Universality in spin glass models: chaos and ultrametricity

Antonio Auffinger*  Wei-Kuo Chen†
Northwestern University  University of Chicago

Abstract

We establish several results concerning the universality of spin glass models. First, in the mixed $p$-spin model, we prove that the Ghirlanda-Guerra identities hold if the disorder has zero mean and finite fourth moment. This establishes ultrametricity under these assumptions and therefore the long-standing belief among physicists that the Parisi solution in mean-field models is universal. Second, for both mixed $p$-spin models and the Edwards-Anderson (EA) model we establish disorder and temperature chaos phenomena for non-gaussian environments. Last, for random field Ising model we show that under a small external field and finite second moment the overlap is self-averaging, extending the results of [6, 10].

1 Introduction

It is widely expected that many statistical quantities and properties in disordered systems should not depend on the particular distribution of the environment. This phenomena, described as universality, is a major topic of research within the probability and mathematical physics communities. The simplest example is the central limit theorem, where the limiting distribution of sum of independent and identically distributed random variables is gaussian provided that these variables have just a finite second moment.

In spin glasses, universality is broadly accepted. Most of the fascinating predictions made by physicists should hold assuming that the environment satisfies some moment conditions. Among these predictions, the Parisi solution for mean-field spin glasses, which prescribes an ultrametric structure for the Gibbs measure in the thermodynamical limit, stands as one of the most ingenious and important ideas of the past decades.

Two major steps in the Parisi solution are known rigorously. First, the Parisi formula, which describes the limiting free energy [15, 19, 25], and, more recently, the proof of the ultrametricity conjecture by Panchenko [20]. In his celebrated work, Panchenko related the ultrametricity conjecture to the validity of the Ghirlanda-Guerra identities. These identities are known in several mean-field spin glass models with gaussian disorder, including the Sherrington-Kirkpatrick model and mixed $p$-spin models. They were first

*Department of Mathematics, Northwestern University. Email: auffing@math.northwestern.edu
†Department of Mathematics, University of Chicago. Email: wkchen@math.uchicago.edu
proved in [14] on average over the inverse temperature parameters and later, in a closely related formulation, by introducing a perturbation term to the Hamiltonian. For generic mixed spin-models they are known in a strong sense without perturbation [18].

The first main goal of this paper is to remove the hypothesis of gaussian disorder and to show that these identities are universal: we prove that they hold only assuming mean zero and finite fourth moment. This combined with the result of [20] establishes ultrametricity under these assumptions.

A second main direction in this manuscript is to establish chaos phenomena in a non-gaussian environment. Chaos is a classical old problem in spin glasses. It arose from the discovery that in some models, a small change in the external parameters, such as the disorder or the temperature implies a dramatic change to the overall energy landscape. This phenomenon was first predicted by Fisher and Huse [12] for the Edwards-Anderson model, although it seems the term chaos first appeared in the widely cited paper of Bray and Moore [1] on disorder chaos for the Sherrigton-Kirkpatrick model. It has attracted a lot of recent attention from mathematicians. Chaos in disorder for mixed even-spin models was established in [4, 5, 7]. Temperature chaos, a more intricate subject according to physicists [11], was obtained in [8, 9]. Despite the remarkable progress, all these results require gaussian disorder for both original and perturbed Hamiltonian. The second main goal of this paper is to remove this assumption. We establish both temperature and disorder chaos in the mixed \( p \)-spin model and in the EA model for environments with zero mean and finite second moment.

To be more precise, a typical way of measuring the instability of a spin system is to sample independently a configuration from a Gibbs measure \( G \) and a second configuration from a new Gibbs measure \( G' \) constructed as a perturbation of \( G \), and consider the behavior of the overlap of these two configurations under \( G \times G' \). The phenomenon of chaos states that this overlap behaves differently and is concentrated near a constant no matter if the two systems \( G \) and \( G' \) are in the high or low temperature regime.

A word of comment is needed here. In a first sight, one may be puzzled by the joint presence of universality and disorder chaos. Roughly speaking, universality means that statistics of configurations sampled independently from the same random Gibbs measure do not depend on the law of the disorder. On the other hand, if one sample two configurations from different Gibbs measures then they behave differently. Our results show that regardless of the law of the disorder, if one adds a perturbation to the Hamiltonian then chaos holds. In other words, we show that disorder chaos is also universal.

Our methods also allow us to study other universal properties of non mean-field models. In the EA model we show self-averaging of the magnetization under an assumption of finite third moment for the disorder and a decay rate for the external field. Under these same assumptions, in the random field Ising model, we extend the results of [6, 10] by showing self-averaging of the site overlap.

Rigorous progress on universality of spin glasses was obtained in the past. It is known that the limiting free energy for the Sherrington-Kirkpatrick does not depend on the particular distribution of the environment. This result was first obtained by Talagrand [24] who showed that Bernoulli and Gaussian disorder share the same limiting free energy and then generalized in two beautiful papers, one by Guerra and Toninelli [16] under assumption of symmetric laws with finite fourth moment and another by Carmona and
Hu [2] only assuming mean zero and finite third moment for the environment. The third moment condition was reduced to finite second moment in [3]. It was also extended to other types of mean-field models, for instance, the bipartite model [13].

The method of our proofs is as follows. As in the papers dealing with the limiting free energy, our first step is inspired in Guerra’s interpolation technique: we apply an approximate integration by parts lemma to an appropriate observable (that changes depending on the problem). The more moments we assume for the disorder, the better is this approximation. The core of the argument and the main novelty of our proofs is however on how we handle and control the error terms coming from these approximations. We do this by estimating appropriate derivatives, and although not very enlightening, the proof goes through an unavoidable sequence of careful computations.

In this regard, we hope to help the reader through the organization of the paper. In the next section, we show how the method works in a very simple setting, where all computations are almost straightforward. There, we establish self-averaging of the magnetization and overlap in the EA and random field Ising model, respectively. In section 3, we raise the stakes and prove universality of disorder chaos. Our main results, universality of ultrametricity and temperature chaos, are in section 4 where Lemma 2 plays a critical role and universality of the Gibbs measure is proven.

1.1 Acknowledgments

Both authors would like to thank the 2014 Midwest Probability Colloquium where part of this work was discussed. W.-K. C. would like to thank D. Panchenko for previous conversations. The research of A. A. is supported by NSF Grant DMS-1407554.

2 Self-averaging of the magnetization

Let $\Sigma$ be a finite set and $\mu$ be a random measure on $\Sigma$. For a given countable set $E$, consider a family of measurable functions $(f_e)_{e \in E}$ with $|f_e| \leq 1$ on $\Sigma$. Let $(y_e)_{e \in E}$ be independent random variables with mean zero, variance one and $B_3 := \sup_{e \in E} \mathbb{E}|y_e|^3 < \infty$. These are all independent of $\mu$. We consider the Hamiltonian for $\gamma \geq 0$,

$$H_y(\sigma) = \gamma \sum_{e \in E} y_e f_e(\sigma)$$

and define its Gibbs measure as

$$G_y(\sigma) = \frac{\exp H_y(\sigma)d\mu(\sigma)}{Z_y},$$

where $Z_y$ is the normalizing factor. Denote by the Gibbs average by $\langle \cdot \rangle_y$.

**Theorem 1.** Let $|E|$ be finite. Define the magnetization by

$$M(\sigma) = \frac{1}{|E|} \sum_{e \in E} f_e(\sigma).$$
If $\gamma > 0$, then we have that

$$\mathbb{E}\langle (M(\sigma) - \langle M(\sigma) \rangle_y)^2 \rangle_y \leq 9B_3\gamma + \frac{1}{\gamma\sqrt{|E|}}.$$  

**Lemma 1** (Approximate integration by parts). Let $y$ be a random variable such that its first $k$ moments match with those in the Gaussian random variable and $\mathbb{E}|y|^{k+1} < \infty$. Suppose that $F \in C^{k+1}(\mathbb{R})$ with $\|F^{(k)}\|_\infty < \infty$. Then

$$|\mathbb{E}yF(y) - \mathbb{E}F'(y)| \leq \frac{(k + 1)}{k!} \|F^{(k)}\|_\infty \mathbb{E}|y|^{k+1}.$$  

**Proof of Theorem 1.** A direct computation gives

$$\frac{d}{dy_e} \langle M(\sigma) \rangle = \gamma (\langle f_e(\sigma)M(\sigma) \rangle_y - \langle f_e(\sigma) \rangle \langle M(\sigma) \rangle_y),$$

$$\frac{d^2}{dy_{2,e}} \langle M(\sigma) \rangle = \gamma^2 (\langle f_e(\sigma)^2 M(\sigma) \rangle_y - \langle f_e(\sigma)M(\sigma) \rangle_y \langle f_e(\sigma) \rangle_y)
- \gamma^2 (\langle f_e(\sigma^1)M(\sigma^2)(f_e(\sigma^1) + f_e(\sigma^2)) \rangle_y - 2\langle f_e(\sigma^1)M(\sigma^2) \rangle_y \langle f_e(\sigma) \rangle_y).$$

Using approximate integration by parts with $k = 2$,

$$|\mathbb{E}y_e\langle M(\sigma) \rangle_y - \mathbb{E}(\langle f_e(\sigma)M(\sigma) \rangle_y - \langle f_e(\sigma) \rangle_y \langle M(\sigma) \rangle_y)| \leq 9B_3\gamma^2.$$  

Dividing this inequality by $\gamma|E|$ and summing over all $e \in E$, the triangle inequality gives

$$\mathbb{E}\langle (M(\sigma) - \langle M(\sigma) \rangle_y)^2 \rangle_y \leq 9B_3\gamma + \frac{1}{\gamma|E|} \mathbb{E}\left| \sum_{e \in E} y_e \right| \langle M(\sigma) \rangle_y
\leq 9B_3\gamma + \frac{1}{\gamma|E|} \mathbb{E}\left| \sum_{e \in E} y_e \right|
\leq 9B_3\gamma + \frac{1}{\gamma\sqrt{|E|}},$$

where the last inequality used Jensen’s inequality and the fact that $(y_e)_{e \in E}$ are independent with mean zero and variance one. \qed

### 2.1 Applications of Theorem 1

Now, we illustrate three applications of Theorem 1. We also use this subsection to introduce the three models we consider in the paper: the mixed $p$-spin model, the EA model and the random field Ising model.

**Example 1** (mixed $p$-spin model). Let $(\beta_p)_{p \geq 2}$ be a sequence of real numbers with $\sum_{p \geq 2} 2^p \beta_p^2 < \infty$ and $h \in \mathbb{R}$. For any $N \geq 1$, the Hamiltonian of the mixed $p$-spin model is defined as

$$X(\sigma) = \sum_{p \geq 2} \beta_p X_p(\sigma) + h \sum_{i=1}^{N} \sigma_i,$$
for all $\sigma = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N := \{-1, 1\}^N$, where $X_p$ is the pure $p$-spin Hamiltonian,

$$X_p(\sigma) = \frac{1}{N(p-1)/2} \sum_{1 \leq i_1, \ldots, i_p \leq N} y_{i_1, \ldots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}.$$

Here, $y_{i_1, \ldots, i_p}$'s are independent random variables with mean zero and variance one.

When $p = 2$, and the disorder $(y_{i_1, \ldots, i_p})$ is Gaussian we have the famous Sherrington-Kirkpatrick model [23]. We denote the Gibbs measure and the Gibbs expectation of the mixed $p$-spin model as $G$ and $\langle \cdot \rangle$. Suppose that $\beta_p > 0$ for some $p \geq 2$. Set $E = \{1, \ldots, N\}^p$ and $\gamma = \beta_p N^{-(p-1)/2}$. Define $y_e = y_{i_1, \ldots, i_p}$ and $f_e(\sigma) = \sigma_{i_1} \cdots \sigma_{i_p}$ if $e = (i_1, \ldots, i_p) \in E$. Then using this notation, $G$ can be rewritten as

$$G(\sigma) = \frac{\exp H_y(\sigma) d\mu(\sigma)}{Z_y},$$

where

$$d\mu(\sigma) = \exp(X(\sigma) - H_y(\sigma)).$$

Therefore, Theorem 1 implies that

$$\mathbb{E}\langle (m(\sigma)^p - \langle m(\sigma)^p \rangle)^2 \rangle \leq \frac{9\beta_p B_3}{N^{(p-1)/2}} + \frac{1}{\beta_p \sqrt{N}},$$

where $m(\sigma) = \sum_{i=1}^{N} \sigma_i / N$ is the usual magnetization. This example extends the result obtained in Gaussian disorder in [10, Example 5].

**Example 2** (Edwards-Anderson model). Let $A$ be a finite undirected graph with vertex set $v(A)$ and edge set $e(A)$. The Edwards-Anderson model with temperature $\beta$ and external field $\gamma$ is defined as

$$X(\sigma) = \beta \sum_{(i,j) \in e(A)} y_{i,j} \sigma_i \sigma_j + \gamma \sum_{i \in v(A)} y_i \sigma_i$$

for $\sigma \in \{-1, +1\}^{v(A)}$, where $(y_{i,j})_{(i,j) \in e(A)}$ and $(y_i)_{i \in v(A)}$ are independent random variables with mean zero and variance one. Let $G$ and $\langle \cdot \rangle$ be the Gibbs measure and Gibbs expectation associated to $X$. Set $E = v(A)$ and also $y_e = y_i$ and $f_e(\sigma) = \sigma_i$ if $e = i$. Then we can rewrite

$$G(\sigma) = \frac{\exp H_y(\sigma) \mu(\sigma)}{Z_y},$$

where

$$\mu(\sigma) = \exp(X(\sigma) - H_y(\sigma)).$$

From Theorem 1, for all $\gamma > 0$,

$$\mathbb{E}\langle (m(\sigma) - \langle m(\sigma) \rangle)^2 \rangle \leq 9B_3 \gamma + \frac{1}{\gamma \sqrt{|v(A)|}},$$

5
where \( m(\sigma) := |v(A)|^{-1} \sum_{i \in v(A)} \sigma_i \). If the external field tends to zero in the rate that \( \gamma = (3B_3^{1/2}|v(A)|^{1/4})^{-1} \), then the magnetization is self-averaged,

\[
\mathbb{E}(\langle m(\sigma) - \langle m(\sigma) \rangle \rangle \leq \frac{6B_3^{1/2}}{|v(A)|^{1/4}}.
\]

**Example 3** (Random field Ising model). Let \( A \) be the undirected graph with vertex set \( v(A) = \mathbb{Z}^d \cap [0,N]^d \) and edge set \( e(A) = \{(i,j) : i,j \in v(A) \text{ with } |i-j| = 1 \} \). Let \( \beta, \gamma > 0 \). The Hamiltonian of the random field Ising model on \( A \) is defined as

\[
X(\sigma) = \beta \sum_{(i,j) \in e(A)} \sigma_i \sigma_j + \gamma \sum_{i \in v(A)} y_i \sigma_i
\]

for \( \sigma \in \{-1,+1\}^{v(A)} \), where \((y_i)_{i \in v(A)}\) are independent random variables with mean zero and variance one. As before, we use \( \langle \cdot \rangle \) to denote the Gibbs expectation with respect to the Hamiltonian. One important feature in this model is that its spin correlation satisfies the FKG inequality,

\[
\langle \sigma_i \sigma_j \rangle \geq \langle \sigma_i \rangle \langle \sigma_j \rangle
\]

for all \( i, j \in v(A) \). This implies that the site overlap \( R_{1,2} = |v(A)|^{-1} \sum_{i \in v(A)} \sigma_i^1 \sigma_i^2 \) satisfies

\[
\mathbb{E}(\langle (R_{1,2} - \langle R_{1,2} \rangle)^2 \rangle) = \frac{1}{|v(A)|^2} \sum_{i,j} \mathbb{E}(\langle \sigma_i^1 \sigma_j^2 \rangle^2 - \langle \sigma_i^2 \rangle^2 \langle \sigma_j^2 \rangle^2)
\]

\[
= \frac{1}{|v(A)|^2} \sum_{i,j} \mathbb{E} \big| \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \big| \big| \langle \sigma_i \sigma_j \rangle + \langle \sigma_i \rangle \langle \sigma_j \rangle \big|
\]

\[
\leq \frac{2}{|v(A)|^2} \sum_{i,j} \mathbb{E} \big| \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \big|
\]

\[
= \frac{2}{|v(A)|^2} \sum_{i,j} \mathbb{E}(\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle)
\]

\[
= 2\mathbb{E}(\langle m(\sigma) - \langle m(\sigma) \rangle \rangle^2),
\]

where the third equality used the FKG inequality (4). Let \( \gamma = (3B_3^{1/2}|v(A)|^{1/4})^{-1} \). Following the same argument as in the previous example, (3) holds for the random field Ising model and thus, we obtain the self-averaging of the overlap,

\[
\mathbb{E}(\langle (R_{1,2} - \langle R_{1,2} \rangle)^2 \rangle) \leq \frac{12B_3^{1/2}}{|v(A)|^{1/4}}.
\]

### 3 Quenched disorder chaos

In this section we prove disorder chaos for an environment with mean zero, variance one and finite third moment. Recall the notation used in section 2. Let \((y_{1,e})_{e \in E}\) and \((y_{2,e})_{e \in E}\) be independent random variables with mean zero and variance one. We assume that these
are also independent of \((y_e)_{e \in E}\). Denote \(B_{1,3} = \sup_{e \in E} \mathbb{E}|y_{1,e}|^3\) and \(B_{2,3} = \sup_{e \in E} \mathbb{E}|y_{2,e}|^3\). For \(0 \leq t \leq 1\), we consider the Hamiltonians on \(\Sigma\),

\[
H_{1,y}(\rho) = \gamma_1 \sum_{e \in E} f_\epsilon(\rho) (\sqrt{1 - ty_e} + \sqrt{1 - ty_{1,e}}), \\
H_{2,y}(\tau) = \gamma_2 \sum_{e \in E} f_\epsilon(\tau) (\sqrt{1 - ty_e} + \sqrt{1 - ty_{2,e}}).
\]

(5)

In other words, \(H_{1,y}\) and \(H_{2,y}\) are slightly decoupled through the parameter \(t\). Let \(\mu_1, \mu_2\) be two random measures on \(\Sigma\). Set the Gibbs measures as

\[
dG_{1,y}(\rho) = \frac{\exp H_{1,y}(\rho) d\mu_1(\rho)}{Z_{1,y}} , \\
dG_{2,y}(\tau) = \frac{\exp H_{2,y}(\tau) d\mu_2(\tau)}{Z_{2,y}},
\]

(6)

where \(\gamma_1, \gamma_2 \geq 0\) and \(Z_{1,y}, Z_{2,y}\) are the normalizing factors. Denote by \((\rho^\ell, \tau^\ell)_{\ell \geq 1}\) a sequence of i.i.d. samplings from the product measure \(G_{1,y} \times G_{2,y}\). We shall again use the notation \(\langle \cdot \rangle_y\) to stand for the Gibbs expectation but with respect to \(\prod_{\ell \geq 1} (G_{1,y} \times G_{2,y})\). As explained in the introduction, the quantity of great interest is the overlap between the two systems,

\[
Q_{\epsilon, \epsilon'} = \frac{1}{|E|} \sum_{e \in E} f_\epsilon(\rho^\ell) f_\epsilon(\tau^\ell).
\]

We formulate the main result of this section in a way that will cover the mixed p-spin and the EA models as examples.

**Theorem 2.** For any \(\gamma_1, \gamma_2 > 0\) and \(t \in [0, 1]\), we have

\[
\mathbb{E}(\langle Q_{1,1} - \langle Q_{1,1}\rangle_y \rangle_y^2) \leq 36 (B_{1,3} \gamma_1 + B_{2,3} \gamma_2) \sqrt{1 - t} + \frac{4 (\gamma_1 + \gamma_2)}{\gamma_1 \gamma_2 \sqrt{|E|}} \left(1 - \frac{t}{1 - t}\right).
\]

**Proof.** Note that the first and second derivatives of \(\langle Q_{1,1} f_\epsilon(\rho^1) \rangle_y\) with respect to the variable \(y_{2,e}\) equal

\[
\frac{d}{dy_{2,e}} \langle Q_{1,1} f_\epsilon(\rho^1) \rangle_y = \sqrt{1 - t} \gamma_2 \langle Q_{1,1} f_\epsilon(\rho^1)(f_\epsilon(\tau^1) - f_\epsilon(\tau^2)) \rangle_y,
\]

\[
\frac{d^2}{dy_{2,e}^2} \langle Q_{1,1} f_\epsilon(\rho^1) \rangle_y = (1 - t) \gamma_2^2 \langle Q_{1,1} f_\epsilon(\rho^1)(f_\epsilon(\tau^1)^2 - f_\epsilon(\tau^1)f_\epsilon(\tau^2)) \rangle_y.
\]

\[
- (1 - t) \gamma_2^2 \langle Q_{1,1} f_\epsilon(\rho^1)(f_\epsilon(\tau^2)(f_\epsilon(\tau^1) + f_\epsilon(\tau^2)) - f_\epsilon(\tau^3) - f_\epsilon(\tau^4)) \rangle_y.
\]

From the approximate integration by parts with \(k = 3\),

\[
|\mathbb{E}(y_{2,e} Q_{1,1} f_\epsilon(\rho^1))_y - \gamma_2 \sqrt{1 - t} \mathbb{E}(Q_{1,1} f_\epsilon(\rho^1)(f_\epsilon(\tau^1) - f_\epsilon(\tau^2))_y)| \leq 9 B_{2,3} (1 - t) \gamma_2^2.
\]

Summing over all \(e\) and dividing by \(|E| \gamma_2 \sqrt{1 - t}\), the triangle inequality gives

\[
\mathbb{E}(\langle Q_{1,1}^2 - Q_{1,1} Q_{1,2} \rangle_y) \leq 9 B_{2,3} \gamma_2 \sqrt{1 - t} + \frac{1}{\gamma_2 |E| \sqrt{1 - t}} \mathbb{E}(\langle Q_{1,1} \sum_{e \in E} y_{2,e} f_\epsilon(\rho^1) \rangle_y).
\]
Observe that $|E(Q_{1,1} \sum_{e \in E} y_{2,e} f_e(\rho^1))| \leq E(\sum_{e \in E} y_{2,e} f_e(\rho^1))$. Since the Gibbs expectation is only with respect to $\rho^1$ and the disorder in Hamiltonian $H_{1,y}$ is independent of $y_{2,e}$, using conditional expectation and then the Cauchy-Schwarz inequality implies

$$
\left| E\left(Q_{1,1} \left( \sum_{e \in E} y_{2,e} f_e(\rho^1) \right) \right) \right| \leq E\left( E_2 \left( \sum_{e \in E} y_{2,e} f_e(\rho^1) \right) \right)
$$

$$\leq E\left( \left( E_2 \left( \sum_{e \in E} y_{2,e} f_e(\rho^1) \right)^2 \right)^{1/2} \right)
$$

$$= E\left( \left( \sum_{e \in E} f_e(\rho^1)^2 \right)^{1/2} \right) \leq \sqrt{|E|},
$$

where $E_2$ is the expectation with respect to the randomness $(y_{2,e})_{e \in E}$. Thus, we have

$$E(\langle Q_{1,1}^2 - Q_{1,1}Q_{1,2} \rangle_y) \leq 9B_{2,3}\gamma_2 \sqrt{1 - t} + \frac{1}{\gamma_2 \sqrt{|E|}(1 - t)}$$

and similarly,

$$E(\langle Q_{2,2}^2 - Q_{2,2}Q_{1,2} \rangle_y) \leq 9B_{1,3}\gamma_1 \sqrt{1 - t} + \frac{1}{\gamma_1 \sqrt{|E|}(1 - t)}.$$

Finally, these two inequalities lead to

$$E(\langle Q_{1,1} - \langle Q_{1,1} \rangle_y \rangle^2) \leq E(\langle Q_{1,1} - Q_{2,2} \rangle^2)$$

$$\leq 2E(\langle Q_{1,1} - Q_{1,2} \rangle^2) + 2E(\langle Q_{1,2} - Q_{2,2} \rangle^2)$$

$$= 4E(\langle Q_{1,1}^2 - Q_{1,1}Q_{1,2} \rangle) + 4E(\langle Q_{2,2}^2 - Q_{2,2}Q_{1,2} \rangle)$$

$$\leq 36(B_{1,3}\gamma_1 + B_{2,3}\gamma_2) + 4(\gamma_1 + \gamma_2) \sqrt{1 - t} + \frac{4(\gamma_1 + \gamma_2)}{\gamma_1 \gamma_2 \sqrt{|E|}(1 - t)},$$

which finishes our proof. \qed

**Example 4** (Quenched disorder chaos in the mixed $p$-spin model). Recall the Hamiltonians from $X$ and $X_p$ from Example 1. For each $p \geq 2$, consider two pure $p$-spin Hamiltonians on $\Sigma_N$,

$$X_{1,p}(\rho) = \frac{1}{N(p-1)/2} \sum_{1 \leq i_1, \ldots, i_p \leq N} y_{1,i_1,\ldots,i_p} \rho_{i_1} \cdots \rho_{i_p},$$

$$X_{2,p}(\tau) = \frac{1}{N(p-1)/2} \sum_{1 \leq i_1, \ldots, i_p \leq N} y_{2,i_1,\ldots,i_p} \tau_{i_1} \cdots \tau_{i_p}.$$

Here the random variables $y_{1,i_1,\ldots,i_p}$ and $y_{2,i_1,\ldots,i_p}$ are independent random variables with mean zero and variance one for all $i_1, \ldots, i_p$ and $p \geq 2$. These are also independent of $(y_e)_{e \in E}$. Let $(\beta_{1,p})_{p \geq 2}$ and $(\beta_{2,p})_{p \geq 2}$ be two nonnegative sequences with finite $\sum_{p \geq 2} 2^p \beta_{1,p}$ and $\sum_{p \geq 2} 2^p \beta_{2,p}$, $h_1, h_2 \in \mathbb{R}$ and $(t_p)_{p \geq 2}$ be a sequence with $0 \leq t_p \leq 1$ for $p \geq 2$. We
consider two mixed $p$-spin models,

$$X_1(\rho) = \sum_{p \geq 2} \beta_{1,p} \left( \sqrt{t_p} X_p(\rho) + \sqrt{1 - t_p} X_{1,p}(\rho) \right) + h_1 \sum_{1 \leq i \leq N} \rho_i, \quad (7)$$

$$X_2(\tau) = \sum_{p \geq 2} \beta_{2,p} \left( \sqrt{t_p} X_p(\tau) + \sqrt{1 - t_p} X_{2,p}(\tau) \right) + h_2 \sum_{1 \leq i \leq N} \tau_i. \quad (8)$$

and set their Gibbs measures as

$$G_1(\rho) = \frac{\exp X_1(\rho)}{Z_1},$$
$$G_2(\tau) = \frac{\exp X_2(\tau)}{Z_2},$$

where $Z_1, Z_2$ are the partition functions. Let $\langle \cdot \rangle$ denote the Gibbs expectation with respect to $G_1 \times G_2$.

Suppose that $\beta_{1,p}, \beta_{2,p} > 0$ and $0 \leq t_p < 1$ for some $p \geq 2$. In the notations of (5) and (6), we set $E = \{1, \ldots, N\}^p$ and $t = t_p$. Set for all $e = (i_1, \ldots, i_p) \in E$, $f_e(\sigma) = \sigma_{i_1} \cdots \sigma_{i_p}$, $y_{1,e} = y_{1,i_1,\ldots,i_p}$, $y_{2,e} = y_{2,i_2,\ldots,i_p}$.

We can then rewrite

$$G_1(\rho) = \frac{\exp H_{1,y}(\rho) \mu_1(\rho)}{Z_{1,y}},$$
$$G_2(\tau) = \frac{\exp H_{2,y}(\tau) \mu_2(\tau)}{Z_{2,y}},$$

where $\gamma_1 = \beta_{1,p} N^{-(p-1)/2}$, $\gamma_2 = \beta_{2,p} N^{-(p-1)/2}$ and

$$\mu_1(\rho) = \exp(X_1(\rho) - H_{1,y}(\rho)), \quad \mu_2(\tau) = \exp(X_2(\tau) - H_{2,y}(\tau)).$$

Applying Theorem 2, we obtain

$$\mathbb{E}(\langle R_{1,1}^p \rangle - \langle R_{1,1}^p \rangle^2) \leq \frac{36(B_{1,3} \beta_{1,p} + B_{2,3} \beta_{2,p})}{N(p-1)/2} \sqrt{1 - t_p} + \frac{4(\beta_{1,p} + \beta_{2,p})}{\beta_{1,p} \beta_{2,p} \sqrt{N(1 - t_p)}},$$

where $R_{1,1} = N^{-1} \sum_{i=1}^N \rho_i^1 \tau_i^1$ is the overlap between the two systems.

**Example 5** (Quenched disorder chaos in the Edwards-Anderson model). Recall the random variables $(y_{(i,j)})_{(i,j) \in e(A)}$ and $(y_i)_{i \in v(A)}$ in the Edwards-Anderson model from Example 2. Consider independent random variables $(y_{1,(i,j)})_{(i,j) \in e(A)}$, $(y_{2,(i,j)})_{(i,j) \in e(A)}$, $(y_{1,i})_{i \in e(A)}$.
and \((y_{2,i})_{i \in v(A)}\) with mean zero and variance one. For \(\beta_1, \beta_2, h_1, h_2 \geq 0\) and \(0 \leq t_b, t_s \leq 1\), we consider two Hamiltonians of the Edwards-Anderson model,

\[
X_1(\rho) = \beta_1 \sum_{(i,j) \in e(A)} (\sqrt{t_b}y_{(i,j)} + \sqrt{1-t_e}y_{1,(i,j)})\rho_i\rho_j + h_1 \sum_{i \in v(A)} (\sqrt{t_s}y_i + \sqrt{1-t_v}y_{1,i})\rho_i,
\]

\[
X_2(\tau) = \beta_2 \sum_{(i,j) \in e(A)} (\sqrt{t_b}y_{(i,j)} + \sqrt{1-t_e}y_{2,(i,j)})\tau_i\tau_j + h_2 \sum_{i \in v(A)} (\sqrt{t_s}y_i + \sqrt{1-t_v}y_{2,i})\tau_i.
\]

Let

\[
B_{1,3}^b = \max_{(i,j) \in e(A)} \mathbb{E}|y_{1,(i,j)}|^3, \quad B_{1,3}^s = \max_{i \in v(A)} \mathbb{E}|y_{1,i}|^3,
\]

\[
B_{2,3}^b = \max_{(i,j) \in e(A)} \mathbb{E}|y_{2,(i,j)}|^3, \quad B_{2,3}^s = \max_{i \in v(A)} \mathbb{E}|y_{2,i}|^3.
\]

Following similar arguments as in Example 4, we get that if \(\beta_1, \beta_2 > 0\) and \(0 \leq t_e < 1\), then

\[
\mathbb{E}((Q_{1,1}^b - \langle Q_{1,1}^b \rangle)^2) \leq 36(B_{1,3}^b,\beta_1 + B_{2,3}^s,\beta_2)\sqrt{1-t_b} + \frac{(\beta_1 + \beta_2)}{\beta_1\beta_2\sqrt{\mathbb{E}(|e(A)|)(1 - t_b)}},
\]

where \(Q_{1,1}^b = |e(A)|^{-1} \sum_{(i,j) \in e(A)} \rho_i^1 \rho_j^1 \tau_i^1 \tau_j^1\) is the bond overlap. If

\[
\lim_{|e(A)| \to \infty} t_b = 1 \quad \text{and} \quad \lim_{|e(A)| \to \infty} |e(A)|(1 - t_b) = \infty,
\]

then \(Q_{1,1}^b\) is essentially concentrated around its quenched average. Similarly, if \(h_1, h_2 > 0\) and \(0 \leq t_s < 1\), then

\[
\mathbb{E}((Q_{1,1}^s - \langle Q_{1,1}^s \rangle)^2) \leq 36(B_{1,3}^s, h_1 + B_{2,3}^s, h_2)\sqrt{1-t_s} + \frac{(\beta_1 + \beta_2)}{\beta_1\beta_2\sqrt{\mathbb{E}(|e(A)|)(1 - t_s)}},
\]

where \(Q_{1,1}^s = |v(A)|^{-1} \sum_{i \in v(A)} \rho_i^1 \tau_i^1\) is the site overlap. If

\[
\lim_{|v(A)| \to \infty} t_s = 1 \quad \text{and} \quad \lim_{|v(A)| \to \infty} |v(A)|(1 - t_s) = \infty,
\]

then \(Q_{1,1}^s\) is concentrated around its quenched average.

### 4 Universality of the Gibbs measure

In this section, we establish our two main results: ultrametricity and temperature chaos for disorders with finite fourth moment.

#### 4.1 Ultrametricity

Recall the Hamiltonian \(H_y\) from (5). For \(k \geq 2\), we assume that \((y_e)_{e \in E}\) in \(H_y\) are independent of each other such that their first \(k\) moments match with those in the standard
Gaussian random variable and $B_{k+1} := \sup_{e \in E} \mathbb{E}|y_e|^k < \infty$. In addition, we allow $\gamma$ in $H_y$ to depend on $e \in E$. Now set the Hamiltonian

$$H_y(\sigma) = \sum_{e \in E} \gamma_e y_e f_e(\sigma),$$

where $\sum_{e \in E} \gamma_e^2 < \infty$. We again use $G_y$ and $\langle \cdot \rangle_y$ to denote the Gibbs measure and the Gibbs expectation associated to this Hamiltonian. In particular, if $(y_e)_{e \in E}$ are i.i.d. standard Gaussian random variables $(g_e)_{e \in E}$, we shall denote all these by $H_g$, $G_g$ and $\langle \cdot \rangle_g$.

The results of this section will be consequences of the following theorem.

**Theorem 3.** Suppose that $L$ is a measurable function on $\Sigma^n$ with $\|L\|_\infty \leq 1$. We have that

$$|\mathbb{E}(L)_g - \mathbb{E}(L)_y| \leq D_{k,n} \sum_{e \in E} \gamma_e^{k+1},$$

where $D_{k,n} := 2^{k(k+3)/2} n^{k+1} B_{k+1}$.

**Lemma 2.** Let $\nu$ be a measure on $\Sigma$. Suppose that $f$ is a measurable function on $\Sigma$ with $|f| \leq 1$ and $L$ is a measurable function on $\Sigma^n$ with $\|L\|_\infty \leq 1$. Consider the Gibbs measure

$$G(\sigma) = \frac{\exp(\gamma x f(\sigma)) \nu(\sigma)}{Z},$$

where $Z$ is the normalizing factor. Denote by $\langle \cdot \rangle_x$ the Gibbs expectation associated to $G$. For any $k \geq 1$, we have

$$\left| \frac{d^k}{dx^k} \langle L \rangle_x \right| \leq \gamma^k C_{k,n},$$

where $C_{k,n} := 2^{k(k+1)/2} n^k$.

**Proof.** We claim that for $k \geq 1$,

$$\frac{d^k}{dx^k} \langle L \rangle_x = \gamma^k \sum_{\ell_1=1}^{2^n} \sum_{\ell_2=1}^{2^n} \cdots \sum_{\ell_k=1}^{2^n} c_{\ell_1, \ell_2, \ldots, \ell_k} \langle L f(\sigma^{\ell_1}) f(\sigma^{\ell_2}) \cdots f(\sigma^{\ell_k}) \rangle_x,$$

for some constants $|c_{\ell_1, \ell_2, \ldots, \ell_k}| = 1$. If this holds, then clearly,

$$\left| \frac{d^k}{dx^k} \langle L \rangle_x \right| \leq \gamma^k \prod_{i=1}^k (2^n) = \gamma^k 2^{k(k+1)/2} n^k = \gamma^k C_{k,n}.$$

To show (9), we proceed by induction. If $k = 1$, then

$$\frac{d}{dx} \langle L \rangle_x = \gamma \sum_{\ell=1}^n \langle L f(\sigma^{\ell}) - f(\sigma^{\ell+n}) \rangle_x = \gamma \sum_{\ell=1}^{2^n} c_{\ell} \langle L f(\sigma^{\ell}) \rangle_x,$$
where \( c_{\ell_1} = 1 \) for \( 1 \leq \ell_1 \leq n \) and \(-1\) for \( n + 1 \leq \ell_1 \leq 2n \). Assume that (9) holds for some \( k \geq 1 \). For any \( Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k}) \), we shall regard it as a function depending on \((\sigma^\ell)_{1 \leq \ell \leq 2^kn}\) and we compute directly to get

\[
\frac{d}{dx} \langle Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k}) \rangle_x = \gamma \sum_{\ell=1}^{2^kn} \langle Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k})(f(\sigma^{\ell}) - f(\sigma^{2^kn+\ell})) \rangle_x
\]

\[
= \gamma \sum_{\ell=1}^{2^kn} d_{\ell_1,\ldots,\ell_k,\ell_{k+1}} \langle Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k})f(\sigma^{\ell}) \rangle_x,
\]

where \( d_{\ell_1,\ldots,\ell_k,\ell_{k+1}} = 1 \) if \( 1 \leq \ell_{k+1} \leq 2^kn \) and \(-1\) if \( 2^kn \leq \ell_{k+1} \leq 2^{k+1}n \). It follows that

\[
\frac{d^{k+1}}{dx^{k+1}} \langle L \rangle_x = \gamma^{k+1} \sum_{\ell_1=1}^{2^n} \cdots \sum_{\ell_k=1}^{2^n} c_{\ell_1,\ell_2,\ldots,\ell_k} \sum_{\ell=1}^{2^{k+1}n} d_{\ell_1,\ldots,\ell_k,\ell_{k+1}} \langle Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k})f(\sigma^{\ell}) \rangle_x
\]

\[
= \gamma^{k+1} \sum_{\ell_1=1}^{2^n} \cdots \sum_{\ell_k=1}^{2^n} \sum_{\ell=1}^{2^{k+1}n} c_{\ell_1,\ell_2,\ldots,\ell_k} d_{\ell_1,\ldots,\ell_k,\ell_{k+1}} \langle Lf(\sigma^{\ell_1}) \cdots f(\sigma^{\ell_k})f(\sigma^{\ell}) \rangle_x.
\]

Letting \( c_{\ell_1,\ldots,\ell_k,\ell_{k+1}} = c_{\ell_1,\ell_2,\ldots,\ell_k} d_{\ell_1,\ldots,\ell_k,\ell_{k+1}} \) implies (9) in the case of \( k+1 \). This completes the proof of our claim. \( \square \)

**Proof of Theorem 3.** Consider the interpolated Hamiltonian between \( H_g \) and \( H_y \),

\[
H_\kappa(\sigma) = \sum_{e \in E} \gamma_e (\sqrt{s_e} + \sqrt{1 - s_e}) f_e(\sigma).
\]

Let \( \langle \cdot \rangle_\kappa \) be the corresponding Gibbs average and set \( \phi(s) = \mathbb{E}(L)_s \). A direct computation gives that

\[
\phi'(t) = \sum_{e \in E} \sum_{\ell=1}^n \gamma_e \mathbb{E}\left( L_{e,\ell} \left( \frac{g_e}{\sqrt{t}} - \frac{y_e}{\sqrt{1-t}} \right) \right)_s,
\]

where \( L_{e,\ell} = 2^{-1}L(f_e(\sigma^\ell) - f_e(\sigma^{\ell+n})) \). Note that \( L_{e,\ell} \) is a function of \((\sigma^\ell)_{1 \leq \ell \leq n}\) with \(|L_{e,\ell}| \leq 1\). We shall think of \( L_{e,\ell} \) as a function depending on \((\sigma)_{1 \leq \ell \leq 2n}\). For each \( e \in E \), using Gaussian integration by parts,

\[
\mathbb{E}(L_{e,\ell}g_e)_t = \sqrt{t} \gamma_e \sum_{\ell=1}^{2n} \mathbb{E}(L_{e,\ell}(f_e(\sigma^\ell) - f_e(\sigma^{\ell+2n})))_s
\]

and from the approximate integration by parts together with Lemma 2, we obtain

\[
\left| \mathbb{E}(L_{e,\ell}y_e)_t - \sqrt{1-t} \gamma_e \sum_{\ell=1}^{2n} \mathbb{E}(L_{e,\ell}(f_e(\sigma^\ell) - f_e(\sigma^{\ell+n})))_t \right| \leq \gamma_e^k (1-t)^{k/2} C_{k,2n} \mathbb{E}|y_e|^{k+1}.
\]
Combining these two inequalities together and noting that $0 \leq t \leq 1$, the triangle inequality yields

$$\left| \frac{1}{\sqrt{s}} \mathbb{E}\langle L_e, t g_e \rangle_s - \frac{1}{\sqrt{1-s}} \mathbb{E}\langle L_e, t y_e \rangle_s \right| \leq \gamma_e^k (1 - s)^{(k-1)/2} C_{k,2n} \mathbb{E}|y_e|^{k+1} \leq \gamma_e^k C_{k,2n} B_{k+1} \quad (10)$$

and consequently,

$$|\phi'(s)| \leq n C_{k,2n} B_{k+1} \sum_{e \in E} \gamma_e^{k+1},$$

which clearly gives the announced result.

\( \square \)

**Example 6** (mixed $p$-spin model). Set $E_p = \{1, \ldots, N\}^p$ for all $p \geq 2$ and $E = \bigcup_{p \geq 1} E_p$. Let $\gamma_e = \beta_p N^{-\left(p-1\right)/2}$ and $f_e(\sigma) = \sigma_i \cdots \sigma_p$, if $e = (i_1, \ldots, i_p) \in E_p$. The Hamiltonian of the mixed $p$-spin model in Example 1 can be written as

$$X(\sigma) = H_y(\sigma) + h \sum_{i=1}^N \sigma_i.$$

Since

$$\sum_{e \in E} \gamma_e^{k+1} \mathbb{E}|y_e|^{k+1} = \sum_{p \geq 2} \frac{\beta_p^{k+1}}{N^{(p-1)/2}} \sum_{e \in E_p} \mathbb{E}|y_e|^{k+1} \leq B_{k+1} \sum_{p \geq 2} \frac{\beta_p^{k+1}}{N^{(k+1)(p-1)/2}} N^p = B_{k+1} \sum_{p \geq 2} \frac{\beta_p^{k+1}}{N^{(k+1)(p-1)/2-p}},$$

Theorem 3 yields

$$|\mathbb{E}\langle L \rangle_g - \mathbb{E}\langle L \rangle_y| \leq B_{k+1} D_{n,k} \sum_{p \geq 2} \frac{\beta_p^{k+1}}{N^{(k+1)(p-1)/2-p}}.$$

There are two cases that are of great importance. First, if $\beta_p = 0$ for $p = 2, 3$,

$$|\mathbb{E}\langle L \rangle_g - \mathbb{E}\langle L \rangle_y| \leq B_{k+1} D_{n,k} \sum_{p \geq 4} \frac{\beta_p^{k+1}}{N^{(k+1)(p-1)/2-p}} \leq \frac{B_{k+1} D_{n,k}}{N^{p/2-3/2}} \sum_{p \geq 4} \beta_p^{k+1} \rightarrow 0 \quad (11)$$

as $N \rightarrow \infty$, where the last inequality used $k \geq 2$. Second, if $\beta_2 \neq 0$, then the moment matching assumption $k \geq 4$ gives that

$$|\mathbb{E}\langle L \rangle_g - \mathbb{E}\langle L \rangle_y| \leq B_{k+1} D_{n,k} \sum_{p \geq 2} \frac{\beta_p^{k+1}}{N^{(k+1)(p-1)/2-p}} \leq \frac{B_{k+1} D_{n,k}}{N^{(k-3)/2}} \sum_{p \geq 2} \beta_p^{k+1} \rightarrow 0 \quad (12)$$

as $N \rightarrow \infty$, where the last inequality used

$$\frac{1}{N^{(k+1)(p-1)/2-p}} \leq \frac{1}{N^{(k+1)/2}} = \frac{1}{N^{(k-3)/2}}.$$

In other words, the behavior of the mixed $p$-spin model at the level of Gibbs measure is indeed universal in both cases. This implies the following result.
Theorem 4. Consider the mixed $p$-spin model with i.i.d. disorder $Y := (y_{1,\ldots,i_p}), p \geq 2$. Assume that the first four moments of the disorder match with the first four moments of standard Gaussian. If the Ghirlanda-Guerra identities hold for the model with Gaussian disorder then it holds for the model with disorder $Y$.

Proof. A sequence of random measure $G_N$ is said to satisfy the Ghirlanda-Guerra identities if for any $n \geq 2$, any bounded measurable function $f$ of the overlaps $(R_{l,l'})_{l,l' \leq n} := N^{-1} \sum_{i=1}^N \sigma_l \cdot \sigma_{l'}$ and any bounded measurable function $\psi$ of one overlap,

$$\left| \mathbb{E}\langle f \psi(R_{1,n+1}) \rangle - \frac{1}{n} \mathbb{E}\langle f \rangle \mathbb{E}\langle \psi(R_{1,2}) \rangle - \frac{1}{n} \mathbb{E}\langle f \psi(R_{1,1}) \rangle \right| \to 0 \quad \text{as} \quad N \to \infty$$  \hspace{1cm} (13)

where we used $\langle \cdot \rangle$ to denote expectation with respect to the Gibbs average $G_N$. Now one can see that if (13) holds for the model with Gaussian disorder then (12) implies that it holds for the model with disorder $Y$. $\square$

Remark 1. If $\beta_2 = \beta_3 = 0$ then the result of the theorem above holds under the assumption of mean zero and variance one for the disorder $Y$. (See (11)).

Remark 2. As mentioned in the introduction, the GG identities as stated in (13) are known for generic $p$-spin models ($\beta_p \neq 0$, for all $p \geq 2$). Under a perturbation of the temperature parameters, they are also known for all mixed $p$-spin models, including the Sherrington-Kirkpatrick and the pure $p$-spin model.

Combining with of [20, Theorem 1], we obtain:

Corollary 1. Ultrametricity holds under the assumptions of Theorem 4.

4.2 Coupled Gibbs measures and temperature chaos

We now proceed to establish the universality for the coupled Gibbs measure. Recall the Hamiltonians $H_{1,y}$ and $H_{2,y}$ in (5). As above, we assume that the first $k$ moments of the random variables $(y_e), (y_{1,e})$ and $(y_{2,e})$ match with those in the standard Gaussian random variable. In addition, the parameters $t, \gamma_1, \gamma_2$ in $H_{1,y}, H_{2,y}$ are allowed to depend on $e \in E$. Consider the following generalized Hamiltonians,

$$H_{1,y}(\rho) = \sum_{e \in E} \gamma_{1,e} f_e(\rho)(\sqrt{t_e y_e} + \sqrt{1-t_e y_{1,e}}),$$
$$H_{2,y}(\tau) = \sum_{e \in E} \gamma_{2,e} f_e(\tau)(\sqrt{t_e y_e} + \sqrt{1-t_e y_{2,e}}),$$

where $(t_e) \subset [0, 1], \sum \gamma_{1,e}^2$ and $\sum \gamma_{2,e}^2$ are finite. We again use the notation $G_{1,y}, G_{2,y}$ to denote the Gibbs measures associated to these two Hamiltonians and $\langle \cdot \rangle_y$ is the Gibbs expectation for $\prod_{e=1}^{\infty} (G_{1,y} \times G_{2,y})$. In the case that $(y_e), (y_{1,e})$ and $(y_{2,e})$ are i.i.d. standard Gaussian random variables $(g_e), (g_{1,e}), (g_{2,e})$, we shall use the notations $H_{1,g}, H_{2,g}$ and $\langle \cdot \rangle_g$. 

14
Set

\[ B_{k+1} = \sup_{e \in E} \mathbb{E}|y_e|^{k+1}, \]
\[ B_{1,k+1} = \sup_{e \in E} \mathbb{E}|y_{1,e}|^{k+1}, \]
\[ B_{2,k+1} = \sup_{e \in E} \mathbb{E}|y_{2,e}|^{k+1}. \]

**Theorem 5.** Let \( L \) be a function depending on \((\rho^\ell, \tau^\ell)_{\ell \leq n}\) with \( \|L\|_\infty \leq 1 \). We have that

\[ |\mathbb{E}(L)_g - \mathbb{E}(L)_y| \leq D'_{k,n} \sum_{e \in E} (\gamma_{1,e} + \gamma_{2,e})^{k+1}, \]

where \( D'_{k,n} := 2nC_{k,2n}(B_{k+1} + B_{1,k+1} + B_{2,k+1}) \).

**Proof.** Consider the interpolated Hamiltonian for \(0 \leq s \leq 1\),

\[ H_s(\rho, \tau) = \sqrt{s}(H_{1,g}(\rho) + H_{2,g}(\tau)) + \sqrt{1-s}(H_{1,y}(\rho) + H_{2,y}(\tau)). \]

Denote by \( \langle \cdot \rangle_s \) its Gibbs expectation and set \( \phi(s) = \mathbb{E}(L)_s \). Note that \( \phi(0) = \mathbb{E}(L)_y \) and \( \phi(1) = \mathbb{E}(L)_g \). Computing directly gives

\[ \phi'(s) = \sum_{e \in E} \sum_{\ell \leq n} \gamma_{1,e} \sqrt{t_e} \mathbb{E}\langle L_{1,e,\ell} \left( \frac{g_{e}}{\sqrt{s}} - \frac{y_{e}}{\sqrt{1-s}} \right) \rangle_s \]
\[ + \sum_{e \in E} \sum_{\ell \leq n} \gamma_{2,e} \sqrt{t_e} \mathbb{E}\langle L_{2,e,\ell} \left( \frac{g_{e}}{\sqrt{s}} - \frac{y_{e}}{\sqrt{1-s}} \right) \rangle_s \]
\[ + \sum_{e \in E} \sum_{\ell \leq n} \gamma_{1,e} \sqrt{1-t_e} \mathbb{E}\langle L_{1,e,\ell} \left( \frac{g_{1,e}}{\sqrt{s}} - \frac{y_{1,e}}{\sqrt{1-s}} \right) \rangle_s \]
\[ + \sum_{e \in E} \sum_{\ell \leq n} \gamma_{2,e} \sqrt{1-t_e} \mathbb{E}\langle L_{2,e,\ell} \left( \frac{g_{2,e}}{\sqrt{s}} - \frac{y_{2,e}}{\sqrt{1-s}} \right) \rangle_s, \]

where \( L_{1,e,\ell} := 2^{-1}L(f_{e}(\rho^\ell) - f_{e}(\rho^{2\ell})) \) and \( L_{2,e,\ell} := 2^{-1}L(f_{e}(\tau^\ell) - f_{e}(\tau^{2\ell})) \). Note that we shall regard \( L_{1,e,\ell} \) and \( L_{2,e,\ell} \) as functions of \((\rho^\ell, \tau^\ell)_{\ell \leq 2n}\) with \(|L_{1,e,\ell}| \leq 1\) and \(|L_{2,e,\ell}| \leq 1\).

Following the same derivation as (10), each term in the first two summations can be controlled by

\[ \left| \mathbb{E}\langle L_{1,e} \left( \frac{g_{e}}{\sqrt{s}} - \frac{y_{e}}{\sqrt{1-s}} \right) \rangle_s \right| \leq t_{e}^{k/2}(\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{k+1} \leq (\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{k+1}, \]
\[ \left| \mathbb{E}\langle L_{2,e} \left( \frac{g_{e}}{\sqrt{s}} - \frac{y_{e}}{\sqrt{1-s}} \right) \rangle_s \right| \leq t_{e}^{k/2}(\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{k+1} \leq (\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{k+1}, \]

while each term in the last two summations can be controlled through

\[ \left| \mathbb{E}\langle L_{1,e} \left( \frac{g_{1,e}}{\sqrt{s}} - \frac{y_{1,e}}{\sqrt{1-s}} \right) \rangle_s \right| \leq (1 - t_{e})^{k/2}(\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{1,k+1} \leq (\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{1,k+1}, \]
\[ \left| \mathbb{E}\langle L_{2,e} \left( \frac{g_{2,e}}{\sqrt{s}} - \frac{y_{2,e}}{\sqrt{1-s}} \right) \rangle_s \right| \leq (1 - t_{e})^{k/2}(\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{2,k+1} \leq (\gamma_{1,e} + \gamma_{2,e})^{k}C_{k,2n}B_{2,k+1}. \]
Therefore, our proof is finished since

\[ |\phi'(s)| \leq nC_{k,2n} \left( B_{k+1} \sum_{e \in E} (\gamma_{1,e} + \gamma_{2,e})^{k+1} + B_{1,k+1} \sum_{e \in E} \gamma_{1,e}^{k+1} + B_{2,k+1} \sum_{e \in E} \gamma_{2,e}^{k+1} \right) \]

\[ \leq 2nC_{k,2n}(B_{k+1} + B_{1,k+1} + B_{2,k+1}) \sum_{e \in E} (\gamma_{1,e} + \gamma_{2,e})^{k+1}. \]

\[ \square \]

**Example 7** (Universality of chaos phenomena in the mixed \( p \)-spin model). Recall the two mixed \( p \)-spin Hamiltonians from Example 4. As in Example 6, we define \( E = \cup_{p \geq 2} E_p \) with \( E_p = \{1, \ldots, N\}^p \). For any \( e = (i_1, \ldots, i_p) \in E \), set \( t_e = t_p \), \( f(e) = \sigma_{i_1, \ldots, i_p} \) and

\[ \gamma_{1,e} = \beta_{1,p} N^{-(p-1)/2}, \quad \gamma_{2,e} = \beta_{2,p} N^{-(p-1)/2}, \]

\[ y_e = y_{i_1, \ldots, i_p}, \quad y_{1,e} = y_{1,i_1, \ldots, i_p}, \quad y_{2,e} = y_{2,i_1, \ldots, i_p}. \]

Then the two mixed \( p \)-spin Hamiltonians \( X_1, X_2 \) in Example 4 can be written as

\[ X_1(\rho) = H_{1,y}(\rho) + h_1 \sum_{i=1}^N \rho_i, \]

\[ X_2(\rho) = H_{2,y}(\tau) + h_2 \sum_{i=1}^N \tau_i, \]

where \( H_{y,1}, H_{2,y} \) are defined in (14). From Theorem 5, we obtain that

\[ |\mathbb{E}(L|_y - \mathbb{E}(L|_y)| \leq D_{k,n}' \sum_{p \geq 2} \frac{(\beta_{1,p} + \beta_{2,p})^{k+1}}{N^{(k+1)(p-1)/2-p}}. \] (15)

The right-side of (15) goes to 0 as \( N \) goes to infinity as long as \( k \geq 4 \) or \( k \geq 2 \) and \( \beta_{1,2} = \beta_{2,2} = \beta_{1,3} = \beta_{2,3} = 0 \). Once (15) goes to zero, we can approximate any coupled Gibbs measure with disorder \( y \) by a Gaussian disorder, thus establishing universality for coupled statistics.

Using (15), we now prove temperature chaos for disorders with four moments matching a standard Gaussian. Temperature chaos is known under the general assumptions on the temperature parameters (see [9, Page 4]). Here, for simplicity, we consider the following important example. Let \( p_0 \in \mathbb{N} \). Consider two mixed \( p \) spins with all \( \beta_{1,p}, \beta_{2,p} \neq 0 \), \( \beta_{1,p} = \beta_{2,p} \), for all \( p \neq p_0 \geq 2 \) and \( \beta_{1,p_0} \neq \beta_{2,p_0} \). Let \( \langle \cdot \rangle \) denote the coupled Gibbs measure.

**Theorem 6** (Universality of temperature chaos). If the disorders \((y_{1,i_1, \ldots, i_p})\) and \((y_{2,i_1, \ldots, i_p})\) are i.i.d. with four moments matching a standard Gaussian then

\[ \lim_{N \to \infty} \mathbb{E} \langle I(|R(\sigma, \tau)| > \varepsilon) \rangle = 0, \quad \forall \varepsilon > 0. \] (16)

**Proof.** The above result holds for Gaussian environments by [9, Theorem 1]. Since \( I(R(\sigma, \tau) > \varepsilon) \) is a bounded measurable function for any \( \varepsilon > 0 \), the theorem follows by (15).

\[ \square \]
Appendix

Proof of Lemma 1. Using Taylor’s theorem for $F$ and $F'$,

\[
yF(y) = \sum_{n=0}^{k-1} \frac{F^{(n)}(0)}{n!} y^{n+1} + \frac{F^{(k)}(a(y))}{k!} y^k,
\]

\[
F'(y) = \sum_{n=0}^{k-2} \frac{F^{(n+1)}(0)}{n!} y^n + \frac{F^{(k)}(b(y))}{(k-1)!} y^{k-1} = \sum_{n=1}^{k-1} \frac{F^{(n)}(0)}{(n-1)!} y^{n-1} + \frac{F^{(k)}(b(y))}{(k-1)!} y^{k-1}
\]

for two functions $a(y), b(y)$. Since their difference is equal to

\[
yF(y) - F'(y) = F'(0)y + \sum_{n=1}^{k-1} \frac{F^{(n)}(0)}{n!} \left( \frac{y^{n+1}}{n!} - \frac{y^{n-1}}{(n-1)!} \right) + \frac{F^{(k)}(a(y))}{k!} y^k - \frac{F^{(k)}(b(y))}{(k-1)!} y^{k-1}.
\]

Taking expectation and noting that $E|y| = 0$ and $(n-1)!E|y|^{n+1} = n!E|y|^{n-1}$, we obtain that

\[
|EyF(y) - EF'(y)| \leq \frac{\|F^{(k)}\|_{\infty}}{k!} E|y|^{k+1} + \frac{\|F^{(k)}\|}{(k-1)!} E|y|^{k-1}
\]

\[
\leq \frac{(k+1)}{k!} \|F^{(k)}\|_{\infty} E|y|^{k+1},
\]

where the second term on the right-hand side of the first inequality used Jensen’s inequality, $E|y|^{k-1} = E|y|^2 E|y|^{k-1} \leq E|y|^{k+1}$. 

\[\square\]

References


