+h)=\frac{1}{2(1-q)}(x-(\gamma_N\pm\beta(1-q)))^2\mp(\beta\gamma_N+h)-\frac{\beta^2}{2}(1-q)$$

in the right-hand side of

$$\begin{aligned}&E_\xi\left[U(\xi+\gamma_N)\cosh(\beta(\xi+\gamma_N)+h)\right]\\&=\int\frac{U(x)}{\sqrt{2\pi(1-q)}}\frac{e^{\beta x+h}+e^{-\beta x-h}}{2}\exp\left(-\frac{(x-\gamma_N)^2}{2(1-q)}\right),\end{aligned}$$

then (1.14) holds by (2.47) and we are done. □

Appendix

Proof of Lemma 2.1. It is easy to see that U_0, U_1, and U_2 are of moderate growth. For U_3, since as $|x|\to\infty$,

$$e^{-ax^2}U_3(x)=\exp\left(-\frac{a}{2}(x+r)^2\right)U(x+r)\exp\left(-\frac{a}{2}(x^2-2rx-r^2)\right)\to 0,$$

for all $a>0$, it follows that U_3 is also of moderate growth. The function U_4 is of moderate growth since

$$\begin{aligned}&\limsup_{x^2+y^2\to\infty}|U_4(x,y)|\exp\left(-a(x^2+y^2)\right)\\&\le\limsup_{x^2+y^2\to\infty}1_{\{|x+y|\ge M\}}|U(x+y)|\exp\left(-a(x^2+y^2)\right)\\&+\limsup_{x^2+y^2\to\infty}1_{\{|x+y|<M\}}|U(x+y)|\exp\left(-a(x^2+y^2)\right)\\&=\limsup_{x^2+y^2\to\infty}1_{\{|x+y|\ge M\}}|U(x+y)|\exp\left(-a(x^2+y^2)\right)\\&\le\limsup_{x^2+y^2\to\infty}1_{\{|x+y|\ge M\}}|U(x+y)|\exp\left(-\frac{a}{2}(x+y)^2\right),\end{aligned}$$

for all $M>0$ and also the fact that U is of moderate growth. Since U_3 is of moderate growth, U_5 is well-defined. For $0<b<1$, write

$$(x-y)^2-by^2=\left(\sqrt{1-b}y-\frac{1}{\sqrt{1-b}}x\right)^2-\frac{b}{1-b}x^2.$$

Thus,

$$\begin{aligned}
&\exp\left(-ax^2\right)U_5(x)\\
&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}U(y)\exp\left(-ax^2\right)\exp\left(-\frac{(x-y)^2}{2\sigma^2}\right)dy\\
&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}U(y)\exp\left(-\frac{by^2}{2\sigma^2}\right)\exp\left(-\frac{(x-y)^2-by^2}{2\sigma^2}-ax^2\right)dy\\
&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}U(y)\exp\left(-\frac{by^2}{2\sigma^2}\right)\exp\left(-\frac{1}{2\sigma^2}\left(\sqrt{1-b}y-\frac{1}{\sqrt{1-b}}x\right)^2\right)dy\\
&\quad\times\exp\left(-\left(a-\frac{b}{2\sigma^2(1-b)}\right)x^2\right).
\end{aligned}\tag{2.48}$$

Since U is of moderate growth, $U(y)\exp\left(-\frac{by^2}{2\sigma^2}\right)$ can be regarded as a bounded function in y. On the other hand, since

$$\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2\sigma^2}\left(\sqrt{1-b}y-\frac{1}{\sqrt{1-b}}x\right)^2\right)dy$$

is finite and independent of x, by taking b to be small enough and letting $|x|$ tend to infinity, we conclude, from (2.48), that U_5 is of moderate growth. This completes the proof. □

Proof of Lemma 2.2. For any $x, x', y \in \mathbb{R}$, by the mean value theorem, we can find some $z(x,x',y)$ between x and x' so that $U(x+y)-U(x'+y)=U'(z(x,x',y)+y)(x-x')$. Since U' is of moderate growth, for any $M_1, M_2 > 0$,

$$K_1 := \sup_{|y|\geq M_1, |z|\leq M_2}|U'(z+y)|\exp\left(-\frac{y^2}{4\sigma^2}\right)<\infty.$$

By the continuity of U', $K_2 := \sup_{|y|\leq M_1,|z|\leq M_2}|U'(z+y)|\exp\left(-y^2/2\sigma^2\right)<\infty$. Therefore, for $|x|, |x'|\leq M_2$,

$$\begin{aligned}
&\left|\frac{U(x+y)-U(x'+y)}{x-x'}\right|\exp\left(-\frac{y^2}{2\sigma^2}\right)\\
&\leq K_1\exp\left(-\frac{y^2}{4\sigma^2}\right)1_{\{|y|\geq M_1\}}+K_2 1_{\{|y|<M_2\}}
\end{aligned}$$

and by the dominated convergence theorem, $\psi'(x)=E_\xi U'(\xi+x)$. □

References

[1] Chatterjee, S.: Spin Glasses and Stein's method. *Probab. Theory Relat. Fields* 148 (2010), no. 3-4, 567-600. MR-2678899

[2] Mézard, M., Parisi, G., and Virasoro, M. A.: Spin Glass Theory and Beyond. World Scientific Lecture Notes in Physics, 9. World Scientific Publishing Co., Inc., Teaneck, NJ, (1987). MR-1026102

[3] Sherrington, D. and Kirkpatrick, S.: Solvable model of a spin glass. *Phys. Rev. Lett.* 35 (1975), no. 26, 1792-1796.

[4] Talagrand, M.: Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge A Series of Modern Surveys in Mathematics, 46. Springer-Verlag, Berlin (2003). MR-1993891

[5] Talagrand, M.: Mean Field Models for Spin Glasses. Volume I. Basic Examples. Ergebnisse der Mathematik und ihrer Gren-zgebiete. 3. Folge A Series of Modern Surveys in Mathematics, 54. Springer-Verlag, Berlin (2011). MR-2731561

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

[7] Thouless, D. J., Anderson, P. W., and Palmer, R. G.: Solution of 'solvable model of a spin glass'. *Philos. Mag.* 35:3 (1977), 593-601.

Acknowledgments. The author would like to thank Michel Talagrand for sharing a preliminary version of his book [5], which motivated the present paper, and also Alexander Vandenberg-Rodes for several helpful suggestions regarding the presentation of the paper.