DISORDER CHAOS IN SOME DILUTED SPIN GLASS MODELS

BY WEI-KUO CHEN AND DMITRY PANCHENKO

University of Minnesota and University of Toronto

We prove disorder chaos at zero temperature for three types of diluted models with large connectivity parameter: K-spin antiferromagnetic Ising model for even $K \geq 2$, K-spin spin glass model for even $K \geq 2$, and random K-sat model for all $K \geq 2$. We show that modifying even a small proportion of clauses results in near maximizers of the original and modified Hamiltonians being nearly orthogonal to each other with high probability. We use a standard technique of approximating diluted models by appropriate fully connected models and then apply disorder chaos results in this setting, which include both previously known results as well as new examples motivated by the random K-sat model.

1. Introduction. We begin by defining three types of diluted spin glass models that will be considered in this paper. Let $K \geq 2$ be a fixed integer. A random clause with K variables is a random function $\theta(\sigma_1, \ldots, \sigma_K)$ on $\{-1, +1\}^K$. We will consider the following three examples.

EXAMPLE 1 (K-spin antiferromagnetic model). The function θ is given by

$$\theta(\sigma_1, \ldots, \sigma_K) = -\sigma_1 \cdots \sigma_K, \tag{1}$$

so in this case it is nonrandom.

EXAMPLE 2 (K-spin spin glass). The random function θ is given by

$$\theta(\sigma_1, \ldots, \sigma_K) = J\sigma_1 \cdots \sigma_K, \tag{2}$$

where J is a Rademacher random variable with $\mathbb{P}(J = \pm 1) = 1/2$.

EXAMPLE 3 (K-sat model). The random function θ is given by

$$\theta(\sigma_1, \ldots, \sigma_K) = -\prod_{k \leq K} \frac{1 + J_k \sigma_k}{2}, \tag{3}$$

where $(J_k)_{k \geq 1}$ are i.i.d. Rademacher random variables.

We choose random variables J to be Rademacher only for simplicity, and one can also consider other symmetric distributions. We will denote by θ_j independent copies of the function θ for various indices j. In Example 1, θ is nonrandom so $\theta_j = \theta$, and in Examples 2 and 3,

$$\theta_j(\sigma_1, \ldots, \sigma_K) = J_j \sigma_1 \cdots \sigma_K \quad \text{and} \quad \theta_j(\sigma_1, \ldots, \sigma_K) = - \prod_{k \leq K} \frac{1 + J_{k,j} \sigma_k}{2}$$

with i.i.d. copied J_j or $J_{k,j}$ of J. Given a parameter $\lambda > 0$, called the connectivity parameter, the Hamiltonian of the models we will be studying is defined by

$$(4) \qquad H_\lambda(\sigma) = \sum_{j \leq \pi(\lambda N)} \theta_j(\sigma_{i_{1,j}}, \ldots, \sigma_{i_{K,j}}),$$

where $\pi(\lambda N)$ is a Poiss(λN) random variable with the mean λN, and the coordinate indices $i_{k,j}$ are independent for different pairs (k, j) and are chosen uniformly from $\{1, \ldots, N\}$. In Example 1, we will be interested only in balanced configurations belonging to the set

$$(5) \qquad V = \left\{ \sigma : \sum_{i \leq N} \sigma_i = 0 \text{ if } N \in 2\mathbb{N} \text{ and } \sum_{i \leq N} \sigma_i = 1 \text{ if } N \in 2\mathbb{N} + 1 \right\},$$

while in Examples 2 and 3 the set will be the entire hypercube $V = \{-1, +1\}^N$. Our main result below will describe a chaotic behaviour of near maximizers of this Hamiltonian under resampling a positive proportion of clauses, even arbitrarily small.

Example 1 with $K = 2$ is related to the maximum bisection of the sparse Erdős–Rényi random graph $G(N, 2\lambda/N)$, since we can represent a bisection by a balanced configurations σ and write the number of edges between the two groups (up to the usual Poisson approximation) as

$$\sum_{j \leq \pi(\lambda N)} \mathrm{I}(\sigma_{i_{1,j}} \neq \sigma_{i_{2,j}}) = \frac{\pi(\lambda N)}{2} - \sum_{j \leq \pi(\lambda N)} \sigma_{i_{1,j}} \sigma_{i_{2,j}} = \frac{\pi(\lambda N)}{2} + H_\lambda(\sigma).$$

Example 2 is a diluted version of the K-spin Sherrington–Kirkpatrick spin glass model [28], and Example 3 corresponds to the random K-sat model and, in particular, finding the assignment of variables (σ_i) that maximizes the number of satisfied clauses, which for a given clause means that at least one $\sigma_k = -J_k$ for $1 \leq k \leq K$.

For the rest of the paper, we fix a correlation/proportion parameter

$$(6) \qquad\qquad\qquad\qquad t \in (0, 1),$$

and consider two copies $H_\lambda^1(\sigma)$ and $H_\lambda^2(\sigma)$ of (4) defined in two different ways as follows.

(*Resampling clauses.*) In this case, the two Hamiltonians $H_\lambda^1(\sigma)$ and $H_\lambda^2(\sigma)$ will have Poiss$(t\lambda N)$ common clauses and two independent Poiss$((1-t)\lambda N)$ independently generated clauses, which means that we resample both indices of variables

and random signs. For instance, in Example 1 the only randomness is in the choice of the clause indices so, for $\ell = 1, 2$, we define

$$(7) \qquad H_\lambda^\ell(\sigma) = - \sum_{j \leq \pi(t\lambda N)} \sigma_{i_{1,j}} \cdots \sigma_{i_{K,j}} - \sum_{j \leq \pi_\ell((1-t)\lambda N)} \sigma_{i_{\ell,1,j}} \cdots \sigma_{i_{\ell,K,j}},$$

where $\pi(t\lambda N)$, $\pi_1((1-t)\lambda N)$ and $\pi_2((1-t)\lambda N)$ are independent Poisson random variables with the means $t\lambda N$ and $(1-t)\lambda N$ and all indices $i_{k,j}$ and $i_{\ell,k,j}$ are independent and uniform on $\{1, \ldots, N\}$. In Examples 2 and 3, we can also resample only random signs J without resampling indices, as follows.

(*Resampling random signs.*) In this case, the number of clauses, $\pi(\lambda N)$, will be the same, as well as their indices, and only random signs will be resampled. In the K-spin spin glass, we will replace the random variable J_j in each clause by two correlated copies J_j^1 and J_j^2 such that $\mathbb{E}J_j^1 J_j^2 = t$. In the K-sat model, we will consider two versions of the correlated copies $H_\lambda^1(\sigma)$ and $H_\lambda^2(\sigma)$.

(a) In the first version, independently for each clause θ_j, with probability $1 - t$ we resample all random signs $J_{1,j}^1, \ldots, J_{K,j}^1$ simultaneously to produce $J_{1,j}^2, \ldots, J_{K,j}^2$. In this case, $\mathbb{E}J_{k,j}^1 J_{k,j}^2 = t$ but the pairs $(J_{k,j}^1, J_{k,j}^2)$ within the same clause are correlated.

(b) In the second version, we will simply replace each $J_{k,j}$ in each clause by two copies such that $\mathbb{E}J_{k,j}^1 J_{k,j}^2 = t$. In other words, we resample each $J_{k,j}^1$ with probability $1 - t$ to produce $J_{k,j}^2$, but $(J_{k,j}^1, J_{k,j}^2)$ are independent for all different pairs (k, j).

The main result of this paper is that, for all decouplings of the Hamiltonian described above and for large connectivity λ, with high probability all near maximizers are nearly orthogonal to each other. For technical reasons related to Theorem 2 below, we will assume from now on that $K \geq 2$ is even in Examples 1 and 2, and arbitrary $K \geq 2$ in Example 3.

THEOREM 1. *For any $\varepsilon, t \in (0, 1)$ there exists small enough $\eta > 0$ such that for large enough λ the following holds for large enough N with probability at least $1 - Le^{-N\eta^2/L}$: for any configurations $\sigma^1, \sigma^2 \in V$ that nearly maximize the corresponding Hamiltonian,*

$$(8) \qquad \frac{1}{N} H_\lambda^\ell(\sigma^\ell) \geq \frac{1}{N} \max_{\sigma \in V} H_\lambda^\ell(\sigma) - \eta\sqrt{\lambda} \qquad \text{for } \ell = 1, 2,$$

the overlap $R_{1,2} = N^{-1} \sum_{i \leq N} \sigma_i^1 \sigma_i^2$ between them satisfies $|R_{1,2}| \leq \varepsilon$.

In other words, the set of near maximizers is chaotic under resampling even a small proportion of clauses. Here the constant L depends only on K, $L = L(K) > 0$, and we will see in the proof that one can take $\lambda \geq L\eta^{-6}$ also for some large enough constant $L = L(K) > 0$. The definition of near maximizers is taken on

the scale $\sqrt{\lambda}$, which corresponds to the scale of the maximum [see (11) and (12) below] except that in the random K-sat model one needs to subtract $-\lambda/2^K$ corresponding to the nonrandom part of the clauses in (3).

Let us mention one standard consequence of Theorem 1—the existence of exponentially many in N near maximizers of the Hamiltonian H_λ in Examples 1 and 2 (and in Example 3 after subtracting $-\lambda/2^K$) which are all nearly orthogonal to each other. To see this, consider the case of resampling clauses and notice that, by (11), the normalized maximum over Poiss$((1 - t)\lambda N)$ clauses will be of order $\sqrt{(1 - t)\lambda}$ so, for t close to 1, the maximizer of the second Hamiltonian can be considered a near maximizer of the first one. Since we have exponential control of all probabilities, we can resample the Hamiltonian exponentially many times to find exponentially many near maximizers that are all nearly orthogonal to each other.

The proof of Theorem 1 will be based on some known as well as new examples of disorder chaos at zero temperature for mixed p-spin models, and a standard comparison technique of the diluted models with large connectivity with appropriate mixed p-spin models using the Guerra–Toninelli interpolation from [14]. This technique has been used in various settings in [11, 15, 23, 27] to show that the diluted K-spin model and K-spin spin glass can be approximated by the fully connected pure K-spin spin glass model with the Hamiltonian

$$
(9) \qquad H(\sigma) = \frac{1}{N^{(K-1)/2}} \sum_{1 \le i_1,\ldots,i_K \le N} g_{i_1,\ldots,i_K} \sigma_{i_1} \cdots \sigma_{i_K}
$$

and the K-sat model can be approximated by the following mixed p-spin Hamiltonian

$$
(10) \qquad H(\sigma) = \sum_{p=1}^{K} \sqrt{\binom{K}{p} \frac{1}{N^{p-1}}} \sum_{1 \le i_1,\ldots,i_p \le N} g_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p},
$$

where the coefficients $(g_{i_1,\ldots,i_p})$ are standard Gaussian random variables independent for all $p \ge 1$ and all indices $(i_1, \ldots, i_p)$. From now on, whenever H_λ and H appear in the same formula, we assume without explicitly mentioning again that Examples 1 and 2 are paired with (9) and Example 3 is paired with (10). For the K-spin antiferromagnetic Ising model from Example 1 and K-spin spin glass from Example 2, it was shown in [11, 27] that

$$
(11) \qquad \frac{1}{N} \mathbb{E} \max_{\sigma \in V} H_\lambda(\sigma) = \sqrt{\lambda} \frac{1}{N} \mathbb{E} \max_{\sigma \in V} H(\sigma) + O(\lambda^{1/3})
$$

as $\lambda \to \infty$, uniformly in N. For the K-sat model, it was shown in [23] that

$$
(12) \qquad \frac{1}{N} \mathbb{E} \max_{\sigma \in V} H_\lambda(\sigma) = -\frac{\lambda}{2^K} + \frac{\sqrt{\lambda}}{2^K} \frac{1}{N} \mathbb{E} \max_{\sigma \in V} H(\sigma) + O(\lambda^{1/3})
$$

as $\lambda \to \infty$, uniformly in N (which was obtained earlier in [19] by the nonrigorous replica method). All the above results were stated with the error $o(\sqrt{\lambda})$ instead of $O(\lambda^{1/3})$; however, a (not so) careful inspection of the arguments as well as the proof below will make it clear that the error term can be chosen to be $O(\lambda^{1/3})$. Upper and lower bounds on the factor in front of $\sqrt{\lambda}$ were obtained earlier in [10]. The main motivation for the above approximation results was due to the fact that the limit of $N^{-1}\mathbb{E}\max_{\sigma \in V} H(\sigma)$ on the right-hand side is well known and is given by the Parisi formula [25, 26] proved for various mixed p-spin models in [29] and [21, 22] and extended to zero temperature in [7]. This zero temperature formula appears below in the equation (30).

To prove Theorem 1, we will connect by a similar technique two resampled diluted systems coupled by some overlap constraint to two fully connected systems correlated in an appropriate way as follows. The covariance of the Gaussian Hamiltonians above is given by

$$\tag{13} \mathbb{E}H(\sigma^1)H(\sigma^2) = N\xi(R_{1,2}),$$

where in the case of (9) and (10) correspondingly,

$$\tag{14} \xi(s) = s^K \quad \text{and} \quad \xi(s) = \sum_{p=1}^{K}\binom{K}{p}s^p = (1+s)^K - 1.$$

For $t \in (0, 1)$, we will consider two correlated copies H^1, H^2 of these Hamiltonians such that either

$$\tag{15} \mathbb{E}H^1(\sigma^1)H^2(\sigma^2) = Nt\xi(R_{1,2})$$

or

$$\tag{16} \mathbb{E}H^1(\sigma^1)H^2(\sigma^2) = N\xi(tR_{1,2}).$$

The first type of correlation (15) will be used to approximate correlated copies of the diluted models in all cases except one – the K-sat model with the resampling as in (b) above, in which case the second type of correlation (16) will be used. Once this approximation is established, at the core of the proof of Theorem 1 will be the following disorder chaos result for fully connected models.

THEOREM 2. *Assume that one of the following conditions is satisfied*:

  (i) *H is given by (9) with even K and (15) holds*;
  (ii) *H is given by (10) for any $K \geq 2$ and either (15) or (16) holds.*

For any $\varepsilon, t \in (0, 1)$, there exist $\eta > 0$ such that, for large enough N,

$$\tag{17} \frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}(H^1(\sigma^1) + H^2(\sigma^2)) \leq \frac{2}{N}\mathbb{E}\max_{\sigma} H(\sigma) - \eta.$$

The case of the K-spin model for even K was already included in the disorder chaos result in [8], which in fact covers mixed p-spin models with even p. The other two cases corresponding to the Hamiltonian (10) will be proved in this paper and, since the proof is very similar to [8], only necessary modifications will be detailed. For a number of earlier results about disorder chaos in various settings, see [1–5, 9].

REMARK 1. The proof of Theorem 2 applies to general covariance structures under the following technical assumptions. Let $\xi_0(s) := t\xi(s)$ in the case (15) and $\xi_0(s) := \xi(ts)$ in the case (16). If $\xi(s)$ and $\xi_0(s)$ are both convex functions on $[-1, 1]$ that satisfy $\xi_0''(s) < \xi''(|s|)$ on $[-1, 1] \setminus \{0\}$ and $\zeta_+(s)$ and $\zeta_-(s)$ defined in the equation (36) in Lemma 2 below are nondecreasing on $(0, 1]$, then chaos in disorder (17) remains true.

2. Proof of Theorem 1. We will first show how Theorem 2 implies Theorem 1. The proofs are very similar in all cases, so we will only detail one case when only random signs are resampled and one case when clauses are resampled.

2.1. *Resampling random signs.* First, we will describe the analogue of the Guerra–Toninelli interpolation in the case of resampling of random signs in the K-spin spin glass and random K-sat model. In the next susbsection, we will describe the interpolation in the case of resampling clauses.

For $s \in [0, 1]$, let us consider the interpolating Hamiltonian

$$(18) \qquad H(s, \sigma^1, \sigma^2) = \sum_{\ell=1}^{2} (\delta H^\ell_{\lambda(1-s)}(\sigma^\ell) + \sqrt{s}\beta H^\ell(\sigma^\ell)).$$

Here the correlated Hamiltonians $H^1_{\lambda(1-s)}$ and $H^2_{\lambda(1-s)}$ are defined in the same way as H^1_λ and H^2_λ in the Introduction, only with the connectivity parameter λ replaced by $\lambda(1-s)$, while H^1 and H^2 are defined through (13) to (16). The inverse temperature parameters $\delta > 0$ and $\beta > 0$ will be chosen later. Let

$$(19) \qquad \varphi(s) = \frac{1}{N}\mathbb{E}\log \sum_{|R_{1,2}|>\varepsilon} \exp H(s, \sigma^1, \sigma^2)$$

be the interpolating free energy of these correlated systems coupled by the overlap constraint $|R_{1,2}| > \varepsilon$. We will now compute the derivative $\varphi'(s) = \mathrm{I} + \mathrm{II}$ as a sum of two terms coming from the Gaussian integration by parts and Poisson integration by parts. Let us denote by $\langle \cdot \rangle_s$ the average with respect to the Gibbs measure on $\{(\sigma^1, \sigma^2) \in V^2 : |R_{1,2}| > \varepsilon\}$,

$$(20) \qquad G_s(\sigma^1, \sigma^2) = \frac{\exp H(s, \sigma^1, \sigma^2)}{\sum_{|R_{1,2}|>\varepsilon} \exp H(s, \sigma^1, \sigma^2)},$$

corresponding to the Hamiltonian $H(s, \sigma^1, \sigma^2)$. Recall that the K-spin spin glass and random K-sat model are defined on $V = \{-1, +1\}^N$. Let us denote the i.i.d. replicas from this measure by $(\sigma^{\ell,1}, \sigma^{\ell,2})$ for $\ell \geq 1$ and let us denote

$$R^{j,j'}_{\ell,\ell'} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{\ell,j} \sigma_i^{\ell',j'}.$$

Taking the derivative in $\sqrt{s}$ in the second term in (18) and using standard Gaussian integration by parts (see, e.g., [31] or Section 1.2 in [20]) gives

$$\mathrm{I} = \frac{\beta^2}{2} \left(2\xi(1) + 2\mathbb{E}\langle t\xi(R^{1;2}_{1,1}) \rangle_s - \mathbb{E}\langle \xi(R^{1;1}_{1,2}) + \xi(R^{2;2}_{1,2}) + t\xi(R^{1;2}_{1,2}) + t\xi(R^{2;1}_{1,2}) \rangle_s \right)$$

in the case when the correlation of H^1 and H^2 is given by (15) and

$$\mathrm{I} = \frac{\beta^2}{2} \left(2\xi(1) + 2\mathbb{E}\langle \xi(tR^{1;2}_{1,1}) \rangle_s - \mathbb{E}\langle \xi(R^{1;1}_{1,2}) + \xi(R^{2;2}_{1,2}) + \xi(tR^{1;2}_{1,2}) + \xi(tR^{2;1}_{1,2}) \rangle_s \right)$$

in the case when the correlation of H^1 and H^2 is given by (16). The rest of the calculation below is quite similar to the one in the proof of the Franz-Leone upper bound for the free energy in diluted models in [12, 24]. To differentiate with respect to s in the Poisson random variable $\pi(\lambda(1-s)N)$ in the first term in (18) we use that

$$\frac{d}{ds} \mathbb{E} f(\pi(s)) = \mathbb{E} f(\pi(s) + 1) - \mathbb{E} f(\pi(s))$$

for a Poisson random variables $\pi(s)$ with the mean s. Therefore, the derivative of $\varphi(s)$ with respect to s in $\pi(\lambda(1-s)N)$ equals

$$\mathrm{II} = -\lambda \left(\mathbb{E} \log \sum_{|R_{1,2}|>\varepsilon} \exp H^+(s, \sigma^1, \sigma^2) - \mathbb{E} \log \sum_{|R_{1,2}|>\varepsilon} \exp H(s, \sigma^1, \sigma^2) \right),$$

where $H^+(s, \sigma^1, \sigma^2)$ differs from $H(s, \sigma^1, \sigma^2)$ by the addition of one more clause in each of the correlated diluted Hamiltonians,

$$H^+(s, \sigma^1, \sigma^2) = H^+(s, \sigma^1, \sigma^2) + \delta\theta^1(\sigma^1_{i_1}, \ldots, \sigma^2_{i_K}) + \delta\theta^2(\sigma^2_{i_1}, \ldots, \sigma^2_{i_K}),$$

and these clauses are independent of $H(s, \sigma^1, \sigma^2)$ and are given by

$$\theta^\ell(\sigma^\ell_{i_1}, \ldots, \sigma^\ell_{i_K}) = -\prod_{k \leq K} \frac{1 + J^\ell_k \sigma^\ell_{i_k}}{2},$$

where the random signs J^ℓ_k are correlated as in the case (a) or case (b), that is, they are resampled with probability $1 - t$ either independently or simultaneously within this one clause. Clearly, we can rewrite the derivative above as

$$\mathrm{II} = -\lambda \mathbb{E} \log \langle \exp \delta\theta^1(\sigma^1_{i_1}, \ldots, \sigma^1_{i_K}) \exp \delta\theta^2(\sigma^2_{i_1}, \ldots, \sigma^2_{i_K}) \rangle_s.$$

Since $\theta \in \{-1, 0\}$, we can write $\exp \delta\theta = 1 + (1 - e^{-\delta})\theta$ and

$$\exp \delta\theta^1 \exp \delta\theta^2 = 1 - (1 - e^{-\delta})\Delta(\sigma^1, \sigma^2),$$

where

$$\Delta(\sigma^1, \sigma^2) = \prod_{k \leq K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} + \prod_{k \leq K} \frac{1 + J_k^2 \sigma_{i_k}^2}{2}$$

$$- (1 - e^{-\delta}) \prod_{k \leq K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2}.$$

Since $0 \leq \Delta(\sigma^1, \sigma^2) \leq 1 + e^{-\delta}$ and $(1 - e^{-\delta})\Delta(\sigma^1, \sigma^2) \leq 1 - e^{-2\delta}$, we can express the logarithm above using the Taylor series as

$$\mathrm{II} = \lambda \sum_{n \geq 1} \frac{(1 - e^{-\delta})^n}{n} \mathbb{E}\langle \Delta(\sigma^1, \sigma^2)\rangle_s^n.$$

Using replicas, we can represent

$$\mathbb{E}\langle \Delta(\sigma^1, \sigma^2)\rangle_s^n = \mathbb{E}\Big\langle \prod_{\ell \leq n} \Delta(\sigma^{\ell,1}, \sigma^{\ell,2})\Big\rangle_s = \mathbb{E}\Big\langle \mathbb{E}' \prod_{\ell \leq n} \Delta(\sigma^{\ell,1}, \sigma^{\ell,2})\Big\rangle_s,$$

where $\mathbb{E}'$ is the expectation with respect to the randomness J_k^1, J_k^2 and i_k of the clauses θ^1 and θ^2, which is independent of the randomness in $\langle \cdot \rangle_s$. To compute this expectation, let us first note that the expectation $\mathbb{E}_J$ in the random variables J_k^ℓ satisfies

$$\mathbb{E}_J \prod_{k \leq K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^1 \sigma_{i_k}^2}{2} = \mathbb{E}_J \prod_{k \leq K} \frac{1 + J_k^2 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2}$$

$$= \prod_{k \leq K} \frac{1 + \sigma_{i_k}^1 \sigma_{i_k}^2}{4},$$

in the case (a) of the correlations between J_k^1 and J_k^2 we have

$$\mathbb{E}_J \prod_{k \leq K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2} = t \prod_{k \leq K} \frac{1 + \sigma_{i_k}^1 \sigma_{i_k}^2}{4} + \frac{1 - t}{4^K},$$

and in the case (b) of the correlations between J_k^1 and J_k^2 we have

$$\mathbb{E}_J \prod_{k \leq K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2} = \prod_{k \leq K} \frac{1 + t \sigma_{i_k}^1 \sigma_{i_k}^2}{4}.$$

Taking expectation with respect to the random indices i_k, we get

$$\mathbb{E}' \prod_{k \le K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^1 \sigma_{i_k}^2}{2} = \mathbb{E}' \prod_{k \le K} \frac{1 + J_k^2 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2}$$

$$= \frac{(1 + R_{1,2})^K}{4^K} = \frac{1 + \xi(R_{1,2})}{4^K},$$

in the case (a) of the correlations between J_k^1 and J_k^2 we have

$$\mathbb{E}' \prod_{k \le K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2} = \frac{1 + t\xi(R_{1,2})}{4^K},$$

and in the case (b) of the correlations between J_k^1 and J_k^2 we have

$$\mathbb{E}' \prod_{k \le K} \frac{1 + J_k^1 \sigma_{i_k}^1}{2} \cdot \frac{1 + J_k^2 \sigma_{i_k}^2}{2} = \frac{1 + \xi(t R_{1,2})}{4^K}.$$

From now on, we will consider the case (a) since the second case (b) is similar. Using the above formulas, let us compute $\mathbb{E}\langle \Delta(\sigma^1, \sigma^2) \rangle_s^n$ above for $n = 1, 2$ first. For $n = 1$,

$$\mathbb{E}\langle \Delta(\sigma^1, \sigma^2) \rangle_s = \frac{2}{2^K} - \frac{1 - e^{-\delta}}{4^K} \left(1 + t \mathbb{E}\langle \xi(R_{1,1}^{1;2}) \rangle_s \right).$$

For $n = 2$, we will separate the terms that contain the factor $(1 - e^{-\delta})$ to obtain

$$\mathbb{E}\langle \Delta(\sigma^1, \sigma^2) \rangle_s^2 = \mathbb{E} \Big\langle \frac{1 + \xi(R_{1,2}^{1;1})}{4^K} + \frac{1 + \xi(R_{1,2}^{2;2})}{4^K} + \frac{1 + t\xi(R_{1,2}^{1;2})}{4^K} + \frac{1 + t\xi(R_{1,2}^{2;1})}{4^K} \Big\rangle_s$$

$$+ \text{III}_1,$$

where $|\text{III}_1| \le L(1 - e^{-\delta})$. Finally, using that $\Delta(\sigma^1, \sigma^2) \le 1 + e^{-\delta}$, we can bound the sum over $n \ge 3$ in II above by

$$\text{III}_2 := \lambda \sum_{n \ge 3} \frac{(1 - e^{-\delta})^n}{n} \mathbb{E}\langle \Delta(\sigma^1, \sigma^2) \rangle_s^n \le \lambda \sum_{n \ge 3} \frac{(1 - e^{-2\delta})^n}{n} \le L\lambda (1 - e^{-2\delta})^3.$$

Plugging all these expressions back into II,

$$\text{II} = \frac{2\lambda(1 - e^{-\delta})}{2^K} - \frac{\lambda(1 - e^{-\delta})^2}{4^K} \left(1 + t\mathbb{E}\langle \xi(R_{1,1}^{1;2}) \rangle_s \right)$$

$$+ \frac{4\lambda(1 - e^{-\delta})^2}{2 \cdot 4^K} + \frac{\lambda(1 - e^{-\delta})^2}{2 \cdot 4^K}$$

$$\times \mathbb{E}\langle \xi(R_{1,2}^{1;1}) + \xi(R_{1,2}^{2;2}) + t\xi(R_{1,2}^{1;2}) + t\xi(R_{1,2}^{2;1}) \rangle_s + \text{III},$$

where $|\text{III}| \le L\lambda (1 - e^{-2\delta})^3$.

Next, given λ and δ, we are going to make the following choice of β:

$$(21) \qquad \frac{\beta^2}{2} = \frac{\lambda(1 - e^{-\delta})^2}{2 \cdot 4^K} \quad \text{or} \quad \beta = \frac{\sqrt{\lambda}(1 - e^{-\delta})}{2^K}.$$

With this choice, all the terms containing overlaps $R_{\ell,\ell'}^{j,j'}$ cancel out and, since $2\xi(1) = 2 \cdot 2^K - 2$,

$$\varphi'(s) = \mathrm{I} + \mathrm{II} = \frac{2\lambda(1 - e^{-\delta})}{2^K} + \frac{2\lambda(1 - e^{-\delta})^2}{2 \cdot 2^K} + \mathrm{III},$$

where $|\mathrm{III}| \le L\lambda(1 - e^{-2\delta})^3 \le L\lambda\delta^3$. Recall that we are interested in the regime when λ is large enough, and below we will take $\delta = \lambda^{-1/3}$. By Taylor's expansion,

$$\frac{2\lambda(1 - e^{-\delta})}{2^K} + \frac{2\lambda(1 - e^{-\delta})^2}{2 \cdot 2^K} = \frac{2\lambda\delta}{2^K} + O(\lambda\delta^3),$$

so $\varphi'(s) = 2\lambda\delta/2^K + O(\lambda\delta^3)$. Integrating between 0 and 1, we get

$$\left| \varphi(0) + \frac{2\lambda\delta}{2^K} - \varphi(1) \right| = O(\lambda\delta^3)$$

and, dividing both sides by δ,

$$\left| \frac{1}{\delta}\varphi(0) + \frac{2\lambda}{2^K} - \frac{1}{\delta}\varphi(1) \right| = O(\lambda\delta^2).$$

Using elementary estimates

$$\frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} H(s, \sigma^1, \sigma^2) \le \varphi(s) \le 2\log 2 + \frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} H(s, \sigma^1, \sigma^2)$$

for $s = 0$ and $s = 1$, we get

$$\left| \frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} (H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)) - \frac{1}{\delta}\varphi(0) \right| \le \frac{2\log 2}{\delta}$$

and

$$\left| \frac{\beta}{\delta} \frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} (H^1(\sigma^1) + H^2(\sigma^2)) - \frac{1}{\delta}\varphi(1) \right| \le \frac{2\log 2}{\delta}.$$

By Taylor's expansion and our choice of β above,

$$\frac{\beta}{\delta} = \frac{\sqrt{\lambda}(1 - e^{-\delta})}{2^K\delta} = \frac{\sqrt{\lambda}}{2^K} + O(\sqrt{\lambda}\delta^2)$$

and, therefore,

$$\frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} (H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)) = -\frac{2\lambda}{2^K} + \frac{\sqrt{\lambda}}{2^K} \frac{1}{N}\mathbb{E} \max_{|R_{1,2}|>\varepsilon} (H^1(\sigma^1) + H^2(\sigma^2))$$

$$+ O\left(\frac{1}{\delta} + \lambda\delta^2 + \sqrt{\lambda}\delta^2 \right).$$

With the choice of $\delta = \lambda^{-1/3}$, the error term here is $O(\lambda^{1/3})$. By Theorem 2, we conclude that there exists $\eta > 0$ such that

$$\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)\right)$$

$$\leq -\frac{2\lambda}{2^K} + \frac{2\sqrt{\lambda}}{2^K}\frac{1}{N}\mathbb{E}\max_\sigma H(\sigma) - \frac{2\sqrt{\lambda}\eta}{2^K} + L\lambda^{1/3}$$

and, using (12), we get

$$\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)\right) \leq \frac{2}{N}\mathbb{E}\max_\sigma H_\lambda(\sigma) - \frac{2\sqrt{\lambda}\eta}{2^K} + L\lambda^{1/3}.$$

For $\lambda \geq L\eta^{-6}$ for large enough constant $L = L(K)$, this implies that

$$\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)\right) \leq \frac{2}{N}\mathbb{E}\max_\sigma H_\lambda(\sigma) - \frac{\sqrt{\lambda}\eta}{L}.$$

By Azuma's inequality, this implies that (increasing value of the constant $L = L(K)$)

$$\frac{1}{N}\max_{|R_{1,2}|>\varepsilon}\left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2)\right) \leq \frac{1}{N}\left(\max_\sigma H_\lambda^1(\sigma) + \max_\sigma H_\lambda^2(\sigma)\right) - \frac{\sqrt{\lambda}\eta}{L}$$

with probability at least $1 - Le^{-N\eta^2/L}$. On this event, the existence of σ^1, σ^2 such that

$$\frac{1}{N}H_\lambda^\ell(\sigma^\ell) \geq \frac{1}{N}\max_\sigma H_\lambda^\ell(\sigma) - \frac{\sqrt{\lambda}\eta}{3L}$$

and such that $|R_{1,2}| > \varepsilon$ would, obviously, lead to contradiction, so the proof is finished in the case of resampling random signs.

2.2. *Resampling clauses.* Next we describe the interpolation in the case of resampling clauses. We only consider the first example of antiferromagnetic K-spin model. Under resampling of clauses, the Hamiltonians are defined in (7). Now, for $s \in [0, 1]$, we will replace these Hamiltonians by

$$H_{\lambda,s}^\ell(\sigma) = -\sum_{j \leq \pi((1-s)t\lambda N)}\sigma_{i_{1,j}}\cdots\sigma_{i_{K,j}} - \sum_{j \leq \pi_\ell((1-s)(1-t)\lambda N)}\sigma_{i_{\ell,1,j}}\cdots\sigma_{i_{\ell,K,j}}$$

and consider the interpolating Hamiltonian

$$(22) \qquad H(s, \sigma^1, \sigma^2) = \sum_{\ell=1}^2\left(\delta H_{\lambda,s}^\ell(\sigma^\ell) + \sqrt{s}\beta H^\ell(\sigma^\ell)\right).$$

Here, H^1 and H^2 are pure fully connected K-spin Hamiltonians (9) with the correlation in (15). We will use the same notation $\varphi(s)$ and G_s as in (19) and (20).

Recall that in this model, the pairs of configurations as well as the Gibbs measure are defined on $\{(\sigma^1, \sigma^2) \in V^2 : |R_{1,2}| > \varepsilon\}$, where V is the set of configurations in (5) with zero magnetization.

As in the previous section, the derivative of the interpolating free energy $\varphi(s)$ in (19) with respect to $\sqrt{s}$ in front of the second term gives

$$\mathrm{I} = \frac{\beta^2}{2}\big(2 + 2\mathbb{E}\langle t(R_{1,1}^{1,2})^K \rangle_s - \mathbb{E}\langle (R_{1,2}^{1,1})^K + (R_{1,2}^{2,2})^K + t(R_{1,2}^{1,2})^K + t(R_{1,2}^{2,1})^K \rangle_s\big).$$

On the other hand, the derivative with respect to s in the Poisson random variables will now be applied to three different terms $\pi((1-s)t\lambda N), \pi_1((1-s)(1-t)\lambda N)$ and $\pi_2((1-s)(1-t)\lambda N)$ and, similarly to the computation above,

$$\mathrm{II} = -t\lambda\mathbb{E}\log\langle\exp\big(-\delta\sigma_{i_1}^1 \cdots \sigma_{i_K}^1 - \delta\sigma_{i_1}^2 \cdots \sigma_{i_K}^2\big)\rangle_s$$
$$- (1-t)\lambda\mathbb{E}\log\langle\exp\big(-\delta\sigma_{i_1}^1 \cdots \sigma_{i_K}^1\big)\rangle_s - (1-t)\lambda\mathbb{E}\log\langle\exp\big(-\delta\sigma_{i_1}^2 \cdots \sigma_{i_K}^2\big)\rangle_s.$$

Since the product of spins takes values ± 1, we can represent

$$\exp\big(-\delta\sigma_{i_1}^\ell \cdots \sigma_{i_K}^\ell\big) = \mathrm{ch}(\delta)\big(1 - \mathrm{th}(\delta)\sigma_{i_1}^\ell \cdots \sigma_{i_K}^\ell\big)$$

and rewrite their product as

$$\exp\big(-\delta\sigma_{i_1}^1 \cdots \sigma_{i_K}^1 - \delta\sigma_{i_1}^2 \cdots \sigma_{i_K}^2\big) = \mathrm{ch}(\delta)^2\big(1 - \mathrm{th}(\delta)\Delta(\sigma^1, \sigma^2)\big)$$

with the notation

$$\Delta(\sigma^1, \sigma^2) = \sigma_{i_1}^1 \cdots \sigma_{i_K}^1 + \sigma_{i_1}^2 \cdots \sigma_{i_K}^2 - \mathrm{th}(\delta)\sigma_{i_1}^1 \cdots \sigma_{i_K}^1 \sigma_{i_1}^2 \cdots \sigma_{i_K}^2.$$

Expressing the logarithm by its Taylor series, we can rewrite

$$\mathrm{II} = -2\lambda\log\mathrm{ch}\,\delta + \lambda\sum_{n \geq 1}\frac{\mathrm{th}(\delta)^n}{n}\bigg(t\mathbb{E}\langle\Delta(\sigma^1, \sigma^2)\rangle_s^n + (1-t)\sum_{\ell=1}^2 \mathbb{E}\langle\sigma_{i_1}^\ell \cdots \sigma_{i_K}^\ell\rangle_s^n\bigg).$$

Let us recall that, in this example, we restrict configurations to the set V with zero magnetization, $m(\sigma^\ell) = N^{-1}\sum_{i \leq N}\sigma_i^\ell = 0$. For odd N, magnetization equals $1/N$ which for simplicity of notation we denote by 0. Averaging in the random signs, we get

$$\mathbb{E}\langle\sigma_{i_1}^\ell \cdots \sigma_{i_K}^\ell\rangle_s = \mathbb{E}\langle m(\sigma^\ell)^K\rangle_s = 0$$

and

$$\mathbb{E}\langle\Delta(\sigma^1, \sigma^2)\rangle_s = -\mathrm{th}(\delta)\mathbb{E}\langle (R_{1,1}^{1,2})^K\rangle_s.$$

This is the only place where we used the assumption that magnetization is zero. Using replicas as above, averaging in random signs and collecting all the terms of the order $O(\lambda\,\mathrm{th}(\delta)^3)$ into error term, it is easy to check that

$$\mathrm{II} = -2\lambda\log\mathrm{ch}\,\delta - t\lambda\,\mathrm{th}(\delta)^2\mathbb{E}\langle (R_{1,1}^{1,2})^K\rangle_s$$

$$+ \frac{\lambda\,\mathrm{th}(\delta)^2}{2}\mathbb{E}\langle (R_{1,2}^{1,1})^K + (R_{1,2}^{2,2})^K + t(R_{1,2}^{1,2})^K + t(R_{1,2}^{2,1})^K\rangle_s + O(\lambda\delta^3).$$

If we now take $\beta = \sqrt{\lambda}\,\text{th}(\delta)$, then all the terms containing overlaps cancel out and we get

$$\varphi'(s) = \text{I} + \text{II} = -2\lambda\log\text{ch}\,\delta + \lambda\,\text{th}(\delta)^2 + O(\lambda\delta^3).$$

One can check that $-2\log\text{ch}\,\delta + \text{th}(\delta)^2 = O(\delta^4)$ as $\delta \to 0$ and, therefore, $\varphi'(s) = O(\lambda\delta^3)$. By Taylor's expansion and our choice of β,

$$\frac{\beta}{\delta} = \frac{\sqrt{\lambda}\,\text{th}(\delta)}{\delta} = \sqrt{\lambda} + O(\sqrt{\lambda}\delta^2)$$

and, using elementary estimates connecting the free energy and maximum as above, one can check that the obtained control of the derivative implies that

$$\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\big(H^1_\lambda(\sigma^1) + H^2_\lambda(\sigma^2)\big)$$
$$= \sqrt{\lambda}\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\big(H^1(\sigma^1) + H^2(\sigma^2)\big) + O(\lambda^{1/3}),$$

if we again take $\delta = \lambda^{-1/3}$. The maximum on the right-hand side is taken over $\sigma^1, \sigma^2 \in V$ with zero magnetization and we can bound it from above by removing this magnetization constraint and then applying Theorem 2 to the right-hand side,

$$\frac{1}{N}\mathbb{E}\max_{|R_{1,2}|>\varepsilon}\big(H^1(\sigma^1) + H^2(\sigma^2)\big) \le \frac{2}{N}\mathbb{E}\max_\sigma H(\sigma) - \eta,$$

where the maximum on the right-hand side is now taken over all $\sigma \in \{-1,+1\}^N$. However, since the ground state energy over the whole space is essentially the same as over subset V of configurations with zero magnetization as $N \to \infty$ (see Lemma 1 below), by reducing η we can replace the maximum above by the one over $\sigma \in V$. Then, using (11), the proof is finished in exactly the same way as above.

LEMMA 1.   *We have that*

$$\lim_{N\to\infty}\mathbb{E}\max_\sigma\frac{H(\sigma)}{N} = \lim_{N\to\infty}\mathbb{E}\max_{\sigma\in V}\frac{H(\sigma)}{N}$$

PROOF.   Obviously the left-hand side is no less than the right-hand side. It remains to show the reverse inequality. First, we note that it is already known from the proof of [8], Proposition 9, that for any $\varepsilon > 0$,

$$(23)\qquad \lim_{N\to\infty}\mathbb{E}\max_{|m(\sigma)|<\varepsilon}\frac{H(\sigma)}{N} = \lim_{N\to\infty}\mathbb{E}\max_\sigma\frac{H(\sigma)}{N},$$

where $m(\sigma) := N^{-1}\sum_{i=1}^N \sigma_i$ is the magnetization of σ. Let $\varepsilon > 0$ be fixed. For any σ satisfying $|m(\sigma)| < \varepsilon$, we can find $\pi(\sigma) \in \{-1,1\}^N$ with $m(\pi(\sigma)) \in V$

such that the Hamming distance between σ and $\pi(\sigma)$ satisfies

$$(24) \qquad d(\sigma, \pi(\sigma)) = \frac{1}{N} \sum_{i=1}^{N} I(\sigma_i \neq \pi(\sigma)_i) < \varepsilon$$

provided that $N\varepsilon > 1$. Indeed, let $A = \{i : \sigma_i = 1\}$ and $B = \{i : \sigma_i = -1\}$ and assume $|A| < |B|$ ($|A| \geq |B|$ is similar). If N is even, we consider a partition $\{B_1, B_2, B_3\}$ of B with $|B_1| = |B_2| = (|B| - |A|)/2$. Set $\pi(\sigma)$ as

$$\pi(\sigma)_i = \begin{cases} 1 & \text{if } i \in A \cup B_1, \\ -1 & \text{if } i \in B_2 \cup B_3. \end{cases}$$

Then clearly $m(\pi(\sigma)) = 0$ and

$$d(\sigma, \pi(\sigma)) = \frac{1}{N}|B_1| = \frac{|B| - |A|}{2N} = \frac{|m(\sigma)|}{2} < \frac{\varepsilon}{2}.$$

Similarly, if N is odd, then we consider a partition $\{B_1, B_2, B_3, B_4\}$ of B with $|B_1| = |B_2| = (|B| - |A| - 1)/2$ and $|B_4| = 1$. Set $\pi(\sigma)$ by

$$\pi(\sigma)_i = \begin{cases} 1 & \text{if } i \in A \cup B_1 \cup B_4, \\ -1 & \text{if } i \in B_2 \cup B_3. \end{cases}$$

Then $m(\pi(\sigma)) = 1/N$ and if $N\varepsilon > 1$,

$$d(\sigma, \pi(\sigma)) = \frac{1}{N}(|B_1| + |B_4|) = \frac{|B| - |A| + 1}{2N} < \frac{\varepsilon}{2} + \frac{1}{2N}.$$

All these imply that $\pi(\sigma) \in V$ and (24) holds. Now, for any $N\varepsilon > 1$, if $|m(\sigma)| < \varepsilon$, then

$$\begin{aligned} \mathbb{E}(H(\sigma) - H(\pi(\sigma)))^2 &= 2N(\xi(1) - \xi(R(\sigma, \pi(\sigma)))) \\ &\leq 2N\xi'(1)d(\sigma, \pi(\sigma)) \\ &< 2N\xi'(1)\varepsilon. \end{aligned}$$

We apply the Slepian inequality (see, e.g., [18]) to get

$$\mathbb{E} \max_{|m(\sigma)| < \varepsilon} |H(\sigma) - H(\pi(\sigma))| \leq N\sqrt{2\varepsilon\xi'(1)\log 2}.$$

Consequently,

$$\mathbb{E} \max_{|m(\sigma)| < \varepsilon} \frac{H(\sigma)}{N} \leq \mathbb{E} \max_{\sigma \in V} \frac{H(\sigma)}{N} + \sqrt{2\varepsilon\xi'(1)\log 2}.$$

From (23), sending $N \to \infty$ and then $\varepsilon \downarrow 0$ gives

$$\lim_{N \to \infty} \mathbb{E} \max_{\sigma} \frac{H(\sigma)}{N} \leq \lim_{N \to \infty} \mathbb{E} \max_{\sigma \in V} \frac{H(\sigma)}{N}$$

and this completes our proof. $\quad \square$

3. Proof of Theorem 2. In section, we establish Theorem 2 assuming the case (ii). In this case H is associated to

$$\xi(s) = (1+s)^K - 1 \tag{25}$$

for $K \geq 2$. For notational convenience, we denote the covariance between H^1 and H^2 by ξ_0,

$$\mathbb{E}H^1(\sigma)H^2(\sigma^2) = \xi_0(R_{1,2}).$$

Throughout the remainder of this section, we assume that $t \in (0,1)$ and H^1, H^2 have the covariance structure (15) or (16), that is, either

$$\xi_0(s) = t\xi(s) \tag{26}$$

or

$$\xi_0(s) = \xi(ts). \tag{27}$$

Recently, chaos in disorder for the ground energy was proved in [8] for the case when ξ is an infinite series of even mixtures and (26) holds. The major difference between [8] and our current situation is that ξ now includes odd p-spin interactions and H^1 and H^2 possess a new type of covariance structure (27). Our proof of Theorem 2 is essentially based on [6], where the author established chaos in disorder at positive temperature using the Guerra replica symmetry breaking (RSB) bound [13] and its two-dimensional extension (in the spirit of Talagrand [29, 30]) under general assumptions on ξ and ξ_0. While several properties of the Parisi measure $\gamma_P(ds)$ and the Parisi PDE Φ_γ established in [8] will be used, our proof will follow closely the one in [6] only now at zero temperature.

3.1. *Parisi formula and the Guerra–Talagrand RSB bound.* The first key ingredient is played by the Parisi formula for the maximum Hamiltonian in [7]. Let $\mathcal{U}$ be the set of all nonnegative and nondecreasing right-continuous functions γ on $[0,1]$ such that

$$\int_0^1 \gamma(s)\,ds < \infty.$$

We equip the space $\mathcal{U}$ with $L^1(dx)$ norm. Define the Parisi functional on $\mathcal{U}$ by

$$\mathscr{P}(\gamma) = \Phi_\gamma(0,0) - \frac{1}{2}\int_0^1 s\xi''(s)\gamma(s)\,ds, \tag{28}$$

where $\Phi_\gamma(0,0)$ is defined through the weak solution of the Parisi PDE with boundary condition $\Phi_\gamma(1,x) = |x|$,

$$\partial_s \Phi_\gamma(s,x) = -\frac{\xi''(s)}{2}\big(\partial_{xx}\Phi_\gamma(s,x) + \gamma(s)(\partial_x\Phi_\gamma(s,x))^2\big) \tag{29}$$

for $(s, x) \in [0, 1) \times \mathbb{R}$. The existence and regularity properties of Φ_γ can be found in [8]. The Parisi formula for the maximum energy states that, for $ME_N = N^{-1} \max_\sigma H(\sigma)$,

$$(30) \qquad ME := \lim_{N \to \infty} ME_N = \inf_{\gamma \in \mathscr{U}} \mathscr{P}(\gamma) \qquad \text{a.s.}$$

Here the minimizer on the right-hand side exists and is unique, see [8]. Denote this minimizer by γ_P and call $\gamma_P(ds)$ the Parisi measure. We mention that (30) was indeed established for general mixtures, see [7].

Next, we state Guerra–Talagrand's RSB bound. For any measurable $A \in [-1, 1]$, consider the normalized maximum

$$MCE_N(A) = \frac{1}{N} \max_{(\sigma^1, \sigma^2): R_{1,2} \in A} \left(H^1(\sigma^1) + H^2(\sigma^2) \right).$$

If A contains only one point q, we simply write $MCE_N(A) = MCE_N(q)$. Let $S_N = \{R_{1,2} : \sigma^1, \sigma^2 \in \{-1, +1\}^N\}$ be the set of possible overlap value for a given N. Let $q \in [-1, 1]$ be fixed. Denote by $\iota = 1$ if $q \geq 0$ and $\iota = -1$ if $q < 0$. Define a matrix-valued function T on $[0, 1]$ by

$$(31) \qquad T(s) = \begin{cases} \begin{bmatrix} \xi''(s) & \iota\xi_0''(\iota s) \\ \iota\xi_0''(\iota s) & \xi''(s) \end{bmatrix} & \text{if } s \in [0, |q|), \\ \begin{bmatrix} \xi''(s) & 0 \\ 0 & \xi''(s) \end{bmatrix} & \text{if } s \in [|q|, 1]. \end{cases}$$

For any $\gamma \in \mathscr{U}$, consider the weak solution $\Psi_\gamma(s, \mathbf{x})$ on $[0, 1] \times \mathbb{R}^2$ of

$$\partial_s \Psi_\gamma(s, \mathbf{x}) = -\frac{1}{2} \left(\langle T(s), \nabla^2 \Psi_\gamma(s, \mathbf{x}) \rangle + \gamma(s) \langle T(s) \nabla \Psi_\gamma(s, \mathbf{x}), \nabla \Psi_\gamma(s, \mathbf{x}) \rangle \right)$$

with boundary condition $\Psi_\gamma(1, \mathbf{x}) = |x_1| + |x_2|$. Here the existence of this PDE and its regularity can be argued in a similar way as those for Φ_γ appearing in [8], Appendix. Define

$$(32) \qquad \mathscr{T}_q(\gamma) = \Psi_\gamma(0, 0, 0) - \frac{1}{2} \left(\int_0^1 s\xi''(s)\gamma(s) \, ds + \int_0^{|q|} s\xi_0''(\iota s)\gamma(s) \, ds \right)$$

for all $\gamma \in \mathscr{U}$. For any $q \in S_N$, the Guerra–Talagrand RSB bound for the expected value of the normalized maximum defined above is given by

$$(33) \qquad \mathbb{E} MCE_N(q) \leq \mathscr{T}_q(\gamma).$$

This inequality is obtained from the usual Guerra–Talagrand upper bound at positive temperature, which holds because of the convexity of ξ and ξ_0 on $[-1, 1]$, by taking zero temperature limit with the same rescaling of the functional order parameter as in one dimensional case explained in Lemma 2 in [7]. In other words, this is a standard two-dimensional analogue of Lemma 2 in [7].

From (30) and (33), the proof of Theorem 2 relies on finding $\gamma \in \mathscr{U}$ such that

$$\mathscr{T}_q(\gamma) < 2\mathscr{P}(\gamma_P),$$

whenever $|q| > \varepsilon$. From (28) and (32), one way of attaining this is to find $\gamma \in \mathscr{U}$ such that the following two conditions are satisfied:

(i) $\int_0^1 s\xi''(s)\gamma(s)\,ds + \int_0^{|q|} s\xi_0''(\iota s)\gamma(s)\,ds = \int_0^1 s\xi''(s)\gamma_P(s)\,ds$;
(ii) $\Psi_\gamma(0,0,0) < 2\Phi_{\gamma_P}(0,0)$.

To obtain (i), we choose

$$(34) \qquad \gamma_q(s) := \begin{cases} \dfrac{\xi''(s)\gamma_P(s)}{\xi''(s) + \xi_0''(\iota s)} & \text{if } s \in [0, |q|), \\ \gamma_P(s) & \text{if } s \in [|q|, 1], \end{cases}$$

and an algebraic manipulation gives

$$(35) \qquad \int_0^1 s\xi''(s)\gamma_q(s)\,ds + \int_0^{|q|} s\xi_0''(\iota s)\gamma_q(s)\,ds = \int_0^1 s\xi''(s)\gamma_P(s)\,ds.$$

One technical condition we need here is the requirement that γ_q must lie in $\mathscr{U}$. The lemma below justifies this condition.

LEMMA 2. *Consider ξ defined through (25) and ξ_0 defined by either (26) or (27). We have that $\xi_0''(s) < \xi''(|s|)$ for all $s \in [-1, 1] \setminus \{0\}$ and ζ_+ and ζ_- are nondecreasing on $(0, 1]$, where for $s \in (0, 1]$,*

$$(36) \qquad \begin{aligned} \zeta_+(s) &:= \frac{\xi''(s)}{\xi''(s) + \xi_0''(s)}, \\ \zeta_-(s) &:= \frac{\xi''(s)}{\xi''(s) + \xi_0''(-s)}. \end{aligned}$$

Since ζ_+ and ζ_- are nondecreasing and ξ'' and ξ_0'' are nonnegative, we see that γ_q is nonnegative and nondecreasing with $\lim_{s \to |q|-} \gamma_q(s) \le \gamma_P(|q|)$. Thus, $\gamma_q \in \mathscr{U}$. The fact that $\xi_0''(s) < \xi''(|s|)$ is not needed for this statement, but will be used in the subsequent sections.

PROOF OF LEMMA 2. Compute directly

$$\zeta_+'(s) = \frac{\xi'''(s)\xi_0''(s) - \xi''(s)\xi_0'''(s)}{(\xi''(s) + \xi_0''(s))^2}$$

and

$$\zeta_-'(s) = \frac{\xi'''(s)\xi_0''(-s) + \xi''(s)\xi_0'''(-s)}{(\xi''(s) + \xi_0''(s))^2}.$$

If (26) holds, then

$$\xi'''(s)\xi_0''(s) - \xi''(s)\xi_0'''(s) = tK^2(K-1)^2(K-2)(1+s)^{2K-5} \geq 0$$

and

$$\xi'''(s)\xi_0''(s) + \xi''(s)\xi_0'''(-s)$$
$$= tK^2(K-1)^2(K-2)\big((1+s)^{K-5} + (1+s)^{K-2}(1-s)^{K-3}\big) \geq 0.$$

Thus, ζ_+ and ζ_- are nondecreasing. The fact that $\xi_0''(s) < \xi(|s|)$ holds for all $s \in [-1, 1] \setminus \{0\}$ is clear since $t \in (0, 1)$. Next, assume that (27) is valid. Since

$$\xi'''(s)\xi_0''(s) - \xi''(s)\xi_0'''(s)$$
$$= K^2(K-1)^2(K-2)$$
$$\times \big(t^2(1+s)^{K-3}(1+ts)^{K-2} - t^3(1+s)^{K-2}(1+ts)^{K-3}\big)$$
$$= (1-t)t^2K^2(K-1)^2(K-2)(1+s)^{K-3}(1+ts)^{K-3}$$
$$\geq 0$$

and

$$\xi'''(s)\xi_0''(-s) + \xi''(s)\xi_0'''(-s)$$
$$= K^2(K-1)^2(K-2)$$
$$\times \big(t^2(1+s)^{K-3}(1-ts)^{K-2} + t^3(1+s)^{K-2}(1-ts)^{K-3}\big)$$
$$= (1+t)t^2K^2(K-1)^2(K-2)(1+s)^{K-3}(1+ts)^{K-3}$$
$$\geq 0.$$

From these, ζ_+ and ζ_- are nondecreasing. On the other hand, clearly

$$\xi_0''(s) = t^2K(K-1)(1+ts)^{K-2} < K(K-1)(1+|s|)^{K-2} = \xi''(|s|)$$

for all $s \in [-1, 1] \setminus \{0\}$. This completes our proof. $\quad\square$

3.2. *Variational representations for Φ_γ and Ψ_γ.* In order to establish the condition (ii) in the previous section, a key ingredient we need is the variational representation for Φ_γ and Ψ_γ in terms of optimal stochastic control problems. Denote by $\mathbf{W} = \{\mathbf{W}(w) = (W_1(w), W_2(w)), \mathscr{G}_w, 0 \leq w \leq 1\}$ a two-dimensional Brownian motion, where the filtration $(\mathscr{G}_w)_{0 \leq w \leq 1}$ satisfies the usual conditions (see Definition 2.25 in Chapter 1 of [17]). Let now $\gamma \in \mathscr{U}$ be fixed. For $0 \leq s \leq 1$, denote by $D[s]$ the space of all two-dimensional progressively measurable processes $v = (v_1, v_2)$ with respect to $(\mathscr{G}_w)_{0 \leq w \leq s}$ satisfying $\sup_{0 \leq w \leq s} |v_1(w)| \leq 1$ and $\sup_{0 \leq w \leq s} |v_2(w)| \leq 1$. Endow the space $D[s]$ with the norm

$$\|v\|_s = \left(\mathbb{E} \int_0^s \big(v_1(w)^2 + v_2(w)^2\big) \, dw\right)^{1/2}.$$

Recall $T(s)$ in (31) and define a functional

$$F_\gamma^s(v) = \mathbb{E}\big[C_\gamma^s(v) - L_\gamma^s(v)\big]$$

for $v \in D[s]$, where

$$C_\gamma^s(v) := \Psi_\gamma\bigg(s, \int_0^s \gamma(w)T(w)v(w)\,dw + \int_0^s T(w)^{1/2}\,d\mathbf{W}(w)\bigg),$$

$$L_\gamma^s(v) := \frac{1}{2}\int_0^s \gamma(w)\langle T(w)v(w), v(w)\rangle\,dw.$$

The functional Ψ_γ defined above as the solution of a PDE can also be written via an optimal stochastic control problem.

PROPOSITION 1. *Let $\gamma \in \mathscr{U}$. For any $s \in [0, 1]$,*

(37) $$\Psi_\gamma(0, 0, 0) = \max\{F_\gamma^s(v) : v \in D[s]\}.$$

The maximum of (37) is attained by $v_\gamma(r) = \nabla\Psi_\gamma(r, \mathbf{X}(r))$, where the two-dimensional stochastic process $(\mathbf{X}_\gamma(w))_{0 \le w \le s}$ satisfies

$$d\mathbf{X}_\gamma(w) = \gamma(w)T(w)\nabla\Psi_\gamma\big(w, \mathbf{X}(w)\big)\,dw + T(w)^{1/2}\,d\mathbf{W}(w),$$

$$\mathbf{X}_\gamma(0) = (0, 0).$$

The derivation of Proposition 1 is a purely an application of Itô's formula. For a detailed proof, we refer the readers to [6], Theorem 5. Although the argument therein is for different boundary condition and $\gamma(1-)$ is bounded, the same argument carries through in the current setting with only minor modification. Note that when $q = 0$, T is a diagonal matrix and

$$\Psi_\gamma(0, x_1, x_2) = \Phi_\gamma(0, x_1) + \Phi_\gamma(0, x_2).$$

Proposition 1 is a zero-temperature two-dimensional analogue of Lemma 2 in [7] (see also [16]). By taking $x_1 = x_2 = 0$, it implies the zero-temperature one-dimensional analogue of Lemma 2 in [7], giving the following representation of Φ_γ. Let W be a one-dimensional standard Brownian motion with respect to the filtration $(\mathscr{G}_w)_{0 \le w \le 1}$ and $D_0[s]$ be the space of all progressively measurable processes u with respect to $(\mathscr{G}_w)_{0 \le w \le s}$ and satisfy $\sup_{0 \le w \le s}|u(w)|\,dw \le 1$.

COROLLARY 1. *Let $\gamma \in \mathscr{U}$. For any $s \in [0, 1]$,*

(38)
$$\begin{aligned}
\Phi_\gamma(0, 0) = \max_{u \in D_0[s]} \mathbb{E}\bigg[&\Phi_\gamma\bigg(s, \int_0^s \xi''(w)\gamma(w)u(w)\,dw \\
&+ \int_0^s \sqrt{\xi''(w)}\,dW(w)\bigg) \\
&- \frac{1}{2}\int_0^s \xi''(w)\gamma(w)u(w)^2\,dw\bigg].
\end{aligned}$$

Here the maximizer is given by $u_\gamma(w) = \partial_x \Phi_\gamma(w, X_\gamma(w))$, where $X_\gamma = (X_\gamma(w))_{0 \le w \le s}$ is the solution to the following SDE with the initial condition $X_\gamma(0) = 0$,

$$(39) \qquad dX_\gamma = \xi''(w)\gamma(w)\partial_x \Phi_\gamma(w, X_\gamma(w))\, dw + \sqrt{\xi''(w)}\, dW(w).$$

REMARK 2. From [8], Lemma 2, this minimizer u_γ is unique if $\gamma(s) > 0$ on $(0, 1]$. In particular, from [8], Section 3.2, 0 lies in the support of $\gamma_P(ds)$, so $\gamma_P(s) > 0$ on $(0, 1]$.

While it is generally not possible to find the solutions Ψ_{γ_q} and Φ_{γ_P} explicitly and compare their values, the variational representations in Proposition 1 and Corollary 1 provide an elementary way to quantify the difference between $\Psi_{\gamma_q}(0, 0, 0)$ and $2\Phi_{\gamma_P}(0, 0)$.

PROPOSITION 2. *The following two statements hold*:

(i) *If $v_{\gamma_q} = (v_1, v_2)$ is the maximizer to the variational problem (37) for $\Psi_{\gamma_q}(0, 0)$ using $s = |q|$, then*

$$
\begin{aligned}
(40) \qquad \Psi_{\gamma_q}(0, 0, 0) &\le 2\Phi_{\gamma_P}(0, 0) - \frac{1}{2}\int_0^{|q|} \frac{\xi''(w)\xi_0''(\iota w)(\xi''(w) - \xi_0''(\iota w))}{2(\xi''(w) + \xi_0''(\iota w))^2} \\
&\quad \times \gamma_P(w)\mathbb{E}(v_1(w) - \iota v_2(w))^2\, dw.
\end{aligned}
$$

(ii) *Let us define*

$$(41) \qquad (u_1(w), u_2(w)) = \frac{1}{\xi''(w) + \xi_0''(\iota w)} T(w)v_{\gamma_q}(w),$$

$$(42) \qquad (B_1(w), B_2(w)) = \frac{1}{\xi''(w)^{1/2}} T(w)^{1/2}\mathbf{W}(w)$$

for $0 \le w \le |q|$. If the following equality holds,

$$\Psi_{\gamma_u}(0, 0, 0) = 2\Phi_{\gamma_P}(0, 0),$$

then u_1 and u_2 are the maximizers of (38) with respect to the Brownian motions B_1 and B_2 in (42) respectively, that is, on the interval $[0, |q|]$,

$$
(43) \qquad
\begin{aligned}
u_1(w) &= \partial_x \Phi_{\gamma_P}(w, X_{1,\gamma_P}(w)), \\
u_2(w) &= \partial_x \Phi_{\gamma_P}(w, X_{2,\gamma_P}(w)),
\end{aligned}
$$

where $(X_{1,\gamma_P}(w))_{0 \le w \le |q|}$ and $(X_{2,\gamma_P}(w))_{0 \le w \le |q|}$ satisfy

$$
(44) \qquad
\begin{aligned}
dX_{1,\gamma_P}(w) &= \gamma_P(w)\xi''(w)\partial_x \Phi_{\gamma_P}(w, X_{1,\gamma_P}(w))\, dw + \xi''(w)^{1/2}\, dB_1(w), \\
dX_{2,\gamma_P}(w) &= \gamma_P(w)\xi''(w)\partial_x \Phi_{\gamma_P}(w, X_{2,\gamma_P}(w))\, dw + \xi''(w)^{1/2}\, dB_2(w)
\end{aligned}
$$

with the initial condition $X_{1,\gamma_P}(0) = X_{2,\gamma_P}(0) = 0$.

Proposition 2 is essentially taken from [6], Proposition 5. Its proof is based on the comparison between (37) and (38) with $s = |q|$. More precisely, note that

$$\Psi_{\gamma_q}(|q|, x_1, x_2) = \Phi_{\gamma_P}(|q|, x_1) + \Phi_{\gamma_P}(|q|, x_2).$$

If one takes the optimizer v_{γ_q} in (37), then (40) can be obtained after some algebraic manipulation, while the statement (43) follows by Remark 2. For details, we refer the readers to [6], Proposition 5. Although the PDEs considered there have different boundary conditions, this does not affect the proof in any essential way.

3.3. *Proof of Theorem* 2 *assuming the case* (ii). First note that $\mathscr{T}_q(\gamma_q) \leq 2\mathscr{P}(\gamma_P)$ from (35) and (40). We claim that

$$\text{(45)} \qquad\qquad\qquad \mathscr{T}_q(\gamma_q) < 2\mathscr{P}(\gamma_P)$$

for all $q \in [-1, 1] \setminus \{0\}$. Assume on the contrary $\mathscr{T}_q(\gamma_q) = 2\mathscr{P}(\gamma_P)$ for some $q \in [-1, 1] \setminus \{0\}$. This and (35) imply $\Psi_{\gamma_q}(0, 0, 0) = 2\Phi_{\gamma_P}(0, 0)$. Note that $\gamma_P(s) > 0$ on $(0, 1]$ by Remark 2. Since $\xi_0''(w) < \xi''(|w|)$ for all $w \in [-1, 1] \setminus \{0\}$ by Lemma 2, applying (40) gives that $v_1 = \iota v_2$ on $[0, |q|]$ and thus, the definition of T in (31) and (41) imply that $(u_1, u_2) = (v_1, \iota v_1)$. By (43),

$$\partial_x \Phi_{\gamma_P}(w, X_{1,\gamma_P}(w)) = u_1(w) = \iota u_2(w) = \iota \partial_x \Phi_{\gamma_P}(w, X_{2,\gamma_P}(w))$$

on $[0, |q|]$. Since $\partial_x \Phi_\gamma(w, \cdot)$ is a strictly increasing odd function on $\mathbb{R}$ (see [8], Lemma 4), $X_{1,\gamma_P} = \iota X_{2,\gamma_P}$ on $[0, |q|]$. Consequently, from (44),

$$0 = X_{1,\gamma_P}(s) - \iota X_{2,\gamma_P}(s) = \int_0^s \xi''(w)^{1/2} \, d\big(B_1(w) - \iota B_2(w)\big)$$

and thus, $B_1(w) = \iota B_2(w)$ on $[0, |q|]$. The definition (42) then implies that

$$0 = \mathbb{E}\big(B_1(w) - \iota B_2(w)\big)^2 = 2\big(\xi''(w) - \xi_0''(\iota w)\big).$$

On the other hand, by Lemma 2, the right-hand side is strictly positive—a contradiction.

Next, from (37) with $s = |q|$, it is easy to see that $\Psi_{\gamma_q}(0, 0, 0)$ is continuous in q. Therefore, by (45), for any $\varepsilon \in (0, 1)$ there exists $\eta > 0$ such that $\mathscr{T}_q(\gamma_q) \leq 2\mathscr{P}(\gamma_P) - 3\eta$ for all $q \in S_N \setminus [-\varepsilon, \varepsilon]$. Applying (30) and (33) yields that, for large enough N,

$$\mathbb{E}MCE_N(u) \leq 2\mathbb{E}ME_N - 2\eta$$

for any $q \in S_N \setminus [-\varepsilon, \varepsilon]$. Furthermore, since $S_N \setminus [-\varepsilon, \varepsilon]$ contains no more than $2N$ distinct elements, using the usual concentration of measure for Gaussian extrema processes $MCE_N(u)$ and ME_N implies that, for large enough N,

$$\mathbb{E}MCE_N\big([-1, 1] \setminus [-\varepsilon, \varepsilon]\big) \leq 2\mathbb{E}ME_N - \eta.$$

This finishes our proof.

REFERENCES

[1] CHATTERJEE, S. (2008). Chaos, concentration, and multiple valleys. Preprint. Available at arXiv:0810.4221.

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

[3] CHATTERJEE, S. (2014). *Superconcentration and Related Topics*. Springer, Berlin.

[4] CHEN, W.-K. (2013). Disorder chaos in the Sherrington-Kirkpatrick model with external field. *Ann. Probab.* **41** 3345–3391.

[5] CHEN, W.-K. (2014). Chaos in the mixed even-spin models. *Comm. Math. Phys.* **328** 867–901. MR3201215

[6] CHEN, W.-K. (2015). Variational representations for the Parisi functional and the two-dimensional Guerra-Talagrand bound. Preprint. Available at arXiv:1501.06635.

[7] CHEN, W.-K. and AUFFINGER, A. (2016). Parisi formula for the ground state energy in the mixed *p*-spin model. Preprint. Available at arXiv:1606.05335.

[8] CHEN, W.-K., HANDSCHY, M. and LERMAN, G. (2016). On the energy landscape of the mixed even *p*-spin model. Preprint. Available at arXiv:1609.04368.

[9] CHEN, W.-K., HSIEH, H.-W., HWANG, C.-R. and SHEU, Y.-C. (2015). Disorder chaos in the spherical mean-field model. *J. Stat. Phys.* **160** 417–429. MR3360467

[10] COPPERSMITH, D., GAMARNIK, D., HAJIAGHAYI, M. and SORKIN, G. B. (2004). Random MAX SAT, random MAX CUT, and their phase transitions. *Random Structures Algorithms* **24** 502–545.

[11] DEMBO, A., MONTANARI, A. and SEN, S. (2017). Extremal cuts of sparse random graphs. *Ann. Probab.* **45** 1190–1217. MR3630296

[12] FRANZ, S. and LEONE, M. (2003). Replica bounds for optimization problems and diluted spin systems. *J. Stat. Phys.* **111** 535–564. MR1972121

[13] GUERRA, F. (2003). Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** 1–12. MR1957729

[14] GUERRA, F. and TONINELLI, F. L. (2004). The high temperature region of the Viana-Bray diluted spin glass model. *J. Stat. Phys.* **115** 531–555. MR2070106

[15] JAGANNATH, A., KO, J. and SEN, S. (2017). A connection between MAX κ-CUT and the inhomogeneous Potts spin glass in the large degree limit. Preprint. Available at arXiv:1703.03455.

[16] JAGANNATH, A. and TOBASCO, I. (2016). A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.* **144** 3135–3150. MR3487243

[17] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[18] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*: Isoperimetry and Processes. Ergebnisse der Mathematik und Ihrer Grenzgebiete (3) [*Results in Mathematics and Related Areas (3)*] **23**. Springer, Berlin. MR1102015

[19] LEUZZI, L. and PARISI, G. (2001). The *K*-SAT problem in a simple limit. *J. Stat. Phys.* **103** 679–695. MR1851354

[20] PANCHENKO, D. (2013). *The Sherrington–Kirkpatrick Model*. Springer, New York. MR3052333

[21] PANCHENKO, D. (2013). The Parisi ultrametricity conjecture. *Ann. of Math. (2)* **177** 383–393. MR2999044

[22] PANCHENKO, D. (2014). The Parisi formula for mixed *p*-spin models. *Ann. Probab.* **42** 946–958.

[23] PANCHENKO, D. (2016). On the *K*-sat model with large number of clauses. Preprint. Available at arXiv:1608.06256.

[24] PANCHENKO, D. and TALAGRAND, M. (2004). Bounds for diluted mean-fields spin glass models. *Probab. Theory Related Fields* **130** 319–336. MR2095932

[25] PARISI, G. (1980). A sequence of approximate solutions to the S-K model for spin glasses. *J. Phys. A* **13** L-115.

[26] PARISI, G. (1983). Order parameter for spin-glasses. *Phys. Rev. Lett.* **50** 1946–1948. MR0702601

[27] SEN, S. (2016). Optimization on sparse random hypergraphs and spin glasses. Preprint. Available at arXiv:1606.02365.

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

[29] TALAGRAND, M. (2006). The Parisi formula. *Ann. of Math.* (2) **163** 221–263. MR2195134

[30] TALAGRAND, M. (2007). Mean field models for spin glasses: Some obnoxious problems. In *Spin Glasses. Lecture Notes in Math.* **1900** 63–80. Springer, Berlin. MR2309598

[31] TALAGRAND, M. (2011). *Mean Field Models for Spin Glasses. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge A Series of Modern Surveys in Mathematics* **54, 55**. Springer, Berlin.

SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
537 VINCENT HALL
206 CHURCH ST. SE
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: wkchen@umn.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
40 ST. GEORGE STREET
TORONTO, ONTARIO M5S 2E4
CANADA
E-MAIL: panchenk@math.toronto.edu