Partial universality of the superconcentration in the spin glass model

Van Hao Can [*] Van Quyet Nguyen [†] Hong Son Vu [‡]

Abstract

Consider the Sherrington-Kirkpatrick model on the complete graph with N vertices with general disordered environments. In the Gaussian case, Chaterjee has shown in [7] an interesting phenomena so-called superconcentration by proving that the variance of the free energy grows sublinearly in N. In this paper, we aim at proving a partial universality of this phenomenon. More precisely, we prove that the variance is sublinear in N when: (i) the disordered random variable, say y, has the first four moments matching to those of the standard normal distribution and has finite fifth moment; or (ii) the law y is symmetry and y is a smooth functional of the standard Gaussian random variable satisfying some mild conditions. In addition, we also consider the universality of of first and second moments of the free energy of the S-K models on general graphs.

1 Introduction

In this paper, we study the superconcentration in the Sherrington-Kirkpatrick (S-K) model. More concretely, we aim to show that the free energy of this model has the sub-diffusive fluctuation in the sense that its variance grows sublinearly in its expectation, and this phenomena is universal as long as the disordered variable satisfies some moment conditions.

Let $N \geq 1$ be an integer and consider the state space $\Sigma_N = \{+1, -1\}^N$. Let $\beta > 0$ be the inverse temperature and let y be a random variable with mean 0 and variance 1 played the role as the disordered distribution. Let $\{y_{ij}\}_{1\leq i<j\leq N}$ be i.i.d. copies of y and define the Hamiltonian as

$$H_y(\sigma) = \frac{\beta}{\sqrt{N}} \sum_{1\leq i<j\leq N} y_{ij}\sigma_i\sigma_j, \qquad \forall\, \sigma \in \Sigma_N$$

and the Gibbs measure as

$$G_y(\sigma) = \frac{\exp(H_y(\sigma))2^{-N}}{Z_y}, \qquad \forall\, \sigma \in \Sigma_N$$

where Z_y is the partition function defined by

$$Z_y = \sum_{\sigma\in\Sigma} \exp(H_y(\sigma))2^{-N}.$$

The free energy of the model is defined as

$$F_y = \log(Z_y).$$

[*]Department of Statistics and Data Science, National University of Singapore, 6 Science Drive 2 Singapore 117546, and Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam. Email: cvhao89@gmail.com

[†]Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam. Email: quyetnv.math@gmail.com

[‡]Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam. Email: vhson1995@gmail.com

Notice that in the general model the Gibbs measure is defined as $G_y(\sigma) \propto \exp(H_y(\sigma))\nu(\sigma)$ with ν a probability measure on Σ_N. Here, ν is the uniform measure, so $\nu(\sigma) = 2^{-N}$ for all $\sigma \in \Sigma_N$. Though the factor 2^{-N} does not effect to the variance of the free energy, we would like to keep this form to state some results for the general S-K model in Section 2.1.

The S-K model has been extensively investigated in physics community for more than fifty years. In particular, the limit of scaled free energy, say $\alpha_\infty = \lim_{N\to\infty} \frac{1}{N} F_N$, is of special interest. By interpolation and mean-field methods, Parisi gave in [15] a celebrated variational formula for α_∞ in the case of Gaussian disorder (i.e. $y = g$ a standard normal random variable) and this immediately became a hot topic studied by both physicists and mathematicians in two decades. The rigorous proof of this formula is derived by Talagrand in this famous paper [16]. Then the next question is that whether the limiting free energy depend on the distribution of the disorder. Generally, the physicists believe that many limiting properties of the spin glass model including the free energy are universal w.r.t. the change of disorder.

The earlier result for disorder universality is obtained by Talagrand in [16] in which he showed that there is no difference between the limiting free energies of models with Bernoulli and Gaussian disorders. This result is then generalized by Guerra and Toninelli in [13] to the disorder distributions having the first four moments matching with the standard Gaussian random variable, and improved later on by Carmona and Hu to the two moments condition in [5]. The universality for the Gibss measure and chaos phenomena is established by Auffinger and Chen by assuming the first four moments condition in [1] and then is sharpened in [8] to the two moments requirement.

Superconcentration, named by Chaterjee in his monograph [7], is the phenomenon that the usual techniques via concentration measures give sub-optimal bounds on the fluctuation of random objects. In fact, this phenomenon had been observed in probability theory in different names and situations before. Particularly, Benjamini, Kalai, Scharm showed that the first passage time in first passage percolation with Bernoulli edge-weight distribution has the sublinear variance in [3]. The proof is based on the $L_1 - L_2$ inequality developed by Talagrand and a regularization technique which is now called the BKS trick. This result is then generalized to general distribution with $2 + \log$ moment condition in [2, 12] by using an entropy inequality. This powerful scheme is also used for other random growth models in [4, 6, 10].

In [7, 8], Chatterjee give a new approach of proving the sublinearity of the variance in some statistical physics models as S-K model, Gaussian fields, First passage percolation. Moreover, he show a deep connection between the sublinear variance bound and the chaotic phenomena in the model, and then he call this phenomena as superconcentration. For example, he show that in the Gaussian S-K model, the overlap of two spin configurations sampled from the original disorder and from a pertubated one is bounded by a function of the variance of the free energy and the pertubation level. As a result, as long as the variance is sublinear the overlap is concentrated at 0 with a small suitable level of pertubation (that means a small change in disorder may lead nearly orthogonal samples of spins). The main idea behind the approach of Chatterjee is built from a variance formula of Gaussian fields and the spectral analysis of Orstein-Ulencbeck semigroup. Below, we recall the result on the sublinear bound for the variance of free energy in S-K model.

Theorem 1.1. *[9, Theorem 1.6] If y has the distribution as the standard normal random variable g, then*

$$\mathrm{Var}[F_g] \leq \frac{C(\beta)N}{\log N}. \tag{1}$$

As mentioned above, since the approach of Theorem 1.1 is based on spectral analysis of Orstein-Ulencbeck process, it could not be generalized directly to other classes of disordered distribution. Our first result aims to prove the same result for general disorder variables whose the first fourth moments match those of the Gaussian one.

Theorem 1.2. *Suppose that the disorder variable has the first four moments matching those of Gaussian one and has the finite fifth moment, i.e.* $\mathbb{E}[y^i] = \mathbb{E}[g^i]$ *for all* $i = 1, \dots, 4$ *and* $\mathbb{E}[|y|^5] < \infty$. *Then there exists a positive constant* C *depending on* β *and* $\mathbb{E}[|y|^5]$, *such that*

$$|\operatorname{Var}[F_y] - \operatorname{Var}[F_g]| \leq CN^{\frac{3}{4}},$$

and as a consequence

$$\operatorname{Var}[F_y] \leq \frac{2CN}{\log N}.$$

To prove Theorem 1.2, we will use the interpolation technique to get the universality of first and second moments of the free energy which results in the universality of the superconcentration. The interpolation scheme for the first moment of F_y has been proceed successfully in [1, 5]. Moreover, we will work on the S-K model in general graph, say $G = ([N], E)$ with N vertices and the set of edge E, see Section 2 for more details. Roughly speaking, in Theorem 2.1, we show that

$$\begin{aligned} |\mathbb{E}[F_y] - \mathbb{E}[F_g]| &\leq C\gamma^{k+1}|E|, \\ |\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| &\leq C\gamma^k|E|\Big(\mathbb{E}[|y|^{k+1}\mathbb{I}(|y| \geq K)]\frac{B}{K} + \gamma KB\Big) \; \forall K \geq 1, \end{aligned}$$

where γ is the coupling constant ($\gamma = \frac{\beta}{\sqrt{n}}$ in the standard case), B is a number depending on $\gamma, |E|, \mathbb{E}[F_g]$ and the maximal degree $d_{\max}$ and $\nu_{\min} = \min_{\sigma \in \Sigma_N} \nu(\sigma)$ with ν the probability measure in the definition of Gibbs measure.

In the second result, we study the case that disorder variable is a Gaussian functional. More concretely, we assume that $y = h(g)$, where $h : \mathbb{R} \to \mathbb{R}$ is a smooth function satisfying

$$h(g) \overset{(d)}{=} -h(g), \tag{H1}$$

and there exists a strictly increasing function $p_h : \mathbb{N} \to \mathbb{R}$, such that $p_h(k) \to \infty$ as $k \to \infty$,

$$k^{2k}\left[1 + \sup_{1 \leq t \leq k} \sup_{\mathbf{m} \in \mathcal{M}_t} \sup_{\mathbf{a} \in \mathcal{A}_{\mathbf{m}}} \sup_{\mathbf{q} \in \mathcal{Q}_{\mathbf{m}}} \sup_{b_{\mathbf{a},\mathbf{q}} \in \mathcal{B}_{\mathbf{m}}} \mathbb{E}\Big[\big(h^{(b_{\mathbf{a},\mathbf{q}})}(g)\big)^2\Big]\right] \leq p_h(k), \tag{H2}$$

where the indices $\mathbf{a}, \mathbf{q}, b_{\mathbf{a},\mathbf{q}}, \mathcal{M}_t, \mathcal{A}_{\mathbf{m}}, \mathcal{Q}_{\mathbf{m}}, \mathcal{B}_{\mathbf{m}}$ are defined in Lemma 3.3; for $\mathbf{m} \in \mathcal{M}_t$, $\mathbf{a} \in \mathcal{A}_{\mathbf{m}}$ and $b_{\mathbf{a},\mathbf{q}} = (b_{a_1,q_1,1}, \dots b_{a_1,q_1,m_1}, b_{a_2,q_2,1} \dots, b_{a_2,q_2,m_2}, \dots, b_{a_t,q_t,1}, \dots, b_{a_t,q_t,m_t}) \in \mathcal{B}_{\mathbf{m}}$,

$$\mathbb{E}\Big[\big(h^{(b_{\mathbf{a},\mathbf{q}})}(g)\big)^2\Big] = \prod_{s=1}^{t} \mathbb{E}\Big[\prod_{r_s=1}^{m_s} \big(h^{(b_{a_s,q_s,r_s})}(g)\big)^2\Big],$$

where $h^{(b)}$ is the b-th derivative of h with the convention that $h^{(0)} = 1$.

Theorem 1.3. *Assume that* $y = h(g)$ *with* h *a smooth function satisfying (H1) and (H2). Then there exists a positive constant* $C = C(\beta)$, *such that*

$$\operatorname{Var}[F_y] \leq \frac{CN}{p_h^{-1}\big(N^{\frac{1}{6}}\big)},$$

where p_h^{-1} *is the inverse function of* p_h. *In particular, since* $p_h^{-1}(k) \to \infty$ *as* $k \to \infty$, *the superconcentration of the free energy holds.*

The condition (H1) holds when h is an odd function, i.e. $h(x) = -h(x)$ for all $x \in \mathbb{R}$. To quantify the function $p_h(\cdot)$, we consider the condition (H2') below and obtain a clearer bound for $\mathrm{Var}[F_y]$.

Corollary 1.4. *Assume that h is a smooth and odd function and there exists an increasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, such that*

$$|h^{(b)}(x)| \le \exp\left(\frac{x^2}{\varphi(|x|)}\right) \qquad \forall b \ge 0,\, x \in \mathbb{R}. \tag{H2'}$$

Then there exist a positive constant C, such that

$$\mathrm{Var}[F_y] \le \frac{CN}{p_h^{-1}\big(N^{\frac{1}{6}}\big)}, \qquad p_h(k) = 2k^{2k} \sup_{1\le t\le k} \sup_{\boldsymbol{m}\in\mathcal{M}_t} \prod_{s=1}^{t} \big(\exp\big((\varphi^{-1}(8m_s))^2\big) + 2\big),$$

where $\boldsymbol{m} \in \mathcal{M}_t = \{(m_1, \dots, m_t) : m_1 + \ldots + m_t = k\}$. In particular, if $\varphi_1(x) = cx^\alpha$, or $\varphi_2(x) = c\log(x+1)$ with $c, \alpha > 0$, then we have corresponding bounds

$$\mathrm{Var}[F_{\varphi_1}] \le CN\Big(\frac{\log\log N}{\log N} + (\log N)^{\frac{-\alpha}{2}}\Big), \qquad \mathrm{Var}[F_{\varphi_2}] \le \frac{CN\log\log\log N}{\log\log N}.$$

The condition (H2') holds for most of usual functions satisfying that all its derivatives $|h^{(b)}(x)|$ grow like $\exp(o(x^2))$.

To obtain Theorem 1.3, we follow the proof of Theorem 1.1. Since $y = h(g)$, we can consider F_y as a smooth functional of Gaussian fields. Then in order to get an upper bound for the variance of F_y, we need to estimate the expectation of all its derivatives. The derivative formula of F_y for general h is more complicated than the simplest case $h(x) = x$. It leads a quite messy expression of the bound of the variance, though the bound is still sublinear.

Organization of the paper. In Section 2, we consider the S-K model in general graphs and give a comparison for the first and second moments of free energy in arbitrary disordered distribution with the Gaussian case. Then we obtain Theorem 1.2 as a particular case. In Section 3, we use an improved Poincaré inequality and computation of the derivative of the free energy to get an upper bound for the variance and prove Theorem 1.3.

2 The superconcentration of free energy in general environments under the fourth moment conditions

In this section, we prove the universality for the first and second moments (and as consequence the variance) of free energy of a general S-K model.

Let $G = ([N], E)$ be a simple graph with vertex set $[N] = \{1, \dots, N\}$ and the set of edge E. We set $\Sigma_N = \{+1, -1\}^N$ as the state space. Let ν be a probabilistic measure on Σ_N, and let $f = (f_e)_{e\in E}$ family of measurable functions on Σ_N satisfying

$$\max_{\sigma\in\Sigma_N} \max_{e\in E} |f_e(\sigma)| \le 1, \qquad \sum_{\sigma\in\Sigma_N} \sum_{e\in E} x_e f_e(\sigma)\nu(\sigma) \ge 0, \qquad \forall (x_e)_{e\in E} \in \mathbb{R}^E. \tag{Hf}$$

Let $\gamma > 0$ be the temperature parameter and let y be a random variable with mean 0 and variance 1. Given the environment disorder formed by $(y_e)_{e\in E}$ which are i.i.d copies of y, we define the Hamiltonian on Σ_N as

$$H_y(\sigma) = \gamma \sum_{e\in E} y_e f_e(\sigma), \qquad \forall\, \sigma \in \Sigma_N \tag{2}$$

and the Gibbs measure as

$$G_y(\sigma) = \frac{\exp(H_y(\sigma))\nu(\sigma)}{Z_y}, \qquad \forall\, \sigma \in \Sigma_N \tag{3}$$

where Z_y is the partition function defined by

$$Z_y = \sum_{\sigma\in\Sigma} \exp(H_y(\sigma))\nu(\sigma). \tag{4}$$

Denote by $\langle\cdot\rangle_y$ the Gibbs expectation associated to G_y, i.e. for any measure function $L : \Sigma_N \to \mathbb{R}$,

$$\langle L\rangle_y = \sum_{\sigma\in\Sigma_N} L(\sigma)G_y(\sigma).$$

The free energy of the model is defined as

$$F_y = \log(Z_y).$$

Notice that in the standard S-K model, G is the complete graph and ν the uniform measure:

$$E = \{ij, 1 \le i < j \le N\}, \quad \nu(\sigma) = 2^{-N}\, \forall\sigma \in \Sigma_N, \quad \gamma = \frac{\beta}{\sqrt{N}}, \quad f_e(\sigma) = \sigma_i\sigma_j \text{ for } e = ij.$$

Theorem 2.1. *Let $k \ge 2$ and consider the general S-K model assuming that (Hf) holds. Suppose that $\mathbb{E}[y^i] = \mathbb{E}[g^i]$ for all $i = 1, \dots, k$ and $\mathbb{E}[|y|^{k+1}] < \infty$. Then there exists a positive constant C depending on k and $\mathbb{E}[|y|^{k+1}]$, such that the following assertions hold.*

(i) We have

$$|\mathbb{E}[F_y] - \mathbb{E}[F_g]| \le C\gamma^{k+1}|E|.$$

(ii) For any $K \ge 1$,

$$|\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| \le C\gamma^k|E|\Big(\mathbb{E}[|y_e|^{k+1}\mathbb{I}(|y_e| \ge K)]\frac{B}{K} + \gamma KB\Big),$$

where

$$B = \mathbb{E}[F_g] + \gamma^{k+1}|E| + \gamma d_{max} - \log \nu_{min}, \quad d_{max} = \max_{1\le i\le n} \deg(i), \quad \nu_{min} = \min_{\sigma\in\Sigma_N} \nu(\sigma).$$

Proof of Theorem 1.2. We apply Theorem 2.1 to the standard S-K model, where

$$\gamma = \frac{\beta}{\sqrt{N}}, \quad |E| = \frac{N(N-1)}{2}, \quad d_{\max} = N - 1, \quad \nu_{\min} = 2^{-N}, \quad \mathbb{E}[F_g] = \mathcal{O}_\beta(N). \tag{5}$$

By (i),

$$|\mathbb{E}[F_y] - \mathbb{E}[F_g]| \le \mathcal{O}_\beta(1)N^{-\frac{k+1}{2}+2} = \mathcal{O}_\beta(1)N^{\frac{3-k}{2}}.$$

By (5), we have $B = \mathcal{O}_\beta(N)$ and thus (ii) implies

$$\begin{aligned} |\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| &\le C\gamma^k|E|\Big(\mathbb{E}[|y_e|^{k+1}\mathbb{I}(|y_e| \ge K)]\frac{N}{K} + \gamma KN\Big) \\ &= \mathcal{O}_\beta(1)\left(\frac{N^{3-\frac{k}{2}}}{K}\mathbb{E}\big[|y|^{k+1}\mathbb{I}(|y| \ge K)\big] + N^{\frac{5-k}{2}}K\right). \end{aligned}$$

Let $k = 4$ and $K = N^{\frac{1}{4}}$, we obtain that

$$|\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| \le CN^{\frac{3}{4}}, \qquad |\mathbb{E}[F_y] - \mathbb{E}[F_g]| \le CN^{\frac{-1}{2}},$$

for some constant $C > 0$. Therefore, since $|\mathbb{E}[F_y] + \mathbb{E}[F_g]| = \mathcal{O}_\beta(N)$,

$$|\mathrm{Var}[F_y] - \mathrm{Var}[F_g]| \le |\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| + |\mathbb{E}[F_y] - \mathbb{E}[F_g]||\mathbb{E}[F_y] + \mathbb{E}[F_g]| \le CN^{\frac{3}{4}},$$

for some positive constant C. The remaining bound for $\mathrm{Var}[F_y]$ follows from the above estimate and Theorem 1.1. □

The remaining of this section is devoted to prove Theorem 2.1. In the first part (subsection 2.1), we will present some preparatory results on the rate of approximation by interpolating methods and on the estimate of some derivative formulas. In the second part (subsection 2.2), we give the proof of Theorem 2.1.

First, we introduce some notation used throughout this section. Let E be a finite set and F be a function in $C^k(\mathbb{R}^E)$. Then, for any $x = (x_e)_{e\in E}$ we denote the k^{th} partial derivative of F with respect to x_e by

$$\partial_e^k F(x) = \frac{\partial^k}{\partial x_e^k} F(x).$$

For $k = 1$, we simply write $\partial_e F$ for $\partial_e^1 F$.

2.1 Preliminaries

Our major ingredient for the proof of Theorem 2.1 is the following integration by part approximation which inspired by [1, Lemma 2.2].

Lemma 2.2. *Let $(y_e)_{e\in E}$ be a family of i.i.d random variables such that its first $k \ge 2$ moments match those of the standard Gaussian random variable. Let F be in $C^{k+1}(\mathbb{R}^E)$. Then for any $e \in E$ and $K \ge 1$,*

$$\big|\mathbb{E}[y_e F(y_e, \bar{y}_e)] - \mathbb{E}[\partial_e F(y_e, \bar{y}_e)]\big| \le I_1 + I_2, \tag{6}$$

where $\bar{y}_e = (y_{e'})_{e'\in E\setminus e}$, and

$$I_1 = \frac{2}{(k-1)!}\mathbb{E}\Big[\,\big(\sup_{|x_e|\le|y_e|}\big|\partial_e^{k-1}F(x_e,\bar{y}_e)\big| + \sup_{|x_e|\le|y_e|}\big|\partial_e^{k}F(x_e,\bar{y}_e)\big|\big)|y_e|^k\mathbb{I}(|y_e|\ge K)\Big],$$

and

$$I_2 = K\mathbb{E}\Big[\sup_{|x_e|\le|y_e|}\big|\partial_e^{k}F(x_e,\bar{y}_e)\big|\Big(\frac{|y_e|^k}{k!} + \frac{|y_e|^{k-2}}{(k-1)!}\Big)\mathbb{I}(|y_e|\le K)\Big].$$

Assume in addition that $\mathbb{E}[|y_e|^{k+1}] < \infty$. Then

$$\big|\mathbb{E}[y_e F(y_e, \bar{y}_e)] - \mathbb{E}[\partial_e F(y_e, \bar{y}_e)]\big| \le \mathbb{E}\Big[\sup_{|x_e|\le|y_e|}\big|\partial_e^{k}F(x_e,\bar{y}_e)\big|\Big(\frac{|y_e|^{k+1}}{k!} + \frac{|y_e|^{k-1}}{(k-1)!}\Big)\Big]. \tag{7}$$

Proof. Fix $e \in E$, define a function $\hat{F} : \mathbb{R} \to \mathbb{R}$ by

$$\hat{F}(x_e) = \mathbb{E}_{\bar{y}_e}[F(x_e, \bar{y}_e)].$$

Here and below, we denote by $\mathbb{E}_{\bar{y}_e}$ and $\mathbb{E}_{y_e}$ the expectation w.r.t. $\bar{y}_e$ and y_e correspondingly. Then $\hat{F} \in C^{k+1}(\mathbb{R})$, since $F \in C^{k+1}(\mathbb{R}^E)$, and

$$\mathbb{E}[y_e F(y_e, \bar{y}_e)] - \mathbb{E}[\partial_e F(y_e, \bar{y}_e)] = \mathbb{E}[y_e \hat{F}(y_e)] - \mathbb{E}[\hat{F}'(y_e)] \tag{8}$$

By using similar arguments in [1, proof of Lemma 2.2] (see in particular (A.4) and (A.5)), we can show that for any $K \ge 1$,

$$\big|\mathbb{E}[y_e \hat{F}(y_e)] - \mathbb{E}[\hat{F}'(y_e)]\big| \le \hat{I}_1 + \hat{I}_2, \tag{9}$$

where

$$\hat{I}_1 = \frac{2}{(k-1)!}\mathbb{E}\Big[\big(\sup_{|x_e|\le|y_e|}\big|\hat{F}^{(k-1)}(x_e)\big| + \sup_{|\xi_e|\le|y_e|}\big|\hat{F}^{(k)}(x_e)\big|\big)|y_e|^k\mathbb{I}(|y_e|\ge K)\Big],$$

and

$$\hat{I}_2 = K\mathbb{E}\Big[\sup_{|x_e|\le|y_e|}\big|\hat{F}^{(k)}(x_e)\big|\Big(\frac{|y_e|^k}{k!}+\frac{|y_e|^{k-2}}{(k-1)!}\Big)\mathbb{I}(|y_e|\le K)\Big],$$

with $\hat{F}^{(k)}$ the k^{th} derivative of $\hat{F}$. In addition, observe that for any $l\le k$ and any function g

$$\begin{aligned}\mathbb{E}\Big[\sup_{|x_e|\le|y_e|}\big|\hat{F}^{(l)}(x_e)\big||g(y_e)|\Big] &= \mathbb{E}_{y_e}\Big[\sup_{|x_e|\le|y_e|}\big|\mathbb{E}_{\bar{y}_e}[\partial_e^l F(x_e,\bar{y}_e)]\big||g(y_e)|\Big]\\ &\le \mathbb{E}_{y_e}\mathbb{E}_{\bar{y}_e}\Big[\sup_{|x_e|\le|y_e|}\big|\partial_e^l F(x_e,\bar{y}_e)\big||g(y_e)|\Big]\end{aligned}$$

Therefore,

$$\hat{I}_1\le I_1,\qquad \hat{I}_2\le I_2. \tag{10}$$

Combining (8), (9) and (10), we obtain (6). Moreover, if $\mathbb{E}[|y_e|^{k+1}]<\infty$ then by the same arguments in [1, Lemma 2.2], we also have

$$\big|\mathbb{E}[y_e\hat{F}(y_e)]-\mathbb{E}[\hat{F}'(y_e)]\big|\le\mathbb{E}\Big[\sup_{|x_e|\le|y_e|}\big|\hat{F}^{(k)}(x_e)\big|\Big(\frac{|y_e|^{k+1}}{k!}+\frac{|y_e|^{k-1}}{(k-1)!}\Big)\Big].$$

Then the proof of (7) follows from similar arguments as above. □

Remark 1. In [1, Lemma 2.2], the authors assume that $\sup_x|\partial_e^k F(x)|\le C<\infty$. This condition holds for $F_t=\langle\sigma_i\sigma_j\rangle_t$ but not for $F_t=\langle\sigma_i\sigma_j\rangle_t\log Z_t$ (both of these functions will appear in the interpolation of the variance of free energy, see more in the proof of Theorem 1.2). Therefore, in our lemma 2.2, we have to clarify the bounds on the derivatives of F_t as in the formula of I_1 and I_2.

Lemma 2.3. *Let $n\ge1$ and L be a measurable function on Σ_N^n with $\|L\|_\infty\le1$ and consider the free energy Z_y defined in (4) as a smooth function of $y=(y_e)_{e\in E}\in\mathbb{R}^E$. Then for any $k\ge1$, there exists a positive constant C depending on k and n, such that the following assertions hold.*

(i) ([1, Lemma 4.2]) For any $e\in E$ and $y=(y_e)_{e\in E}$,

$$\Big|\partial_e^k(\langle L\rangle_y)\Big|\le C\gamma^k. \tag{11}$$

(ii) For any $e\in E$ and $y=(y_e)_{e\in E}$,

$$\Big|\partial_e^k\Big(\langle L\rangle_y\log Z_y\Big)\Big|\le C\gamma^k\big(1+\log(Z_y)\big). \tag{12}$$

Proof. The first part is exactly [1, Lemma 4.2]. For the second one, we define a function $F\in C^\infty(\mathbb{R}^E)$ by

$$F(y)=\langle L\rangle_y\log Z_y.$$

By the chain rule, for any $l\ge1$ and $e\in E$,

$$\begin{aligned}\partial_e^l[\log(Z_y)] &= \sum_{\mathbf{m}\in\mathbb{N}^l:|\mathbf{m}|_1=l}\binom{l}{\mathbf{m}}\frac{\partial^{|\mathbf{m}|}\log(Z_y)}{\partial Z_y^{|\mathbf{m}|}}\prod_{j=1}^{l}\Big(\frac{\partial_e^j Z_y}{j!}\Big)^{m_j}\\ &= \sum_{\mathbf{m}\in\mathbb{N}^l:|\mathbf{m}|_1=l}\binom{l}{\mathbf{m}}\frac{(-1)^{|\mathbf{m}|-1}(|\mathbf{m}|-1)!}{Z_y^{|\mathbf{m}|}}\prod_{j=1}^{l}\Big(\frac{\gamma^j Z_y\langle(f_e(\sigma))^j\rangle_y}{j!}\Big)^{m_j},\end{aligned}$$

where $\mathbf{m} = (m_1, \dots, m_l)$ and

$$|\mathbf{m}| = \sum_{i=1}^{l} m_i, \quad |\mathbf{m}|_1 = \sum_{i=1}^{l} i m_i, \quad \binom{l}{\mathbf{m}} = \frac{l!}{m_1! m_2! \dots m_l!}.$$

Since $|f_e(\sigma)| \le 1$, it follows that

$$\left|\partial_e^l[\log(Z_y)]\right| \le \sum_{\mathbf{m}\in\mathbb{N}^l:|\mathbf{m}|_1=l} \binom{l}{\mathbf{m}} \frac{(|\mathbf{m}|-1)!}{Z_y^{|\mathbf{m}|}} \prod_{j=1}^{l} \Big(\frac{\gamma^j Z_y}{j!}\Big)^{m_j} = \gamma^l C_l,$$

for some positive constant C_l depending only on l. Combining this estimate with (11), we get that for any $y = (y_e)_{e\in E} \in \mathbb{R}^E$,

$$\begin{aligned} \left|\partial_e^k F(y)\right| &= \Big|\sum_{l=0}^{k} \binom{k}{l} \partial_e^{k-l}[\langle L\rangle_y] \partial_e^l[\log(Z_y)]\Big| \\ &= \Big|\sum_{l=1}^{k} \binom{k}{l} \partial_e^{k-l}[\langle L\rangle_y] \partial_e^l[\log(Z_y)] + \partial_e^k[\langle L\rangle_y]\log(Z_y)\Big| \\ &\le C\gamma^k\big(1 + \left|\log(Z_y)\right|\big) = C\gamma^k\big(1+\log(Z_y)\big) \end{aligned}$$

for some $C = C(k, n)$. In the last equation, we have used $\log(Z_y) \ge 0$, or $Z_y \ge 1$. Indeed, by Jensen's inequality,

$$Z_y = \sum_{\sigma\in\Sigma_N} \exp(H_y(\sigma))\nu(\sigma) = \mathbb{E}_\nu\big[\exp(H_y)\big] \ge \exp(\mathbb{E}_\nu[H_y]) \ge 1,$$

since

$$\mathbb{E}_\nu[H_y] = \sum_{e\in E}\sum_{\sigma\in\Sigma_N} y_e f_e(\sigma)\nu(\sigma) \ge 0,$$

by using the hypothesis (Hf). □

2.2 Proof of Theorem 2.1

As we mentioned, the proof will be based on interpolation technique and integration by parts formulas obtained in Lemma 2.2. Let us consider the interpolated Hamiltonian between H_g and H_y defined as

$$H_t(\sigma) = \sum_{e\in E} \gamma(\sqrt{t}y_e + \sqrt{1-t}g_e) f_e(\sigma), \qquad t \in [0,1],$$

and the corresponding partition function

$$Z_t = \sum_{\sigma\in\Sigma_N} \exp(H_t(\sigma))\nu(\sigma).$$

The interpolated Gibbs measure (denoted by G_t) and its average (denoted by $\langle\cdot\rangle_t$) are defined as usual

$$\langle L\rangle_t = \sum_{\sigma\in\Sigma_N} L(\sigma)G_t(\sigma), \qquad G_t(\sigma) = \frac{\exp(H_t(\sigma))\nu(\sigma)}{Z_t}\ \forall\,\sigma\in\Sigma_N.$$

Consider the interpolated free energy

$$Q_1(t) = \mathbb{E}[\log Z_t].$$

Then $Q_1(0) = \mathbb{E}[F_g]$ and $Q_1(1) = \mathbb{E}[F_y]$. Thus

$$|\mathbb{E}[F_y] - \mathbb{E}[F_g]| \le \sup_{0\le t\le 1} |Q_1'(t)|. \tag{13}$$

By the direct computation, we have

$$Q_1'(t) = \frac{\gamma}{2\sqrt{t}} \sum_{e\in E} \mathbb{E}[y_e \langle f_e(\sigma)\rangle_t] - \frac{\gamma}{2\sqrt{1-t}} \sum_{e\in E} \mathbb{E}[g_e \langle f_e(\sigma)\rangle_t]. \tag{14}$$

Moreover,

$$\frac{\partial}{\partial y_e} \langle f_e(\sigma)\rangle_t = \gamma\sqrt{t}\big(\langle f_e^2(\sigma)\rangle_t - \langle f_e(\sigma)\rangle_t^2\big),$$

and by the Gaussian integration by parts formula,

$$\mathbb{E}[g_e \langle f_e(\sigma)\rangle_t] = \mathbb{E}\Big[\frac{\partial}{\partial g_e}\langle f_e(\sigma)\rangle_t\Big] = \gamma\sqrt{1-t}\mathbb{E}[(\langle f_e^2(\sigma)\rangle_t - \langle f_e(\sigma)\rangle_t^2)].$$

By the above two identities, $\mathbb{E}[\partial_{y_e}\langle f_e(\sigma)\rangle_t] = \sqrt{\frac{t}{1-t}}\mathbb{E}[g_e\langle f_e(\sigma)\rangle_t]$, and thus

$$\Big|\frac{\mathbb{E}[y_e \langle f_e(\sigma)\rangle_t]}{2\sqrt{t}} - \frac{\mathbb{E}[g_e \langle f_e(\sigma)\rangle_t]}{2\sqrt{1-t}}\Big| = \frac{1}{2\sqrt{t}}\Big|\mathbb{E}[y_e\langle f_e(\sigma)\rangle_t] - \mathbb{E}\Big[\frac{\partial\langle f_e(\sigma)\rangle_t}{\partial y_e}\Big]\Big|. \tag{15}$$

Using Lemma 2.3 (i), we get that

$$\sup_{x\in\mathbb{R}^E}\Big|\frac{\partial^k}{\partial y_e^k}\langle f_e(\sigma)\rangle_t\Big|_{y=x} \le C\gamma^k.$$

Plugging this into the approximation (7) yields that for all $t \in [0,1]$ and $e \in E$,

$$\frac{1}{2\sqrt{t}}\Big|\mathbb{E}[y_e\langle f_e(\sigma)\rangle_t] - \mathbb{E}\Big[\frac{\partial\langle f_e(\sigma)\rangle_t}{\partial y_e}\Big]\Big| \le C\gamma^k t^{(k-1)/2}\mathbb{E}[|y_e|^{k+1}] \le C\gamma^k\mathbb{E}[|y_e|^{k+1}],$$

where C is a positive constant depending only on k. Combining this bound with (14) and (15), we obtain that

$$\sup_{0\le t\le 1} |Q_1'(t)| \le C\gamma^{k+1}|E|. \tag{16}$$

This estimate together with (13) yields the result (i).

To compare the second moment and prove (ii), we define

$$Q_2(t) = \mathbb{E}[\log^2(Z_t)].$$

Then $Q_2(1) = \mathbb{E}[F_y^2]$ and $Q_2(0) = \mathbb{E}[F_g^2]$, and hence

$$|\mathbb{E}[F_y^2] - \mathbb{E}[F_g^2]| \le \sup_{0\le t\le 1}|Q_2'(t)|. \tag{17}$$

First, we have

$$\begin{aligned} Q_2'(t) &= \frac{\gamma}{\sqrt{t}} \sum_{e\in E} \mathbb{E}[y_e\langle f_e(\sigma)\rangle_t \log Z_t] - \frac{\gamma}{\sqrt{1-t}} \sum_{e\in E}\mathbb{E}[g_e\langle f_e(\sigma)\rangle_t \log Z_t] \\ &= \frac{\gamma}{\sqrt{t}} \sum_{e\in E}\mathbb{E}[y_e F_{t,e}] - \frac{\gamma}{\sqrt{1-t}} \sum_{e\in E}\mathbb{E}[g_e F_{t,e}], \end{aligned} \tag{18}$$

where

$$F_{t,e} = \langle f_e(\sigma)\rangle_t \log Z_t.$$

For each $e \in E$, we compute

$$\begin{aligned} \frac{\partial}{\partial g_e} G_t(\sigma) &= \frac{\partial}{\partial g_e}\left(\frac{\exp(H_t(\sigma))\nu(\sigma)}{Z_t}\right) = G_t(\sigma)\frac{\partial}{\partial g_e}H_t(\sigma) - G_t(\sigma)\frac{\partial_{g_e} Z_t}{Z_t} \\ &= \gamma\sqrt{1-t}G_t(\sigma)[f_e(\sigma) - \langle f_e(\sigma)\rangle_t], \end{aligned}$$

where $\tilde{\sigma}$ is an independent sample of σ given the environment $(y_e, g_e)_{e\in E}$.

Therefore, using the Gaussian integration by parts,

$$\mathbb{E}[g_e F_{t,e}] = \mathbb{E}\Big[\frac{\partial F_{t,e}}{\partial g_e}\Big] = \mathbb{E}\Big[\log Z_t \sum_{\sigma\in\Sigma_N} f_e(\sigma)\partial_{g_e} G_t(\sigma) + \langle f_e(\sigma)\rangle_t \partial_{g_e}(\log Z_t)\Big]$$
$$= \gamma\sqrt{1-t}\,\mathbb{E}\Big[\log Z_t(\langle (f_e(\sigma))^2\rangle_t - \langle f_e(\sigma)\rangle_t^2) + \langle f_e(\sigma)\rangle_t^2\Big].$$

Similarly,

$$\frac{\partial F_{t,e}}{\partial y_e} = \gamma\sqrt{t}\Big[\log Z_t(\langle (f_e(\sigma))^2\rangle_t - \langle f_e(\sigma)\rangle_t^2) + \langle f_e(\sigma)\rangle_t^2\Big].$$

It follows from the last two equations that

$$\frac{1}{\sqrt{t}}\mathbb{E}[y_e F_{t,e}] - \frac{1}{\sqrt{1-t}}\mathbb{E}[g_e F_{t,e}] = \frac{1}{\sqrt{t}}\left(\mathbb{E}[y_e F_{t,e}] - \mathbb{E}\Big[\frac{\partial F_{t,e}}{\partial y_e}\Big]\right) \tag{19}$$

Moreover, by Lemma 2.2 (in particular (6)) and Lemma 2.3 (ii), for any $K \geq 1$

$$\begin{aligned}
&\frac{1}{\sqrt{t}}\Big|\mathbb{E}[y_e F_{t,e}] - \mathbb{E}\Big[\frac{\partial F_{t,e}}{\partial y_e}\Big]\Big| \\
&\leq \frac{2}{\sqrt{t}(k-1)!}\mathbb{E}\Big[\big(\sup_{|x_e|\leq|y_e|}\big|\partial_e^{k-1}F_{t,e}(x_e,\bar{y}_e)\big| + \sup_{|x_e|\leq|y_e|}\big|\partial_e^k F_t(x_e,\bar{y}_e)\big|\big)|y_e|^k\mathbb{I}(|y_e|\geq K)\Big] \\
&\qquad + \frac{K}{\sqrt{t}}\mathbb{E}\Big[\sup_{|x_e|\leq|y_e|}\big|\partial_e^k F_{t,e}(x_e,\bar{y}_e)\big|\Big(\frac{|y_e|^k}{k!} + \frac{|y_e|^{k-2}}{(k-1)!}\Big)\mathbb{I}(|y_e|\leq K)\Big] \\
&\leq C(A_{1,e}(t) + A_{2,e}(t)),
\end{aligned} \tag{20}$$

where C is a positive constant and,

$$A_{1,e}(t) = \gamma^{k-1}t^{(k-2)/2}\mathbb{E}\Big[\big(1 + \sup_{|x_e|\leq|y_e|}\log Z_t(x_e,\bar{y}_e)\big)|y_e|^k\mathbb{I}(|y_e|\geq K)\Big], \tag{21}$$

and

$$A_{2,e}(t) = \gamma^k t^{(k-1)/2}K\mathbb{E}\Big[\big(1 + \sup_{|x_e|\leq|y_e|}\log Z_t(x_e,\bar{y}_e)\big)\Big(\frac{|y_e|^k}{k!} + \frac{|y_e|^{k-2}}{(k-1)!}\Big)\mathbb{I}(|y_e|\leq K)\Big]. \tag{22}$$

Combining (17)–(20), we get

$$|\mathbb{E}[F_g^2] - \mathbb{E}[F_y^2]| \leq C\gamma|E|\sup_{t\in[0,1]}\max_{e\in E}[A_{1,e}(t) + A_{2,e}(t)] \tag{23}$$

Next, we estimate $A_{1,e}$ and $A_{2,e}$. First, observe that

$$\begin{aligned}
&\sup_{|x_e|\leq|y_e|}\log Z_t(x_e,\bar{y}_e) \leq \sup_{|x_e|\leq|y_e|}\max_{\sigma\in\Sigma_N}H_t(\sigma) = \max_{\sigma\in\Sigma_N}\sup_{|x_e|\leq|y_e|}H_t(\sigma) \\
&\leq \max_{\sigma\in\Sigma_N}\sum_{e'\neq e}\gamma(\sqrt{t}y_{e'} + \sqrt{1-t}g_{e'})f_{e'}(\sigma) + \gamma(\sqrt{t}|y_e| + \sqrt{1-t}|g_e|) \\
&\leq \max_{\sigma\in\Sigma_N}H_{t,e}(\sigma) + \sum_{e'\in E_{1,e}}\gamma(\sqrt{t}|y_{e'}| + \sqrt{1-t}|g_{e'}|) + \gamma(\sqrt{t}|y_e| + \sqrt{1-t}|g_e|),
\end{aligned} \tag{24}$$

where $H_{t,e}(\sigma)$ is the Hamiltonian restricted to $(N-2)$ vertices $[n]\setminus[e]$ with $[e]$ the two extreme points of e, and $E_{1,e}$ is the set of edges sharing with e an extreme point, that means

$$H_{t,e}(\sigma) = \sum_{e'\in E_{2,e}}\gamma(\sqrt{t}y_{e'} + \sqrt{1-t}g_{e'})f_{e'}(\sigma),$$

with

$$E_{1,e} = \{e' : |[e'] \cap [e]| = 1\}, \qquad E_{2,e} = \{e' : [e'] \cap [e] = \varnothing\}.$$

In (24), for the second line we have used the fact that $H_t(\sigma)$ is a linear function of x_e, so it attains the maximum value at the boundary $x_e = |y_e|$.

The key point in the above decomposition is that $H_{t,e}(\sigma)$ and $\{y_{e'}, g_{e'}\}_{e' \in E_{1,e}}$ are independent of y_e and g_e. Therefore, using $\mathbb{E}[|y_e|], \mathbb{E}[|g_e|] = \mathcal{O}(1)$ and $|E_{1,e}| \le 2d_{\max}$, we have

$$\begin{aligned} A_{1,e}(t) &\le C\gamma^{k-1} t^{(k-2)/2} \Big(\mathbb{E}[\max_{\sigma \in \Sigma_N} H_{t,e}(\sigma)] + d_{\max}\gamma + 1 \Big) \mathbb{E}[|y_e|^k \mathbb{I}(|y_e| \ge K)] \\ &\quad + C\gamma^k t^{(k-2)/2} \mathbb{E}[|y_e|^{k+1} \mathbb{I}(|y_e| \ge K)], \end{aligned}$$

for some $C > 0$. In addition,

$$\begin{aligned} \max_{\sigma \in \Sigma_N} H_{t,e}(\sigma) &\le \max_{\sigma \in \Sigma_N} H_t(\sigma) + \gamma \Big[\sum_{e' \in E_{1,e}} (|y_{e'}| + |g_{e'}|) + |y_e| + |g_e| \Big] \\ &\le \log Z_t - \log \nu_{\min} + \gamma \Big[\sum_{e' \in E_{1,e}} (|y_{e'}| + |g_{e'}|) + |y_e| + |g_e| \Big]. \end{aligned}$$

Thus

$$\mathbb{E}[\max_{\sigma \in \Sigma_N} H_{t,e}(\sigma)] \le \mathbb{E}[\log Z_t] - \log \nu_{\min} + C d_{\max}\gamma,$$

for some $C > 0$, since $|E_{1,e}| \le 2d_{\max}$. In summary, for all $t \in [0,1]$,

$$\begin{aligned} A_{1,e}(t) &\le C\gamma^{k-1} t^{(k-2)/2} \Big(\mathbb{E}[\log Z_t] + d_{\max}\gamma - \log \nu_{\min} \Big) \mathbb{E}[|y_e|^k \mathbb{I}(|y_e| \ge K)] \\ &\quad + C\gamma^k t^{(k-2)/2} \mathbb{E}[|y_e|^{k+1} \mathbb{I}(|y_e| \ge K)] \\ &= \mathcal{O}(1)\gamma^{k-1} \mathbb{E}[|y_e|^{k+1} \mathbb{I}(|y_e| \ge K)] \Big(\frac{\mathbb{E}[\log Z_t] + d_{\max}\gamma - \log \nu_{\min}}{K} + \gamma \Big), \end{aligned}$$

where $\mathcal{O}(1)$ depends only on k and $\mathbb{E}[|y|^{k+1}]$. Similarly,

$$\begin{aligned} A_{2,e}(t) &\le C\gamma^k t^{(k-1)/2} K \Big(\mathbb{E}[\log Z_t] + d_{\max}\gamma - \log \nu_{\min} \Big) \mathbb{E}[|y_e|^{k+1}] + C\gamma^{k+1} K \mathbb{E}[|y_e|^{k+1}] \\ &= \mathcal{O}(1)\gamma^k K \Big(\mathbb{E}[\log Z_t] + d_{\max}\gamma - \log \nu_{\min} \Big). \end{aligned}$$

By (16), we have

$$\mathbb{E}[\log Z_t] = Q_1(t) \le (Q_1(0) + \sup_{0 \le s \le t} |Q_1'(s)|) \le \mathbb{E}[F_g] + C\gamma^{k+1}|E|.$$

Therefore,

$$\begin{aligned} \sup_{0 \le t \le 1} \max_{e \in E} A_{1,e}(t) + A_{2,e}(t) &\le C\gamma^{k-1} \mathbb{E}[|y_e|^{k+1} \mathbb{I}(|y_e| \ge K)] \Big(\frac{B}{K} + \gamma \Big) + C\gamma^k K B \\ &\le C\gamma^{k-1} \mathbb{E}[|y_e|^{k+1} \mathbb{I}(|y_e| \ge K)] \frac{B}{K} + 2C\gamma^k K B \end{aligned}$$

where C is a large constant, and

$$B = \mathbb{E}[F_g] + \gamma^{k+1}|E| + \gamma d_{\max} - \log \nu_{\min}.$$

Combining this estimate with (23), we obtain (ii). □

3 The superconcentration of free energy in Gaussian functional environments

Suppose that the environment is formed by a sequence $\{y_{ij}\}_{1\le i<j\le N}$ of i.i.d. random variables with the same distribution as $y = h(g)$. Then the free energy defined by

$$F_y = \log[Z_y],$$

can be viewed as a function of i.i.d. Gaussian variables $\{g_{ij}\}_{1\le i<j\le N}$ by setting $y_{ij} = h(g_{ij})$. Using the Ornstein-Uhlenbeck diffusion, Chatterjee gives in [7] an improved Poincaré inequality allowing him to get a sublinear bound for the variance of F_g (i.e. $h(x) = x$). In this section, we aim at proving a similar result for general function h.

Proposition 2. *[7, Theorem 6.2] Let γ^N be the product measure of N i.i.d. standard normal distribution and let f be a smooth function which is in $L^2(\gamma^N)$. Then for any $m \ge 1$,*

$$\mathrm{Var}_{\gamma^N}[f] \le \sum_{k=1}^{m-1} \frac{\theta_k(f)}{k!} + \frac{\mathbb{E}_{\gamma^N}[|\nabla f|^2]}{m},$$

where

$$\theta_k(f) = \sum_{1\le i_1,\dots,i_k\le N} \Big(\mathbb{E}_{\gamma^N}\Big[\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_k}}\Big]\Big)^2.$$

Proposition 3. *Consider the S-K model whose the disordered variable is $y = h(g)$ with h satisfying (H1) and (H2). For any $k \ge 1$, we have*

$$\theta_k(F_y) = \sum_{1\le i_1<j_1,\dots,i_k<j_k\le N} \Big(\mathbb{E}\Big[\frac{\partial^k F_y}{\partial g_{i_1 j_1}\dots\partial g_{i_k j_k}}\Big]\Big)^2 \le (1+\beta^{2k})k^{8k+1}p_h(k),$$

where $p_h(k)$ is given in (H2).

Proof of Theorem 1.3. First, notice that

$$\mathbb{E}[|\nabla F_y|^2] = \frac{\beta^2}{N}\mathbb{E}\Big[\sum_{1\le i<j\le N}(h'(g_{ij}))^2\langle\sigma_i\sigma_j\rangle_y^2\Big] \le \frac{\beta^2}{N}\sum_{1\le i<j\le N}\mathbb{E}[(h'(g_{ij}))^2] \le CN,$$

where C is positive constant depending on β and $\mathbb{E}[(h'(g))^2]$. Combining this estimate with Propositions 2 and 3, we have for any $m \ge 1$

$$\begin{aligned}\mathrm{Var}[F_y] &\le \sum_{k=1}^{m-1}\frac{(1+\beta^{2k})k^{8k+1}p_h(k)}{k!} + \frac{CN}{m} \le (1+\beta^{2m})m^{8m+2}p_h(m) + \frac{CN}{m}\\ &\le C\left(\frac{(p_h(m))^6}{m} + \frac{N}{m}\right),\end{aligned}$$

since $p_h(m) \ge m^{2m}$. Taking $m = [p_h^{-1}(N^{\frac{1}{6}})]$, we have

$$\mathrm{Var}[F_y] \le \frac{2CN}{p_h^{-1}(N^{\frac{1}{6}})},$$

and Theorem 1.3 follows. □

Proof of Corollary 1.4. Assume that there exists an increasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, such that

$$|h^{(b)}(x)| \le \exp\left(\frac{x^2}{\varphi(|x|)}\right) \qquad \forall b \ge 0,\ x \in \mathbb{R}. \tag{H2'}$$

Then, we have

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{r_s=1}^{m_s}(h^{(b_{a_s,q_s,r_s})}(g))^2\Big] \le \mathbb{E}\Big[\exp\Big(\frac{2m_s g^2}{\varphi(|g|)}\Big)\Big]\\
= &\sqrt{\frac{2}{\pi}}\int_0^\infty \exp\Big(-\frac{x^2}{2}\Big(1-\frac{4m_s}{\varphi(x)}\Big)\Big)dx = \sqrt{\frac{2}{\pi}}\left(\int_0^{\varphi^{-1}(8m_s)} + \int_{\varphi^{-1}(8m_s)}^\infty\right)\\
\le &\sqrt{\frac{2}{\pi}}\left(\varphi^{-1}(8m_s)\exp\big((\varphi^{-1}(8m_s))^2/4\big) + \int_0^\infty e^{-\frac{x^2}{4}}dx\right) \le \exp\big((\varphi^{-1}(8m_s))^2\big)+2.
\end{aligned}$$

Therefore,

$$\sup_{1\le t\le k}\sup_{\mathbf{m}\in\mathcal{M}_t}\sup_{\mathbf{a}\in\mathcal{A}_\mathbf{m}}\sup_{\mathbf{q}\in\mathcal{Q}_\mathbf{m}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_\mathbf{m}}\mathbb{E}\Big[\big(h^{(b_{\mathbf{a},\mathbf{q}})}(g)\big)^2\Big] \le \sup_{1\le t\le k}\sup_{\mathbf{m}\in\mathcal{M}_t}\prod_{s=1}^t\big(\exp\big((\varphi^{-1}(8m_s))^2\big)+2\big),$$

where $\mathbf{m}\in\mathcal{M}_t = \{(m_1,\ldots,m_t): m_1+\ldots+m_t = k\}$, and we can choose

$$p_h(k) = 2k^{2k}\sup_{1\le t\le k}\sup_{\mathbf{m}\in\mathcal{M}_t}\prod_{s=1}^t\big(\exp\big((\varphi^{-1}(8m_s))^2\big)+2\big).$$

Consider $\varphi_1(x) = cx^\alpha$ then $\varphi_1^{-1}(m) = \frac{1}{c}m^{\frac{1}{\alpha}}$ and, we can choose

$$p_{\varphi_1}(k) = \exp(C(k\log k + k^{\frac{2}{\alpha}})),$$

for some large constant C. Therefore

$$\mathrm{Var}[F_{\varphi_1}] \le \frac{CN}{p_h^{-1}\big(N^{\frac{1}{6}}\big)} \le CN\Big(\frac{\log\log N}{\log N} + (\log N)^{\frac{-\alpha}{2}}\Big).$$

The case $\varphi_2(x) = c\log(x+1)$ can be treated similarly, and here

$$p_{\varphi_2}(k) \le \exp(\exp(Ck)),$$

and thus

$$\mathrm{Var}[F_{\varphi_2}] \le \frac{C'N\log\log\log N}{\log\log N},$$

for some large constants C and C'. $\square$

The rest of this section is devoted to the proof of Proposition 3. In the subsection 3.1, we prove some high order derivative formulas of the free energy function. In the subsection 3.2, using the formulas derived, we prove Proposition 3.

3.1 Preliminaries

Fix β and let $F_y = F_y(\beta) = \log[Z_y]$. Then F_y is a smooth function of $(g_{ij})_{1\le i<j\le N}$. Below, we are going to compute the derivatives of F_y.

Let $\sigma^1,\sigma^2,\ldots,\sigma^n$ denote n i.i.d. configurations drawn from the Gibbs measure G_y. Note that $\sigma^1,\sigma^2,\ldots,\sigma^n$ are conditionally independent given the disorder g, but unconditionally dependent. The following key lemma to constitute the general k^{th} derivative of free energy of S-K model. We define the set of edge of the complete graph as

$$E = \{ij : 1\le i<j\le N\}. \tag{25}$$

Lemma 3.1. *For any $ij\in E$ and $k\ge 1$, there exist a collection of real constants $\{\{\gamma_{a,q},\{b_{a,q,r}\}_{r\le k}\}_{q\le k!},\{c_a(\boldsymbol{l})\}_{\boldsymbol{l}\in\mathcal{L}_a}\}_{a\le}$ satisfying the following.*

(a) $\gamma_{a,q} \in \{0,1\}$ *and if* $\gamma_{a,q} = 1$ *then* $\sum_{r=1}^{k} b_{a,q,r} = k$ *for all* $a \le k$ *and* $q \le k!$.

(b) $|c_a(\boldsymbol{l})| \le (a-1)!$ *for all* $1 \le a \le k$ *and* $\boldsymbol{l} \in \mathcal{L}_a$, *where*

$$\mathcal{L}_a = \{(l_1, \dots, l_a) \in \mathbb{N}^a : 1 \le l_s \le s \ \forall\ 1 \le s \le a\}.$$

(c) *We have*

$$\frac{\partial^k F_y}{\partial^k g_{ij}} = \sum_{a=1}^{k} \Big(\frac{\beta}{\sqrt{N}}\Big)^a \sum_{q=1}^{k!} \gamma_{a,q} \prod_{r=1}^{k} h^{(b_{a,q,r})}(g_{ij}) \sum_{\boldsymbol{l} \in \mathcal{L}_a} c_a(\boldsymbol{l}) \langle \sigma_i^{l_1} \sigma_j^{l_1} \dots \sigma_i^{l_a} \sigma_j^{l_a} \rangle_y, \tag{26}$$

where $h^{(b)}$ *is the* b^{th} *derivative of* h *with the convention that* $h^{(0)} = 1$.

Lemma 3.2. *Let* $n \ge 1$ *and* L *be a measurable function on* Σ_N^n. *For any* $ij \in E$ *and* $k \ge 1$, *there exists a collection of constants* $\{\{\gamma_{a,q}, \{b_{a,q,r}\}_{r \le k}\}_{q \le k!}, \{c_a(\boldsymbol{l})\}_{\boldsymbol{l} \in \mathcal{L}_{n,a}}\}_{a \le k}$ *satisfying the conditions (a) Lemma 3.1 and the following.*

(b') $|c_a(\boldsymbol{l})| \le \dfrac{(n+a-1)!}{(n-1)!}$ *for all* $1 \le a \le k$ *and* $\boldsymbol{l} \in \mathcal{L}_{n,a}$, *where*

$$\mathcal{L}_{n,a} = \{(l_1, \dots, l_a) \in \mathbb{N}^a : 1 \le l_s \le n + s\}.$$

(c') *We have*

$$\frac{\partial^k \langle L \rangle_y}{\partial g_{ij}} = \sum_{a=1}^{k} \Big(\frac{\beta}{\sqrt{N}}\Big)^a \sum_{q=1}^{k!} \gamma_{a,q} \prod_{r=1}^{k} h^{(b_{a,q,r})}(g_{ij}) \sum_{\boldsymbol{l} \in \mathcal{L}_{n,a}} c_a(\boldsymbol{l}) \langle L \sigma_i^{l_1} \sigma_j^{l_1} \dots \sigma_i^{l_a} \sigma_j^{l_a} \rangle_y.$$

Remark 4. We notice that the coefficients γ, b, c_a may differ when L, i, j, k change. However, for the simplicity, we use the same notation in the both lemmas 3.1 and 3.2. These coefficients are in fact so complicated that we do not have explicit expressions. Fortunately, the bounds in conditions (a) and (b), or (b') are enough for our purpose.

Proof of Lemmas 3.1 and 3.2. We prove Lemma 3.1 by induction arguments in k. First, the formula (26) is true for $k = 1$ since

$$\frac{\partial}{\partial g_{ij}} F_y = \frac{\beta}{\sqrt{N}} h^{(1)}(g_{ij}) \langle \sigma_i \sigma_j \rangle_y. \tag{27}$$

Next, suppose that the claim holds up to $k \ge 1$. Given $ij \in E$, we need to compute the $(k+1)^{\text{st}}$ derivative of F_y w.r.t. g_{ij}. By the hypothesis induction,

$$\frac{\partial^{k+1} F_y}{\partial^{k+1} g_{ij}} = \sum_{a=1}^{k} \Big(\frac{\beta}{\sqrt{N}}\Big)^a \frac{\partial (T_{1,a} \times T_{2,a})}{\partial g_{ij}}, \tag{28}$$

where for any $1 \le a \le k$

$$\begin{aligned} T_{1,a} &= \sum_{q=1}^{k!} \gamma_{a,q} \prod_{r=1}^{k} h^{(b_{a,q,r})}(g_{ij}), \\ T_{2,a} &= \sum_{\mathbf{l} \in \mathcal{L}_a} c_a(\mathbf{l}) \langle \sigma_i^{l_1} \sigma_j^{l_1} \dots \sigma_i^{l_a} \sigma_j^{l_a} \rangle_y. \end{aligned}$$

We have

$$\frac{\partial T_{1,a}}{\partial g_{ij}} = \sum_{q=1}^{k!} \gamma_{a,q} \sum_{r=1}^{k} \prod_{s\neq r} h^{(b_{a,q,s})}(g_{ij}) h^{(b_{a,q,r}+1)}(g_{i_r j_r}) \mathbb{I}(b_{a,q,r} \neq 0)$$
$$= \sum_{\substack{q'=(q,r)\\ q\le k!,\, r\le k}} \gamma'_{a,q'} \prod_{s=1}^{k} h^{(b'_{a,q',s})}(g_{i_s j_s}),$$

where for $q' = (q,r)$ and $0 \le s, r \le k$,

$$\gamma'_{a,q'} = \gamma_{a,q}\mathbb{I}(b_{a,q,r} \neq 0), \qquad b'_{a,q',s} = \begin{cases} b_{a,q,s} & \text{if } s \neq r \\ b_{a,q,s}+1 & \text{if } s = r. \end{cases} \tag{29}$$

Since $1 \le q \le k!$ and $1 \le r \le k$, we enumerate the q' as ranging from 1 to $k!k$, and obtain that

$$\frac{\partial T_{1,a}}{\partial g_{ij}} = \sum_{q'=1}^{k!k} \gamma'_{a,q'} \prod_{s=1}^{k+1} h^{(b'_{a,q',s})}(g_{ij}), \tag{30}$$

by defining

$$b'_{a,q',k+1} = 0, \tag{31}$$

and noting that $h^{(0)} = 1$. Next, we compute the derivative of $T_{2,a}$. Given a tuple $(l_1, \ldots, l_a)$ with $l_i \le i$ for all $i = 1, \ldots, a$, by the standard computation of Gibbs measure we have

$$\frac{\partial}{\partial g_{ij}} \langle \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a}\rangle_y$$
$$= \frac{\partial}{\partial g_{ij}} \sum_{\sigma^1,\ldots,\sigma^a\in\Sigma_N} \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a} G_y(\sigma^1)\ldots G_y(\sigma^a)$$
$$= \sum_{\sigma^1,\ldots,\sigma^a\in\Sigma_N} \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a} \Big(\sum_{s=1}^{a} \frac{1}{G_y(\sigma^s)} \frac{\partial G_y(\sigma^s)}{\partial g_{ij}}\Big) G_y(\sigma^1)\ldots G_y(\sigma^a)$$
$$= \frac{\beta}{\sqrt{N}} h^{(1)}(g_{ij}) \sum_{\sigma^1,\ldots,\sigma^a\in\Sigma_N} \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a} \Big(\sum_{s=1}^{a} (\sigma_i^s\sigma_j^s - \langle\sigma_j^{a+1}\sigma_j^{a+1}\rangle_y\Big) G_y(\sigma^1)\ldots G_y(\sigma^a)$$
$$= \frac{\beta}{\sqrt{N}} h^{(1)}(g_{ij}) \sum_{l=1}^{a} \langle \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a} (\sigma_i^l\sigma_j^l - \sigma_i^{a+1}\sigma_j^{a+1})\rangle_y$$
$$= \frac{\beta}{\sqrt{N}} h^{(1)}(g_{ij}) \sum_{l_{a+1}=1}^{a+1} d_{a+1}((l_s)_{s\le a+1}) \langle \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a}\sigma_i^{l_{a+1}}\sigma_j^{l_{a+1}}\rangle_y, \tag{32}$$

where $d_{a+1}((l_s)_{s\le a+1}) = 1$ if $1 \le l_{a+1} \le a$ and $d_{a+1}((l_s)_{s\le a+1}) = -a$, otherwise. Using this identity, we compute the derivative of $T_{2,a}$ as below

$$\frac{\partial T_{2,a}}{\partial g_{ij}} = \frac{\partial}{\partial g_{ij}} \sum_{\mathbf{l}\in\mathcal{L}_a} c_a(\mathbf{l}) \langle \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_a}\sigma_j^{l_a}\rangle_y$$
$$= \frac{\beta}{\sqrt{N}} h^{(1)}(g_{ij}) \sum_{\mathbf{l}\in\mathcal{L}_{a+1}} c'_{a+1}(\mathbf{l}) \langle \sigma_i^{l_1}\sigma_j^{l_1} \ldots \sigma_i^{l_{a+1}}\sigma_j^{l_{a+1}}\rangle_y,$$

where for $\mathbf{l} = (l_1, \ldots, l_{a+1})$,

$$c'_a(\mathbf{l}) = c'_{a+1}((l_s)_{s\le a+1}) = c_a((l_s)_{s\le a}) d_{a+1}((l_s)_{s\le a+1}). \tag{33}$$

Therefore,

$$T_{1,a}\frac{\partial T_{2,a}}{\partial g_{ij}} = \frac{\beta}{\sqrt{N}}\sum_{q=1}^{k!}\gamma_{a,q}\prod_{r=1}^{k+1}h^{(b_{a,q,r})}(g_{ij})\sum_{\mathbf{l}\in\mathcal{L}_{a+1}}c'_{a+1}(\mathbf{l})\langle\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_{a+1}}\sigma_j^{l_{a+1}}\rangle_y, \tag{34}$$

where we defined

$$b_{a,q,k+1} = 1. \tag{35}$$

Using (28), (30) and (34), we derive

$$\begin{aligned}
\frac{\partial^{k+1}F_y}{\partial^{k+1}g_{ij}} &= \sum_{a=1}^{k}\Big(\frac{\beta}{\sqrt{N}}\Big)^a\sum_{q'=1}^{k!k}\gamma'_{a,q'}\prod_{s=1}^{k+1}h^{(b'_{a,q',s})}(g_{ij})\sum_{\mathbf{l}\in\mathcal{L}_a}c_a(\mathbf{l})\langle\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_a}\sigma_j^{l_a}\rangle_y\\
&+\sum_{a=1}^{k}\Big(\frac{\beta}{\sqrt{N}}\Big)^{a+1}\sum_{q=1}^{k!}\gamma_{a,q}\prod_{l=r}^{k+1}h^{(b_{a,q,r})}(g_{ij})\sum_{\mathbf{l}\in\mathcal{L}_{a+1}}c'_{a+1}(\mathbf{l})\langle\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_{a+1}}\sigma_j^{l_{a+1}}\rangle_y\\
&=\sum_{a=1}^{k+1}\Big(\frac{\beta}{\sqrt{N}}\Big)^a\sum_{q=1}^{(k+1)!}\gamma^*_{a,q}\prod_{r=1}^{k+1}h^{(b^*_{a,q,r})}(g_{ij})\sum_{\mathbf{l}\in\mathcal{L}_{a+1}}c^*_a(\mathbf{l})\langle\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_a}\sigma_j^{l_a}\rangle_y,
\end{aligned}$$

where

$$\gamma^*_{1,q} = \begin{cases}\gamma'_{1,q} & \text{if } 1\le q\le k!k,\\ 0 & \text{if } k!k+1\le q\le (k+1)!\end{cases},\quad \gamma^*_{k+1,q} = \begin{cases}0 & \text{if } 1\le q\le k!k,\\ \gamma_{k,q-k!k} & \text{if } k!k+1\le q\le (k+1)!,\end{cases}$$

and for each $2\le a\le k$,

$$\gamma^*_{a,q} = \begin{cases}\gamma'_{a,q} & \text{if } 1\le q\le k!k,\\ \gamma_{a,q-k!k} & \text{if } k!k+1\le q\le (k+1)!,\end{cases}$$

The constants $b^*_{a,q,r}$ and $c^*(\mathbf{l})$ are determined correspondingly to $\gamma^*_{a,q}$. Furthermore, we notice that by (33) for all $a=0,\dots,k$ and $\mathbf{l}=(l_1,\dots,l_{a+1})\in\mathcal{L}_{a+1}$,

$$|c^*_{a+1}(\mathbf{l})|\le |c_a((l_s)_{s\le a})d_{a+1}((l_s)_{s\le a+1})|\le (a-1)!a = a!.$$

In addition, $\gamma^*_{a,q}\in\{0,1\}$, and if $\gamma^*_{a,q}=1$ then $\gamma'_{a,q}=1$ or $\gamma_{a,r}=1$ or $\gamma_{k,r}=1$ (when $a=k+1$), with $r=q-k!k$. In all cases, by using (29), (31), (35) we always have

$$\sum_{s=1}^{k+1}b^*_{a,q,s} = 1+\sum_{s=1}^{k}b_{a',q,s} = k+1,\qquad a'=\min\{a,k\}.$$

The proof of Lemma 3.1 is finished.

Now, we prove Lemma 3.2. For any measurable function $L=L(\sigma^1,\dots,\sigma^n)$, we have

$$\begin{aligned}
\frac{\partial\langle L\rangle_y}{\partial g_{ij}} &= \frac{\partial}{\partial g_{ij}}\sum_{\sigma^1,\dots,\sigma^n}L(\sigma^1,\dots,\sigma^n)G_y(\sigma^1)\dots G_y(\sigma^n)\\
&=\sum_{\sigma^1,\dots,\sigma^n}L(\sigma^1,\dots,\sigma^n)\left(\sum_{s=1}^{n}\frac{1}{G_y(\sigma^s)}\frac{\partial G_y(\sigma^s)}{\partial g_{ij}}\right)G_y(\sigma^1)\dots G_y(\sigma^n)\\
&=\frac{\beta}{\sqrt{N}}h^{(1)}(g_{ij})\sum_{\sigma^1,\dots,\sigma^n}L(\sigma^1,\dots,\sigma^n)\left(\sum_{s=1}^{n}\left[\sigma_i^s\sigma_j^s-\langle\sigma_i^{n+1}\sigma_j^{n+1}\rangle_y\right]\right)G_y(\sigma^1)\dots G_y(\sigma^n)\\
&=\frac{\beta}{\sqrt{N}}h^{(1)}(g_{ij})\sum_{1\le l_1\le n+1}c_1(l_1)\langle L\sigma_i^{l_1}\sigma_j^{l_1}\rangle_y,
\end{aligned}\tag{36}$$

where $c_1(l_1) = 1$ for $1 \le l_1 \le n$ and $c_1(n+1) = -n$. This formula is similar to the first derivative of F_y in (27). Then the proof of Lemma 3.2 can be conducted essentially the same as we have done for Lemma 3.1, so we omit it. The only difference is that here we need to compute the derivative of $\langle L\sigma^{l_1}\dots\sigma^{l_a}\rangle_y$ instead of $\langle \sigma^{l_1}\dots\sigma^{l_a}\rangle_y$. Using the same arguments for (32), we have for any $\mathbf{l} \in \mathcal{L}_{n,a} = \{(l_1,\dots,l_a) : 1 \le l_s \le n+s \ \forall 1 \le s \le a\}$,

$$\begin{aligned}
&\frac{\partial}{\partial g_{ij}}\langle L\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_a}\sigma_j^{l_a}\rangle_y \\
&= \frac{\beta}{\sqrt{N}}h^{(1)}(g_{ij})\sum_{l_{a+1}=1}^{a+1} d_{a+1}((l_s)_{s\le a+1})\langle L\sigma_i^{l_1}\sigma_j^{l_1}\dots\sigma_i^{l_a}\sigma_j^{l_a}\sigma_i^{l_{a+1}}\sigma_j^{l_{a+1}}\rangle_y,
\end{aligned}$$

where $d_{a+1}((l_s)_{s\le a+1}) = 1$ if $1 \le l_{a+1} \le n+a$ and $d_{a+1}((l_s)_{s\le a+1}) = -(n+a)$, otherwise. This explains the fact that the new tuple $(l_1,\dots,l_{a+1}) \in \mathcal{L}_{n,a+1}$ instead of $\mathcal{L}_{a+1}$ as in Lemma 3.1. Moreover, the bound for $c_a((l_s)_{s\le a+1})$ is changed as

$$|c_a((l_s)_{s\le a+1})| \le |c_a((l_s)_{s\le a})||d_{a+1}((l_s)_{s\le a+1})| \le \frac{(n+a-1)!}{(n-1)!}(n+a) = \frac{(n+a)!}{(n-1)!},$$

by the hypothesis induction that $|c_a((l_s)_{s\le a})| \le \frac{(n+a-1)!}{(n-1)!}$. □

To apply Proposition 2, we have to estimate all the derivatives of F_y. Let $k \ge 1$ and $T = \{i_1j_1,\dots,i_kj_k\} \in E^k$. We are going to compute the derivative of F_y w.r.t. the variable indexed by T. First, we reorder T as

$$T = \{\underbrace{i'_1j'_1,\dots,i'_1j'_1}_{m_1 \text{ times}},\dots,\underbrace{i'_tj'_t,\dots,i'_tj'_t}_{m_t \text{ times}}\}, \quad 1 \le m_i \le k \ \forall 1 \le i \le t, \quad \sum_{i=1}^t m_i = k, \quad 1 \le t \le k, \tag{37}$$

where $i'_pj'_p \ne i'_qj'_q$ for all $1 \le p \ne q \le t$. For the later convenience, we introduce some notation:

$$\begin{aligned}
E^t_{\ne} &= \{\mathbf{i'j'} = \{i'_1j'_1,\dots,i'_tj'_t\} \in E^t : i'_pj'_p \ne i'_qj'_q \ \forall 1 \le p \ne q \le t\},\\
\mathcal{M}_t &= \Big\{\mathbf{m} = (m_1,\dots,m_t) : 1 \le m_s \le k \,\forall 1 \le s \le t, \sum_{s=1}^t m_s = k\Big\},\\
\mathcal{A}_\mathbf{m} &= \{\mathbf{a} = (a_1,\dots,a_t) : 1 \le a_i \le m_i \,\forall\, i = 1,\dots,t\}, \qquad |\mathbf{a}| = a_1 + \dots + a_t,\\
\mathcal{Q}_\mathbf{m} &= \{\mathbf{q} = (q_1,\dots,q_t) : 1 \le q_i \le m_i! \,\forall\, i = 1,\dots,t\},\\
\mathcal{L}_{|\mathbf{a}|} &= \{\mathbf{l} = (l_1,\dots,l_{|\mathbf{a}|}) : 1 \le l_i \le i \,\forall\, i = 1,\dots,|\mathbf{a}|\},\\
\mathcal{B}_\mathbf{m} &= \Big\{b_{\mathbf{a},\mathbf{q}} = (b^1_{\mathbf{a},\mathbf{q},1},\dots,b^1_{\mathbf{a},\mathbf{q},m_1},\dots,b^t_{\mathbf{a},\mathbf{q},1},\dots,b^t_{\mathbf{a},\mathbf{q},m_t}) \in \mathbb{N}^k :\\
&\qquad\qquad \sum_{r_1=1}^{m_1} b^1_{\mathbf{a},\mathbf{q},r_1} = m_1,\dots,\sum_{r_t=1}^{m_t} b^t_{\mathbf{a},\mathbf{q},r_t} = m_t\Big\},\\
h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i'j'}}) &= \prod_{r_1=1}^{m_1} h^{(b^1_{a_1,q_1,r_1})}(g_{i'_1j'_1})\dots\prod_{r_t=1}^{m_t} h^{(b^t_{a_t,q_t,r_t})}(g_{i'_tj'_t}),\\
\sigma^{\mathbf{l}}_{\mathbf{i'j'}} &= \sigma^{l_1}_{i'_1}\sigma^{l_1}_{j'_1}\dots\sigma^{l_{a_1}}_{i'_1}\sigma^{l_{a_1}}_{j'_1}\sigma^{l_{a_1+1}}_{i'_2}\sigma^{l_{a_1+1}}_{j'_2}\dots\sigma^{l_{a_1+a_2}}_{i'_2}\sigma^{l_{a_1+a_2}}_{j'_2}\dots\sigma^{l_{|\mathbf{a}|}}_{i'_t}\sigma^{l_{|\mathbf{a}|}}_{j'_t}.
\end{aligned}$$

We obtain the general derivative of F_y by the following lemma.

Lemma 3.3. *Let T be the tuple as in (37). There exists a collection of constants $\{\{\gamma_{\boldsymbol{a},\boldsymbol{q}}, b_{\boldsymbol{a},\boldsymbol{q}}\}_{\boldsymbol{q}\in\mathcal{Q}_m}, \{c_{|\boldsymbol{a}|}(\boldsymbol{l})\}_{\boldsymbol{l}\in\mathcal{L}_{|\boldsymbol{a}|}}\}_{\boldsymbol{a}\in\mathcal{A}_m}$ satisfying the following.*

(a) $\gamma_{\boldsymbol{a},\boldsymbol{q}} \in \{0,1\}$ and if $\gamma_{\boldsymbol{a},\boldsymbol{q}=1}$ then $b_{\boldsymbol{a},\boldsymbol{q}} \in \mathcal{B}_{\boldsymbol{m}}$ for all $\boldsymbol{a} \in \mathcal{A}_{\boldsymbol{m}}$ and $\boldsymbol{q} \in \mathcal{Q}_{\boldsymbol{m}}$,

(b) $|c_{|\boldsymbol{a}|}(\boldsymbol{l})| \le (|\boldsymbol{a}|-1)!$ *for all* $\boldsymbol{l} \in \mathcal{L}_{|\boldsymbol{a}|}$ *and* $\boldsymbol{a} \in \mathcal{A}_{\boldsymbol{m}}$,

(c) *We have*

$$\frac{\partial^k F_y}{\partial^{m_1} g_{i_1'j_1'} \dots \partial^{m_t} g_{i_t'j_t'}} = \sum_{\boldsymbol{a}\in\mathcal{A}_{\boldsymbol{m}}} \Big(\frac{\beta}{\sqrt{N}}\Big)^{|\boldsymbol{a}|} \sum_{\boldsymbol{q}\in\mathcal{Q}_{\boldsymbol{m}}} \gamma_{\boldsymbol{a},\boldsymbol{q}} h^{(b_{\boldsymbol{a},\boldsymbol{q}})}(g_{\boldsymbol{i}'\boldsymbol{j}'}) \sum_{\boldsymbol{l}\in\mathcal{L}_{|\boldsymbol{a}|}} c_{|\boldsymbol{a}|}(\boldsymbol{l}) \langle \boldsymbol{\sigma}^{\boldsymbol{l}}_{\boldsymbol{i}'\boldsymbol{j}'}\rangle_y, \quad (38)$$

Proof. We first use Lemma 3.1, then sequentially apply Lemma 3.2 to the function F_y to obtain that

$$\begin{aligned}
&\frac{\partial^k F_y}{\partial^{m_1} g_{i_1'j_1'} \dots \partial^{m_t} g_{i_t'j_t'}} = \frac{\partial^{k-m_1}}{\partial^{m_2} g_{i_2'j_2'} \dots \partial^{m_t} g_{i_t'j_t'}} \Big(\frac{\partial^{m_1} F_y}{\partial^{m_1} g_{i_1'j_1'}}\Big) \\
&= \sum_{a_1=1}^{m_1} \Big(\frac{\beta}{\sqrt{N}}\Big)^{a_1} \sum_{q_1=1}^{m_1!} \gamma^1_{a_1,q_1} \prod_{r_1=1}^{m_1} h^{(b^1_{a_1,q_1,r_1})}(g_{i_1'j_1'}) \\
&\qquad \times \frac{\partial^{k-m_1-m_2}}{\partial^{m_3} g_{i_3'j_3'} \dots \partial^{m_t} g_{i_t'j_t'}} \Big(\frac{\partial^{m_2}}{\partial^{m_2} g_{i_2'j_2'}} \sum_{\mathbf{l}^1\in\mathcal{L}_{a_1}} c^1_{a_1}(\mathbf{l}^1) \langle \sigma^{\mathbf{l}^1}_{i_1'j_1'}\rangle_y\Big) \\
&= \sum_{a_1=1}^{m_1}\sum_{a_2=1}^{m_2} \Big(\frac{\beta}{\sqrt{N}}\Big)^{a_1+a_2} \sum_{q_1=1}^{m_1!}\sum_{q_2=1}^{m_2!} \gamma^1_{a_1,q_1}\gamma^2_{a_2,q_2} \prod_{r_1=1}^{m_1} h^{(b^1_{a_1,q_1,r_1})}(g_{i_1'j_1'}) \prod_{r_2=1}^{m_2} h^{(b^2_{a_2,q_2,r_2})}(g_{i_2'j_2'}) \\
&\times \frac{\partial^{k-m_1-m_2}}{\partial^{m_3} g_{i_3'j_3'} \dots \partial^{m_t} g_{i_t'j_t'}} \Big(\sum_{\mathbf{l}^1\in\mathcal{L}_{a_1}} \sum_{\mathbf{l}^2\in\mathcal{L}_{a_1,a_2}} c^1_{a_1}(\mathbf{l}^1) c^2_{a_2}(\mathbf{l}^2) \langle \sigma^{\mathbf{l}^1}_{i_1'j_1'} \sigma^{\mathbf{l}^2}_{i_2'j_2'}\rangle_y\Big) \\
&= \sum_{a_1=1}^{m_1}\dots\sum_{a_t=1}^{m_t} \Big(\frac{\beta}{\sqrt{N}}\Big)^{|\mathbf{a}|} \sum_{q_1=1}^{m_1!}\dots\sum_{q_t=1}^{m_t!} \gamma^1_{a_1,q_1}\dots\gamma^t_{a_t,q_t} \prod_{r_1=1}^{m_1} h^{(b^1_{a_1,q_1,r_1})}(g_{i_1'j_1'}) \dots \prod_{r_t=1}^{m_t} h^{(b^t_{a_t,q_t,r_t})}(g_{i_t'j_t'}) \\
&\qquad \times \sum_{\mathbf{l}^1\in\mathcal{L}_{a_1}} \dots \sum_{\mathbf{l}^t\in\mathcal{L}_{|\mathbf{a}|-a_t,a_t}} c^1_{a_1}(\mathbf{l}^1)\dots c^t_{a_t}(\mathbf{l}^t) \langle \sigma^{\mathbf{l}^1}_{i_1'j_1'} \dots \sigma^{\mathbf{l}^t}_{i_t'j_t'}\rangle_y \\
&= \sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}} \Big(\frac{\beta}{\sqrt{N}}\Big)^{|\mathbf{a}|} \sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}} \gamma_{\mathbf{a},\mathbf{q}} h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'}) \sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}} c_{|\mathbf{a}|}(\mathbf{l}) \langle \boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y,
\end{aligned}$$

where $\mathbf{a} = (a_1,\dots,a_t), \mathbf{q} = (q_1,\dots,q_t), \mathbf{l} = (\mathbf{l}^1,\dots,\mathbf{l}^t), \gamma_{\mathbf{a},\mathbf{q}} = \prod_{s=1}^t \gamma^s_{a_s,q_s}$, the sets $\mathcal{L}_a, \mathcal{L}_{n,a}$ are defined in Lemmas 3.1 and 3.2, and for $\mathbf{l}^s = (\mathbf{l}^s_1,\dots,\mathbf{l}^s_{a_s}) \in \mathcal{L}_{|\mathbf{a}|_s,a_s}$ with $|\mathbf{a}|_s = a_1+\dots+a_{s-1}$ ($|\mathbf{a}|_1 = 0$), we denoted $\sigma^{\mathbf{l}^s}_{i_s'j_s'} = \sigma^{\mathbf{l}^s_1}_{i_s'}\sigma^{\mathbf{l}^s_1}_{j_s'}\dots\sigma^{\mathbf{l}^s_{a_s}}_{i_s'}\sigma^{\mathbf{l}^s_{a_s}}_{j_s'}$. Notice that $\mathbf{l} = (\mathbf{l}^1,\dots,\mathbf{l}^t) \in \mathcal{L}_{|a|}$, since $\mathbf{l}^s \in \mathcal{L}_{|a|_s,a_s}$ for all $1\le s\le t$.

Here for the second line we used Lemma 3.1, and for the third line we applied Lemma 3.2 to $L = \sigma^{\mathbf{l}^1_1}_{i_1'}\sigma^{\mathbf{l}^1_1}_{j_1'}\dots\sigma^{\mathbf{l}^1_{a_1}}_{i_1'}\sigma^{\mathbf{l}^1_{a_1}}_{j_1'}$, which is a function of a_1 configurations. The next equation is obtained by applying sequentially Lemma 3.2.

Since $\sum_{r_s=1}^{m_s} b^s_{a_s,q_s,r_s} = m_s$ if $\gamma^s_{a_s,q_s} = 1$ for all $1\le s\le t$, we have $b_{\mathbf{a},\mathbf{q}} \in \mathcal{B}_{\mathbf{m}}$. Moreover, we note that $|c^1_{a_1}(\mathbf{l}^1)| \le (a_1-1)!$ and for any $2\le s\le t$ and $\mathbf{l}^s\in\mathcal{L}_{|\mathbf{a}|_s,a_s}$,

$$c^s_{a_s}(\mathbf{l}^s) \le \frac{(|\mathbf{a}|_s + a_s - 1)!}{(|\mathbf{a}|_s-1)!} = \frac{(|\mathbf{a}|_{s+1}-1)!}{(|\mathbf{a}|_s-1)!}.$$

Hence, for all tuples $\mathbf{a}\in\mathcal{A}$ and $\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}$,

$$c_{|\mathbf{a}|}(\mathbf{l}) = \prod_{s=1}^t c^s_{a_s}(\mathbf{l}^s) \le (|\mathbf{a}|_{t+1}-1)! = (|\mathbf{a}|-1)!.$$

The proof of the lemma is completed. ☐

3.2 Proof of Proposition 3

First, we recall the replica trick which will be useful in computing the derivative of the free energy. Let $\{\tilde{g}_{ij}\}_{1\le i<j\le N}$ be an independent copy of $\{g_{ij}\}_{1\le i<j\le N}$, and consider the Gibbs measure, denoted by $G_{\tilde{y}}$, corresponding to the disordered environment $\tilde{y} = (h(\tilde{g}_{ij}))_{1\le i<j\le N}$. Let $\tilde{\sigma}^{l_1}, \tilde{\sigma}^{l_2}, \dots, \tilde{\sigma}^{l_{|\mathbf{a}|}}$ be i.i.d. configurations drawn according to the Gibbs measure induced by the disorders $h(\tilde{g})$. Then for any $\mathbf{a} \in \mathcal{A}_{\mathbf{m}}$ with $\mathbf{m} \in \mathcal{M}_t$ and $1 \le t \le k$,

$$\mathbb{E}\Big[\langle \sigma^{l_1}_{i'_1}\sigma^{l_1}_{j'_1}\dots\sigma^{l_{|\mathbf{a}|}}_{i'_t}\sigma^{l_{|\mathbf{a}|}}_{j'_t}\rangle_y^2\Big] = \mathbb{E}\Big[\langle \sigma^{l_1}_{i'_1}\tilde{\sigma}^{l_1}_{i'_1}\sigma^{l_1}_{j'_1}\tilde{\sigma}^{l_1}_{j'_1}\dots\sigma^{l_{|\mathbf{a}|}}_{i'_t}\tilde{\sigma}^{l_{|\mathbf{a}|}}_{i'_t}\sigma^{l_{|\mathbf{a}|}}_{j'_t}\tilde{\sigma}^{l_{|\mathbf{a}|}}_{j'_t}\rangle_{y,\tilde{y}}\Big], \tag{39}$$

where $\langle\cdot\rangle_{y,\tilde{y}}$ is Gibbs average respect to Gibbs product measure $(G_y \times G_{\tilde{y}})^{\bigotimes|\mathbf{a}|}$. (Here, for simplicity the superscript $|\mathbf{a}|$ is omitted in $\langle\cdot\rangle_{y,\tilde{y}}$).

Let us consider the tuple as in (37):

$$T = \{i_1j_1,\dots,i_kj_k\} = \{\underbrace{i'_1j'_1,\dots,i'_1j'_1}_{m_1 \text{ times}},\dots,\underbrace{i'_tj'_t,\dots,i'_tj'_t}_{m_t \text{ times}}\}, \qquad \{i'_1j'_1,\dots,i'_tj'_t\} \in E^t_{\neq}.$$

Using Lemma 3.3 and the Cauchy-Bunyakovsky-Schwarz inequality, we have

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\frac{\partial F_y}{\partial g_{i_1j_1}\dots\partial g_{i_kj_k}}\Big]\Big)^2\\
&= \Big(\mathbb{E}\Big[\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\Big(\frac{\beta}{\sqrt{N}}\Big)^{|\mathbf{a}|}\sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\gamma_{\mathbf{a},\mathbf{q}}h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}c_{|\mathbf{a}|}(\mathbf{l})\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y\Big]\Big)^2\\
&\le |\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\Big(\frac{\beta}{\sqrt{N}}\Big)^{2|\mathbf{a}|}\Big(\mathbb{E}\Big[\sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\gamma_{\mathbf{a},\mathbf{q}}h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}c_{|\mathbf{a}|}(\mathbf{l})\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y\Big]\Big)^2\\
&\le |\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\Big(\frac{\beta}{\sqrt{N}}\Big)^{2|\mathbf{a}|}\mathbb{E}\Big[\Big(\sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\gamma_{\mathbf{a},\mathbf{q}}h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\Big)^2\Big]\mathbb{E}\Big[\Big(\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}c_{|\mathbf{a}|}(\mathbf{l})\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y\Big)^2\Big].
\end{aligned} \tag{40}$$

Moreover,

$$\begin{aligned}
&\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\sup_{\gamma_{\mathbf{a},\mathbf{q}}}\mathbb{E}\Big[\Big(\sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\gamma_{\mathbf{a},\mathbf{q}}h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\Big)^2\Big] \le |\mathcal{Q}_{\mathbf{m}}|^2\sup_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\mathbb{E}\Big[\Big(h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\Big)^2\Big]\\
= \;& |\mathcal{Q}_{\mathbf{m}}|^2\sup_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\prod_{s=1}^{t}\mathbb{E}\Big[\prod_{r_s=1}^{m_s}\big(h^{(b_{a_s,q_s,r_s})}(g)\big)^2\Big] = |\mathcal{Q}_{\mathbf{m}}|^2\sup_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\mathbb{E}\Big[\big(h^{(b_{\mathbf{a},\mathbf{q}})}(g)\big)^2\Big].
\end{aligned}$$

Hence, by the definition of $p_h(k)$ and the fact that $|\mathcal{Q}_{\mathbf{m}}| \le k^k$ for all $\mathbf{m}$,

$$\begin{aligned}
&\sup_{1\le t\le k}\sup_{\mathbf{m}\in\mathcal{M}_t}\sup_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\sup_{\gamma_{\mathbf{a},\mathbf{q}}}\mathbb{E}\Big[\Big(\sum_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\gamma_{\mathbf{a},\mathbf{q}}h^{(b_{\mathbf{a},\mathbf{q}})}(g_{\mathbf{i}'\mathbf{j}'})\Big)^2\Big]\\
\le\;& k^{2k}\sup_{1\le t\le k}\sup_{\mathbf{m}\in\mathcal{M}_t}\sup_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\sup_{\mathbf{q}\in\mathcal{Q}_{\mathbf{m}}}\sup_{b_{\mathbf{a},\mathbf{q}}\in\mathcal{B}_{\mathbf{m}}}\mathbb{E}\Big[\big(h^{(b_{\mathbf{a},\mathbf{q}})}(g)\big)^2\Big] \le p_h(k).
\end{aligned} \tag{41}$$

By Cauchy-Bunyakovsky-Schwarz inequality again, we have

$$\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}c_{\mathbf{a}}(\mathbf{l})\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y\Big)^2\Big] &\le |\mathcal{L}_{|\mathbf{a}|}|\max_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}c_{|\mathbf{a}|}(\mathbf{l})^2\mathbb{E}\Big[\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y^2\Big]\\
&\le (|\mathbf{a}|!)^3\mathbb{E}\Big[\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y^2\Big],
\end{aligned} \tag{42}$$

since $|\mathcal{L}_{|\mathbf{a}|}| \le |\mathbf{a}|!$ and $|c_{\mathbf{a}}(\mathbf{l})| \le (|\mathbf{a}|-1)!$, by Lemma 3.3. Combining (40), (41) and (42),

$$\Big(\mathbb{E}\Big[\frac{\partial F_y}{\partial g_{i_1j_1}\dots\partial g_{i_kj_k}}\Big]\Big)^2 \le p_h(k)|\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\frac{\beta^{2|\mathbf{a}|}(|\mathbf{a}|!)^3}{N^{|\mathbf{a}|}}\mathbb{E}\Big[\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i}'\mathbf{j}'}\rangle_y^2\Big]. \tag{43}$$

Therefore,

$$
\begin{aligned}
&\sum_{\{i_1j_1,\ldots,i_kj_k\}\in E^k}\Big(\mathbb{E}\Big[\frac{\partial}{\partial g_{i_1j_1}\ldots\partial g_{i_kj_k}}F_y\Big]\Big)^2\\
&\le p_h(k)\sum_{t=1}^{k}\sum_{\mathbf{m}\in\mathcal{M}_t}\sum_{\mathbf{i'j'}\in E^t_{\neq}}|\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\frac{\beta^{2|\mathbf{a}|}(|\mathbf{a}|!)^3}{N^{|\mathbf{a}|}}\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}\mathbb{E}\Big[\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i'j'}}\rangle_y^2\Big]\\
&\le p_h(k)\sum_{t=1}^{k}\sum_{\mathbf{m}\in\mathcal{M}_t}|\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\frac{\beta^{2|\mathbf{a}|}(|\mathbf{a}|!)^3}{N^{|\mathbf{a}|}}\sum_{\mathbf{l}\in\mathcal{L}_{|\mathbf{a}|}}\sum_{\mathbf{i'j'}\in E^t}\mathbb{E}\Big[\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i'j'}}\rangle_y^2\Big]. \qquad (44)
\end{aligned}
$$

Using the replicas trick (39),

$$
\sum_{\mathbf{i'j'}\in E^t}\mathbb{E}\Big[\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i'j'}}\rangle_y^2\Big]=\sum_{\mathbf{i'j'}\in E^t}\mathbb{E}\Big[\langle\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i'j'}}\tilde{\boldsymbol{\sigma}}^{\mathbf{l}}_{\mathbf{i'j'}}\rangle_{y,\tilde y}\Big]. \qquad (45)
$$

For any $\mathbf{l}=(l_1,\ldots,l_{|\mathbf{a}|})\in\mathcal{L}_{|\mathbf{a}|}$,

$$
\begin{aligned}
\sum_{\mathbf{i'j'}\in E^t}\boldsymbol{\sigma}^{\mathbf{l}}_{\mathbf{i'j'}}\tilde{\boldsymbol{\sigma}}^{\mathbf{l}}_{\mathbf{i'j'}}&=\sum_{1\le i'_1,j'_1,\ldots,i'_t,j'_t\le N}\sigma^{l_1}_{i'_1}\tilde\sigma^{l_1}_{i'_1}\sigma^{l_1}_{j'_1}\tilde\sigma^{l_1}_{j'_1}\ldots\sigma^{l_{a_1}}_{i'_1}\tilde\sigma^{l_{a_1}}_{i'_1}\sigma^{l_{a_1}}_{j'_1}\tilde\sigma^{l_{a_1}}_{j'_1}\sigma^{l_{a_1+1}}_{i'_2}\tilde\sigma^{l_{a_1+1}}_{i'_2}\sigma^{l_{a_1+1}}_{j'_2}\tilde\sigma^{l_{a_1+1}}_{j'_2}\\
&\qquad\times\ldots\sigma^{l_{a_1+a_2}}_{i'_2}\tilde\sigma^{l_{a_1+a_2}}_{i'_2}\sigma^{l_{a_1+a_2}}_{j'_2}\tilde\sigma^{l_{a_1+a_2}}_{j'_2}\ldots\sigma^{l_{|\mathbf{a}|}}_{i'_t}\tilde\sigma^{l_{|\mathbf{a}|}}_{i'_t}\sigma^{l_{|\mathbf{a}|}}_{j'_t}\tilde\sigma^{l_{|\mathbf{a}|}}_{j'_t}\\
&=(\sigma^{l_1}\cdot\tilde\sigma^{l_1})^2(\sigma^{l_2}\cdot\tilde\sigma^{l_2})^2\ldots(\sigma^{l_{|\mathbf{a}|}}\cdot\tilde\sigma^{l_{|\mathbf{a}|}})^2, \qquad (46)
\end{aligned}
$$

where

$$
\sigma^{l_s}\cdot\tilde\sigma^{l_s}=\sum_{i=1}^{N}\sigma^{l_s}_i\tilde\sigma^{l_s}_i.
$$

By Hölder's inequality,

$$
\begin{aligned}
&\langle(\sigma^{l_1}\cdot\tilde\sigma^{l_1})^2(\sigma^{l_2}\cdot\tilde\sigma^{l_2})^2\ldots(\sigma^{l_{|\mathbf{a}|}}\cdot\tilde\sigma^{l_{|\mathbf{a}|}})^2\rangle_{y,\tilde y} \qquad (47)\\
&\le\langle(\sigma^{l_1}\cdot\tilde\sigma^{l_1})^{2|\mathbf{a}|}\rangle_{y,\tilde y}^{1/|\mathbf{a}|}\langle(\sigma^{l_2}\cdot\tilde\sigma^{l_2})^{2|\mathbf{a}|}\rangle_{y,\tilde y}^{1/|\mathbf{a}|}\ldots\langle(\sigma^{l_{|\mathbf{a}|}}\cdot\tilde\sigma^{l_{|\mathbf{a}|}})^{2|\mathbf{a}|}\rangle_{y,\tilde y}^{1/|\mathbf{a}|}=\langle(\sigma\cdot\tilde\sigma)^{2|\mathbf{a}|}\rangle_{y,\tilde y},
\end{aligned}
$$

since $\{(\sigma^{l_s},\tilde\sigma^{l_s})\}_{1\le s\le l_{|\mathbf{a}|}}$ are independent copies of $(\sigma,\tilde\sigma)$ given the environment variables $y,\tilde y$. In summary, using (44)–(47), we obtain that

$$
\begin{aligned}
\theta_k(F_y)&=\sum_{\{i_1j_1,\ldots,i_kj_k\}\in E^k}\Big(\mathbb{E}\Big[\frac{\partial}{\partial g_{i_1j_1}\ldots\partial g_{i_kj_k}}F_y\Big]\Big)^2\\
&\le p_h(k)\sum_{t=1}^{k}\sum_{\mathbf{m}\in\mathcal{M}_t}|\mathcal{A}_{\mathbf{m}}|\sum_{\mathbf{a}\in\mathcal{A}_{\mathbf{m}}}\frac{\beta^{2|\mathbf{a}|}(|\mathbf{a}|!)^4}{N^{|\mathbf{a}|}}\mathbb{E}[\langle(\sigma\cdot\tilde\sigma)^{2|\mathbf{a}|}\rangle_{y,\tilde y}], \qquad (48)
\end{aligned}
$$

since $|\mathcal{L}_{|\mathbf{a}|}|=|\mathbf{a}|!$.

Lemma 3.4. *If $y=h(g)\overset{(d)}{=}-h(g)$ then for any measurable function L on Σ_N^2,*

$$
\mathbb{E}[\langle L\rangle_{y,\tilde y}]=\mathbb{E}_{U\times\tilde U}[L], \qquad (49)
$$

where $U,\tilde U$ are the uniform distributions on Σ_N and $\mathbb{E}_{U\times\tilde U}$ is the expectation w.r.t. the product measure of U and $\tilde U$.

Proof. Since

$$
\mathbb{E}[\langle L\rangle_{y,\tilde y}]=\sum_{\sigma,\tilde\sigma\in\Sigma}L(\sigma,\tilde\sigma)\mathbb{E}[G_y(\sigma)]\mathbb{E}[G_{\tilde y}(\tilde\sigma)],
$$

to assert (49), it is enough to show that for any σ, τ in Σ_N,

$$\mathbb{E}[G_y(\sigma)] = \mathbb{E}[G_y(\tau)]. \tag{50}$$

Fix $\sigma, \tau \in \Sigma_N$, we set $x_{ij} := \dfrac{y_{ij}\tau_i\tau_j}{\sigma_i\sigma_j}$ and observe that

$$\begin{aligned}\mathbb{E}[G_y(\tau)] &= \mathbb{E}\left[\frac{\exp\left(\sum_{ij\in E}(y_{ij}\tau_i\tau_j)\right)}{\sum_{\omega\in\Sigma_N}\exp\left(\sum_{ij\in E} y_{ij}\omega_i\omega_j\right)}\right]\\ &= \mathbb{E}\left[\frac{\exp\left(\sum_{ij\in E}(x_{ij}\sigma_i\sigma_j\right)}{\sum_{\omega\in\Sigma_N}\exp\left(\sum_{ij\in E}\frac{x_{ij}\sigma_i\sigma_j}{\tau_i\tau_j}\omega_i\omega_j\right)}\right].\end{aligned}$$

We now define $\omega' \in \Sigma_N$ by setting $\omega'_i = \frac{\sigma_i}{\tau_i}\omega_i$ for $1 \le i \le N$. Then the above equation can be rewritten as

$$\mathbb{E}[G_y(\tau)] = \mathbb{E}\left[\frac{\exp\left(\sum_{ij\in E}(x_{ij}\sigma_i\sigma_j\right)}{\sum_{\omega'\in\Sigma_N}\exp\left(\sum_{ij\in E} x_{ij}\omega'_i\omega'_j\right)}\right] = \mathbb{E}[G_y(\sigma)],$$

since $y_{ij} \overset{(d)}{=} x_{ij}$ (as $y_{ij} = \pm x_{ij}$ and $h(g_{ij}) \overset{(d)}{=} -h(g_{ij})$). Hence, (50) is proved and the result of this lemma follows. □

By Lemma 3.4,

$$\mathbb{E}\left[\langle(\sigma\cdot\tilde{\sigma})^{2|\mathbf{a}|}\rangle_{y,\tilde{y}}\right] = \mathbb{E}_{U\times\tilde{U}}\left[(\sigma\cdot\tilde{\sigma})^{2|\mathbf{a}|}\right].$$

Under $U \times \tilde{U}$, the variable $\sigma\cdot\tilde{\sigma}$ is simply the sum of N i.i.d. symmetric $\{-1, 1\}$-valued random variables. Hence, using Hoeffding's inequality,

$$\mathbb{E}_{U\times\tilde{U}}\left[(\sigma\cdot\tilde{\sigma})^{2|\mathbf{a}|}\right] \le |\mathbf{a}|^{|\mathbf{a}|}N^{|\mathbf{a}|}.$$

Combining the last two equations with (48), we yield that for all $k \ge 1$

$$\theta_k(F_y) \le p_h(k)\sum_{t=1}^{k}\sum_{\mathbf{m}\in\mathcal{M}_t}|\mathcal{A}_\mathbf{m}|\sum_{\mathbf{a}\in\mathcal{A}_\mathbf{m}}\beta^{2|\mathbf{a}|}(|\mathbf{a}|!)^4|\mathbf{a}|^{|\mathbf{a}|} \le (1+\beta^{2k})k^{8k+1}p_h(k),$$

since $|\mathcal{A}_\mathbf{m}| \le k^k$, $|\mathcal{M}_t| \le k^k$ and $|\mathbf{a}| \le k$. This completes the proof of Proposition 3. □

Acknowledgement. The work of V. H. Can is supported by the Singapore Ministry of Education Academic Research Fund Tier 2 Grant MOE2018-T2-2-076. V. Q. Nguyen and H. S. Vu are supported by International Center for Research and Postgraduate Training in Mathematics, Institute of Mathematics, Vietnam Academy of Science and Technology under grant number ICRTM03-2020.02.

References

[1] Auffinger A. and W. -K. Chen. *Universality of chaos and ultrametricity in mixed p-spin models*. Commun. Pure. Appl. Math. **69**, 2107-2130 (2016).

[2] Benaim, M. and Rossignol, R. *Exponential concentration for first passage percolation through modified Poincaré inequalities.* Ann. Inst. H. Poincaré Probab. Statist. **44**, 544-573 (2008).

[3] Benjamini, I., Kalai, G. and Schramm, O. *First passage percolation has sublinear distance variance*. Annals of Probab. **31**, 1970-1978 (2003).

[4] Bernsteina M, Damron. M, Greenwood. T. *Sublinear variance in Euclidean first-passage percolation*. Stoch. Process. Appl. **130** 5060–5099 (2020).

[5] Carmona, P. and Hu, Y. *Universality in Sherrington-Kirkpatrick's spin glass model*. Ann. Inst. Henri Poincaré Probab. Stat. **42**, 215-222 (2006).

[6] Can, H. V. and Nakajima, S. *First passage time of the frog model has a sublinear variance*. Electron. J. Probab. **24**, 1-27 (2019).

[7] Chatterjee, S. *Chaos, concentration and multiple valleys*. Preprint (2008). Available at arXiv:0810.4221.

[8] Chatterjee, S. *Disorder chaos and multiple valleys in spin glasses*. Preprint (2009). Available at arXiv:0907.3381.

[9] Chatterjee, S. *Superconcentration and related topics*. Springer Monographs in Mathematics. Springer, Cham. (2014).

[10] Chatterjee, S. *Superconcentration in surface growth*. Preprint (2021). Available at arXiv:2103.09199.

[11] Chen, Y. T. *Universality of Ghirlanda-Guerra identities and spin distributions in mixed p-spin models*. Ann. Inst. H. Poincaré Probab. Stat. **55**, 528-550 (2019).

[12] Damron, M., Hanson, J. and Sosoe, P. *Sublinear variance in first-passage percolation for general distributions*. Probability Theory and Related Fields. **163**, 223-258 (2015).

[13] Guerra, F. and Toninelli, L. *The thermodynamic limit in mean field spin glass models*. Comm. Math. Phys., **30**, 71-79 (2002).

[14] Panchenko, D. *The Sherrington-Kirkpatrick Model*. Springer-Verlag New York (2013).

[15] Parisi. G. *A sequence of approximate solutions to the S-K model for spin glasses*. J. Phys. A. **13** (1980).

[16] Talagrand, M. *Gaussian averages, Bernoulli averages, and Gibbs measures*. Random Structures Algorithms. **21**, 197-204 (2002).

[17] Talagrand, M. *The Parisi formula*. Ann. of Math. **163**, 221-263 (2006).