VARIATIONAL REPRESENTATIONS FOR THE PARISI FUNCTIONAL AND THE TWO-DIMENSIONAL GUERRA–TALAGRAND BOUND

BY WEI-KUO CHEN

University of Minnesota

The validity of the Parisi formula in the Sherrington–Kirkpatrick model (SK) was initially proved by Talagrand [Ann. of Math. (2) 163 (2006) 221–263]. The central argument relied on a dedicated study of the coupled free energy via the two-dimensional Guerra–Talagrand (GT) replica symmetry breaking bound. It is believed that this bound and its higher dimensional generalization are highly related to the conjectures of temperature chaos and ultrametricity in the SK model, but a complete investigation remains elusive.

Motivated by Bovier–Klimovsky [Electron. J. Probab. 14 (2009) 161–241] and Auffinger–Chen [Comm. Math. Phys. 335 (2015) 1429–1444] the aim of this paper is to present a novel approach to analyzing the Parisi functional and the two-dimensional GT bound in the mixed $p$-spin models in terms of optimal stochastic control problems. We compute the directional derivative of the Parisi functional and derive equivalent criteria for the Parisi measure. We demonstrate how our approach provides a simple and efficient control for the GT bound that yields several new results on Talagrand’s positivity of the overlap and disorder chaos in Chatterjee [Disorder chaos and multiple valleys in spin glasses. Preprint] and Chen [Ann. Probab. 41 (2013) 3345–3391]. In particular, we provide some examples of the models containing odd $p$-spin interactions.

1. Introduction. In 1979, Parisi [16] suggested an ingenious variational formula for the limiting free energy in the Sherrington–Kirkpatrick (SK) model. Its validity was rigorously established by Talagrand [18] following the discovery of Guerra’s beautiful replica symmetry breaking scheme [9]. Parisi’s formula was later shown to be valid in the mixed $p$-spin models by Panchenko [15]. For $N \geq 1$, the Hamiltonian of the mixed $p$-spin model is defined as

$$H_N(\sigma) = X_N(\sigma) + h \sum_{i=1}^{N} \sigma_i$$

Received January 2015; revised November 2015.

1Supported in part by AMS-Simons travel grant, NSF Grant DMS-16-42207 and Hong Kong research grant council GRF 14302515.

MSC2010 subject classifications. 60K35, 82B44.

Key words and phrases. Chaos in disorder, Parisi formula, replica symmetry breaking, Sherrington–Kirkpatrick model.

3929
for $\sigma = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N := \{-1, +1\}^N$, where $X_N$ is the linear combination of the pure $p$-spin Hamiltonians,

$$X_N(\sigma) = \beta \sum_{p \geq 2} \frac{\gamma_p}{N(p-1)/2} \sum_{1 \leq i_1, \ldots, i_p \leq N} g_{i_1, \ldots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}$$

for i.i.d. standard Gaussian random variables, $g_{i_1, \ldots, i_p}$, for all $1 \leq i_1, \ldots, i_p \leq N$ and $p$. In physics, the $g_{i_1, \ldots, i_p}$’s are called the disorder, $h \in \mathbb{R}$ is the strength of the external field and $\beta > 0$ is called the (inverse) temperature. Here, we assume that the nonnegative sequence $(\gamma_p)_{p \geq 2}$ decays fast enough, for example, $\sum_{p \geq 2} 2^p \gamma_p^2 < \infty$, such that the covariance of $X_N$ can be computed as

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

for any two spin configurations $\sigma^1 = (\sigma_1^1, \ldots, \sigma_N^1)$ and $\sigma^2 = (\sigma_1^2, \ldots, \sigma_N^2)$ from $\Sigma_N$, where

$$R_{1,2} := \frac{1}{N} \sum_{i=1}^N \sigma_i^1 \sigma_i^2$$

is called the overlap between $\sigma^1$ and $\sigma^2$ and

$$\xi(s) := \sum_{p \geq 2} \beta_p^2 s^p, \quad \forall s \in [-1, 1]$$

for $\beta_p := \beta \gamma_p$ for all $p$. An important example of $\xi$ is the mixed even $p$-spin model, that is, $\gamma_p = 0$ for all odd $p$. In particular, the SK model corresponds to $\xi(s) = \beta^2 s^2 / 2$. Denote the Gibbs measure associated to $H_N$ by

$$G_N(\sigma) = \frac{\exp H_N(\sigma)}{Z_N},$$

where the normalizing factor $Z_N = \sum_{\sigma \in \Sigma_N} \exp H_N(\sigma)$ is called the partition function.

The Parisi formula is described as follows. Let $\mathcal{M}$ be the space of all probability measures on $[0, 1]$ and $\mathcal{M}_d$ be the collection of all atomic measures in $\mathcal{M}$. Denote by $\alpha_\mu$ the distribution function of $\mu \in \mathcal{M}$. We endow the space $\mathcal{M}$ with the metric:

$$d(\mu, \mu') = \int_0^1 |\alpha_\mu(s) - \alpha_{\mu'}(s)| \, ds.$$  

For any $\mu \in \mathcal{M}$, let $\Phi_\mu$ be the solution to the Parisi PDE on $[0, 1] \times \mathbb{R}$,

$$\partial_s \Phi_\mu(s, x) = -\frac{\xi''(s)}{2} \left( \partial_{xx} \Phi_\mu(s, x) + \alpha_\mu(s) \left( \partial_x \Phi_\mu(s, x) \right)^2 \right),$$

$$\Phi_\mu(1, x) = \log \cosh x.$$  

Here, for any $\mu \in \mathcal{M}_d$, this PDE can be explicitly solved by performing the Hopf–Cole transformation. For an arbitrary probability measure $\mu \in \mathcal{M}$, the solution $\Phi_\mu$
is understood in the weak sense (see Jagannath and Tobasco \[11\]). Define the Parisi functional $\mathcal{P}$ on $\mathcal{M}$ by

$$
\mathcal{P}(\mu) = \log 2 + \Phi_\mu(0, h) - \frac{1}{2} \int_0^1 \alpha_\mu(s)s\xi''(s)\,ds.
$$

Note that this functional is Lipschitz continuous (see Guerra \[9\]). The famous Parisi formula says that

$$
\lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log Z_N = \min_{\mu \in \mathcal{M}} \mathcal{P}(\mu).
$$

Here, the quantity inside the limit of the left-hand side is called the free energy of the model. Recently, Auffinger and Chen \[2\] showed that the Parisi functional is strictly convex, which implies the uniqueness of the minimizer. We will call this minimizer the Parisi measure and denote it by $\mu_P$. In order to classify the structure of $\mu_P$, we say that the Parisi measure is replica symmetric (RS) if it is a Dirac measure, is $k$ replica symmetry breaking ($k$-RSB) if it is atomic with exactly $k + 1$ jumps and is full replica symmetry breaking (FRSB) otherwise. In addition, for given sequence $(\gamma_p)_{p \geq 2}$ and fixed external field $h$, we define the high temperature regime as the collection of all $\beta > 0$ such that the corresponding Parisi measures are RS and the low temperature regime is defined as its complement. An important quantity associated to the mixed $p$-spin model is the overlap $R_{1,2}$ between two independently sampled spin configurations $\sigma_1$ and $\sigma_2$ from the Gibbs measure $G_N$. At very high temperature, that is, when $\beta$ is exceedingly small, this overlap is concentrated around a constant (see Talagrand \[20\], Chapter 13, for the SK model and Jagannath and Tobasco \[10\] for the mixed $p$-spin model), whereas in the low temperature regime, it is typically supported by a set containing more than one point (see Panchenko \[14\]).

Arguably, in the past decade, the most important development in the study of mean-field spin glasses is Guerra’s replica symmetry breaking bound \[9\] for the free energy in the mixed even $p$-spin model. Its statement reads that any $N \geq 1$ and $\mu \in \mathcal{M}$,

$$
\frac{1}{N} \mathbb{E} \log Z_N \leq \mathcal{P}(\mu).
$$

Based on Guerra’s interpolation scheme \[9\], the first proof of Parisi’s formula was obtained in the seminal work of Talagrand \[18\]. The central ingredient was played by a two-dimensional extension of Guerra’s inequality (7) for the coupled free energy with constrained overlaps, which was used to control the error estimate between the two sides of (7) when $\mu$ is very close to the Parisi measure. The full generalization of Guerra’s inequality (7), called the Guerra–Talagrand (GT) bound throughout this paper, was later presented in \[20\], Section 15.7. The two-dimensional GT bound, in particular, has two important consequences regarding the behavior of the overlap under the Gibbs measure. The first is known as the
positivity of the overlap established by Talagrand [20], Section 14.12, in the mixed even $p$-spin model, which says that if the external field is present, $h \neq 0$, then the overlap defined above is essentially bounded from below by some positive constant. Note that this behavior is very different from the one when the external field is absent, $h = 0$, in which case the overlap $R_{1,2}$ is symmetric with respect to the origin.

Another consequence is concerned with the phenomenon of chaos in disorder. It arose from the observation that in some spin glass models, a small perturbation to the disorder will result in a dramatic change to the overall energy landscape (see Rizzo [17] for a recent survey in physics). In the mixed $p$-spin model, one typical way to measure such instability is to consider two Hamiltonians:

$$H_N^1(\sigma^1) = X_N^1(\sigma^1) + h \sum_{1 \leq i \leq N} \sigma_i^1$$
$$H_N^2(\sigma^2) = X_N^2(\sigma^2) + h \sum_{1 \leq i \leq N} \sigma_i^2,$$

where $X_N^1$ and $X_N^2$ are jointly Gaussian with mean zero and covariance structure,

$$\mathbb{E} X_N^1(\sigma^1) X_N^1(\sigma^2) = \xi (R_{1,2}) = \mathbb{E} X_N^2(\sigma^1) X_N^2(\sigma^2),$$

$$\mathbb{E} X_N^1(\sigma^1) X_N^2(\sigma^2) = t \xi (R_{1,2})$$

(8)

for some $t \in [0, 1]$. Let $\sigma^1$ and $\sigma^2$ be independent samplings from $G_N^1$ and $G_N^2$, respectively, and let $R_{1,2}$ be their overlap, which now also depends on $t$. The case $t = 1$ means that the two systems are the same, $H_N^1 = H_N^2 = H_N$, and the overlap has the behavior we described above. From physics literature (e.g., Bray–Morre [4], Fisher–Huse [8], Krzakała–Bouchaud [13]), chaos in disorder is defined by the phenomenon that $R_{1,2}$ is concentrated around a nonrandom number independent of $N$ if the two systems are decoupled, that is, $t \in (0, 1)$. The key point here is that such behavior is predicted to be true at any temperature. The first rigorous result in this direction was justified in the mixed even $p$-spin models without external field in the work of Chatterjee [5] and the situation in the presence of the external field was carried out in Chen [6].

As the above discussion indicates, the Parisi functional and the GT bound play fundamental roles in the study of the mixed $p$-spin model. Several challenging conjectures, such as the strong ultrametricity and temperature chaos (see Talagrand [20], Section 15.7), rely heavily on the subtle control of these two objects and their higher dimensional generalization. To this regard, the aim of this paper is to present a novel approach to analyzing the Parisi functional as well as the two-dimensional GT bound by means of optimal stochastic control theory. Ultimately, we hope that this new method will shed some light on how to tackle the remaining open problems. Our idea is motivated by the observation that the Parisi PDE solution $\Phi_{\mu}$ admits a variational representation (see Theorem 1 below) in terms of an optimal stochastic control problem that corresponds to the Hamilton–Jacobi–Bellman equation induced by a linear diffusion control problem. This was formerly used in Bovier and Klimovsky [3] to study the strict convexity of the Parisi functional for
some cases of the SK model with multidimensional spins. Later it was understood that this approach allows one to derive the strict convexity of the Parisi functional in the mixed \( p \)-spin models by Auffinger and Chen [2].

This article consists of four major results. The first result involves an analytic study of the Parisi formula, where we compute the directional derivative of the Parisi functional and give equivalent criteria for the Parisi measure. As an application, we generalize a theorem of Toninelli [21], which states that the Parisi measure in the SK model is not a Dirac measure when the temperature and external field stay above the Almeida–Thouless transition line [see (15) below]. In addition, we extend Talagrand’s characterization [20], Theorem 13.4.1, of the high temperature regime for the SK model to the temperature regime of \( k \)-RSB Parisi measures for any mixed \( p \)-spin models. Second of all, we establish a variational representation for the two-dimensional Parisi PDE solution in terms of an optimal stochastic control problem and use this to give a new formulation of the original GT bound. Based on this new form, our last two results are devoted to demonstrating a self-contained proof of the positivity of the overlap and disorder chaos in the mixed \( p \)-spin model. We recover the aforementioned results and, furthermore, extend them to many new examples of the model allowing odd \( p \)-spin interactions.

Along the way, we also obtain a nonnegativity principle of the overlap in the mixed \( p \)-spin model, which says that in the absence of the external field, the overlap is basically nonnegative if one adds certain odd \( p \)-spin interactions to the Hamiltonian. In Section 5 below, our approach significantly simplifies and avoids several technicalities in the control of the two-dimensional GT bound compared to the arguments in Talagrand [20], Section 14.12, and Chen [6]. For instance, the error estimate of this bound was previously obtained through a quite involved iteration for certain functions of Gaussian random variables. With the new approach, it now becomes quantitatively simpler in the critical case (see Proposition 5 below).

This paper is organized as follows. In Section 2, we state the four main results described above and their proofs are presented in Sections 3 to 5. The analytic properties of the Parisi functional are investigated in Section 3 and the variational representation for the two-dimensional GT bound is derived in Section 4. Finally, we present the proof for the results on the positivity of the overlap and disorder chaos in Section 5.

2. Main results.

2.1. Some properties of Parisi’s functional and measure. First, we recall the variational representation for the Parisi PDE from Auffinger and Chen [2]. Let \((\mathbb{P}, \mathcal{F}, \mathcal{F}_r)_{0 \leq r \leq 1}\) be a filtered probability space satisfying the usual condition, that is, it is complete and the filtration is right continuous. Let \(B = \{B(r), \mathcal{F}_r, 0 \leq r \leq 1\}\) be a standard Brownian motion. For \(0 \leq s < t \leq 1\), let \(D[s, t]\) be the collection of all progressively measurable processes \(u\) with respect to \((\mathcal{F}_r)_{0 \leq r \leq t}\) satisfying \(\sup_{s \leq r \leq t} |u(r)| \leq 1\). We equip the space \(D[s, t]\) with the norm \(\|u\| = \int_s^t \|u(r)\|^2 dr\).
Let $\xi$ and $h$ be fixed. Set $\zeta = \xi''$. For $\mu \in \mathcal{M}$, we define a functional on $D[s, t] \times \mathbb{R}$ by

$$F_{\mu}^{s,t}(u, x) = \mathbb{E}[C_{\mu}^{s,t}(u, x) - L_{\mu}^{s,t}(u)] \tag{9}$$

where, recalling that $\alpha_{\mu}$ is the distribution function of $\mu$,

$$C_{\mu}^{s,t}(u, x) := \Phi_{\mu}(t, h + \int_{s}^{t} \alpha_{\mu}(w) \zeta(w) w u(w) dw + \int_{s}^{t} \zeta(w) 1/2 dB(w)),$$

$$L_{\mu}^{s,t}(u) := \frac{1}{2} \int_{s}^{t} \alpha_{\mu}(w) \zeta(w) u(w)^2 dw.$$

The Parisi PDE solution can be expressed as the following.

**Theorem 1** ([2], Theorem 3 and Proposition 3). For any $\mu \in \mathcal{M}$,

$$\Phi_{\mu}(s, x) = \max_{u \in D[s, t]} F_{\mu}^{s,t}(u, x). \tag{10}$$

The maximum is attained by $u_{\mu}(r) = \partial_x \Phi_{\mu}(r, X(r))$, where $(X(r))_{s \leq r \leq t}$ satisfies

$$X(r) = x + \int_{s}^{r} \alpha_{\mu}(w) \zeta(w) \partial_x \Phi_{\mu}(w, X(w)) dw + \int_{s}^{r} \zeta(w) 1/2 dB(w). \tag{11}$$

In addition, the maximizer is unique if $\alpha_{\mu} > 0$ on $[s, t]$ and $\int_{s}^{t} \alpha_{\mu}(r) dr < 1$.

Our first main results below are the computation of the directional derivative of the Parisi functional and the equivalent criteria for the Parisi measure.

**Theorem 2.** Let $\mu_0 \in \mathcal{M}$. Define $\mu_\theta = (1 - \theta) \mu_0 + \theta \mu$ for each $\mu \in \mathcal{M}$ and $\theta \in [0, 1]$. We have

$$\frac{d}{d\theta} \mathcal{P}(\mu_\theta) \bigg|_{\theta = 0} = \frac{1}{2} \int_{0}^{1} \zeta(r) (\alpha_{\mu}(r) - \alpha_{\mu_0}(r)) (\mathbb{E} u_{\mu_0}(r)^2 - r) dr \tag{12}$$

for all $\mu \in \mathcal{M}$, where $\frac{d}{d\theta} \mathcal{P}(\mu_\theta) \big|_{\theta = 0}$ is understood as the right derivative at 0 and $u_{\mu_0}$ is the maximizer of (10) using $\mu_0$ and $(s, t) = (0, 1)$. In addition, the following statements are equivalent:

(i) $\mu_0$ is the Parisi measure.

(ii) $\frac{d}{d\theta} \mathcal{P}(\mu_\theta) \big|_{\theta = 0} \geq 0$ for all $\mu \in \mathcal{M}$.

(iii) $\frac{d}{d\theta} \mathcal{P}(\mu_\theta) \big|_{\theta = 0} \geq 0$ for all Dirac measures $\mu = \delta_q$ with $q \in [0, 1]$. 

The equivalence of (i) and (ii) is mainly due to the strict convexity of the Parisi functional. The criterion (iii) essentially says that if one could not lower the Parisi functional by adding one more jump to \( \mu_0 \), then \( \mu_0 \) must be the Parisi measure. There are two immediate consequences that can be drawn from this theorem. For convenience, we set \( \mathcal{M}_d^k \) for \( k \geq 0 \) to be the collection of all members of \( \mathcal{M}_d \) that have no more than \( k+1 \) atoms, that is, \( \mu = \sum_{i=1}^{k+1} a_i \delta_{q_i} \) for some \( 0 \leq q_1 \leq \cdots \leq q_{k+1} \leq 1 \) and \( 0 \leq a_1, \ldots, a_{k+1} \leq 1 \) with \( \sum_{i=1}^{k+1} a_i = 1 \). In particular, \( \mathcal{M}_d^0 \) denotes the space of all Dirac measures on \([0, 1]\). In the first consequence, we extract some information about the support of the Parisi measure.

**Proposition 1.** Let \( S \) be the support of \( \mu_P \). For all \( q \in S \),

\[
\mathbb{E} \partial_x \Phi_{\mu_P}(q, X(q))^2 = q,
\]

\[
\zeta(q) \mathbb{E} \partial_{xx} \Phi_{\mu_P}(q, X(q))^2 \leq 1,
\]

where \((X(s))_{0 \leq s \leq 1}\) satisfies the following stochastic differential equation:

\[
X(s) = h + \int_0^s \alpha_{\mu_P}(r) \zeta(r) \partial_x \Phi_{\mu_P}(w, X(w)) \, dw + \int_0^s \zeta(w)^{1/2} \, dB(w),
\]

\( \forall s \in [0, 1] \).

**Remark 1.** Suppose that \( \mu_P \) is a Dirac measure at some \( q \in [0, 1] \). A direct computation gives

\[
\Phi_{\mu_P}(s, x) = \begin{cases} 
\frac{1}{2}(\xi'(1) - \xi'(s)) + \mathbb{E} \log \cosh(x + z(\xi'(q) - \xi'(s))^{1/2}), & \text{if } (s, x) \in [0, q) \times \mathbb{R}, \\
\frac{1}{2}(\xi'(1) - \xi'(q)) + \log \cosh x, & \text{if } (s, x) \in [q, 1] \times \mathbb{R},
\end{cases}
\]

for \( z \) a standard Gaussian random variable. Since \( \alpha_{\mu_P} = 0 \) on \([0, q)\), Theorem 1 reads

\[
\mathbb{E} \tanh^2(z\xi'(q)^{1/2} + h) = q,
\]

\[
\zeta(q) \mathbb{E} \frac{1}{\cosh^4(z\xi'(q)^{1/2} + h)} \leq 1.
\]

Note that if \( q \in [0, 1] \) minimizes the Parisi functional over all choices in \( \mathcal{M}_d^0 \), then one can get the first equation (by a direct differentiation; see, for example, [19], Chapter 1), but if the temperature and external field are above the Almeida–Thouless line, that is, (15) is violated, then the Parisi measure can not be RS. This generalizes Toninelli’s theorem [21], where he established the same statement for the SK model \( \xi(s) = \beta^2 s^2/2 \).
Remark 2. Consider the SK model without external field, that is, $\xi(s) = \beta^2 s^2 / 2$ and $h = 0$. We now argue that the high temperature regime, defined as the collection of all $\beta$ such that $\mu_P$ is a Dirac measure, is described by $\beta \leq 1$. To see this, note that since $h = 0$, 0 is always in the support of the Parisi measure by [1], Theorem 1. Thus, it suffices to show that $\mu_P = \delta_0$ if and only if $\beta \leq 1$. If $\mu_P = \delta_0$ and $\beta > 1$, we will obtain a contradiction as (15) is violated. Conversely, suppose $\beta \leq 1$. A use of Itô’s formula and (6) gives

$$u_{\delta_0}(r) = \beta \int_0^r \partial_{xX} \Phi_{\delta_0}(w, X(w)) dB(w) + u_{\delta_0}(0)$$

and hence,

$$E u_{\delta_0}(r)^2 = \beta^2 \int_0^r \frac{1}{\cosh^2 X(w)} dw \leq \beta^2 \int_0^r 1 dw \leq r.$$

Therefore, for all $\mu \in \mathcal{M}$,

$$\frac{d}{d\theta} \mathcal{P}(\mu_\theta) \bigg|_{\theta = 0} = \frac{\beta^2}{2} \int_0^1 (\alpha(r) - 1)(E u_{\delta_0}(r)^2 - r) dr \geq 0$$

and Theorem 2 implies that $\delta_0$ is the Parisi measure.

The second consequence of Theorem 2 is a generalization of Talagrand’s characterization [20], Theorem 13.4.1, of the high temperature regime for the SK model, where he showed that this regime is indeed equal to the set of all $\beta$ such that $\inf_{\mu \in \mathcal{M}} \mathcal{P}(\mu) = \mathcal{P}(\mu_0)$ for some $\mu_0 \in \mathcal{M}_d^0$. For any such $\beta$, he proved that $\mu_0$ will automatically be the Parisi measure. With the help of Theorem 2(iii), this result can be generalized to any $k$-RSB Parisi measures.

Proposition 2. Consider arbitrary $\xi$ and $h$. Let $k \geq 0$ and $\mu_0$ be an optimizer of $\mathcal{P}$ over $\mathcal{M}^k_d$. If

$$\inf_{\mu \in \mathcal{M}_d^{k+1}} \mathcal{P}(\mu) = \mathcal{P}(\mu_0),$$

then $\mu_0$ is the Parisi measure.

In other words, for fixed sequence $(\gamma_p)_{p \geq 2}$ and external field $h$, the temperature regime of $k$-RSB Parisi measures is described by the collection of all $\beta > 0$ such that the corresponding Parisi functionals satisfy (16) for some optimizer $\mu_0$ of $\mathcal{P}$ restricted to $\mathcal{M}^k_d$.

It is generally very difficult to compute the Parisi measure as one needs to minimize $\mathcal{P}$ over all probability measures on $[0, 1]$. In principle, Proposition 2 suggests
a heuristic way to simulate $k$-RSB Parisi measures. The procedure is based on the observation that if we restrict $\mathcal{P}$ to $\mathcal{M}_d^k$, then it is a differentiable function that depends only on $2(k+1)$ variables on a compact set:

$$\left\{ (q_1, \ldots, q_{k+1}, a_1, \ldots, a_{k+1}): 0 \leq q_1 \leq \cdots \leq q_{k+1} \leq 1, 0 \leq a_1, \ldots, a_{k+1} \leq 1, \sum_{i=1}^{k+1} a_i = 1 \right\},$$

on which one can compute the derivative of $\mathcal{P}$ and numerically simulate the minimizer of $\mathcal{P}$ over $\mathcal{M}_d^k$. Starting from the case $k = 0$, if (16) is satisfied, then one can stop and obtain the RS Parisi measure; otherwise one must proceed to the case $k = 1$ and continue this process. If eventually there is a smallest integer $k \geq 0$ such that (16) is obtained, then one gets a $k$-RSB Parisi measure.

### 2.2. A variational representation for the two-dimensional GT bound.

The two-dimensional GT bound in the setting of [20], Theorem 15.7, is formulated as follows. Let $h_1, h_2 \in \mathbb{R}$ and $X_N^{1,1}, X_N^{2,2}$ be jointly Gaussian processes indexed by $\Sigma_N$ with mean zero and covariance,

$$\mathbb{E} X_N^{\ell,1}(\sigma^1) X_N^{\ell',1}(\sigma^2) = N \xi_{\ell,1,\ell',1}(R_{1,2})$$

for $1 \leq \ell, \ell' \leq 2$ and $\sigma^1, \sigma^2 \in \Sigma_N$, where $R_{1,2}$ is the overlap between $\sigma^1, \sigma^2$ defined through (2). Here, $\xi_{\ell,1,\ell',1}$’s are convex functions on $[-1,1]$ defined in terms of infinite series similar to the definition of $\xi$ in (3). Consider two mixed $p$-spin Hamiltonians:

$$(17) \quad H_N^{\ell}(\sigma^{\ell}) = X_N^{\ell}(\sigma^{\ell}) + h_\ell \sum_{1 \leq i \leq N} \sigma_i^{\ell}, \quad \ell = 1, 2. \quad$$

Denote by $S_N$ the collection of all possible values of $R_{1,2}$. Fix $q \in S_N$. Assume that $(y_p^{\ell})_{0 \leq p \leq k}$ for $1 \leq \ell \leq 2$ are jointly centered Gaussian random variables such that for certain real sequences $(\rho_p^{\ell,\ell'})_{0 \leq p \leq k+1}$ for $1 \leq \ell, \ell' \leq 2$ with

$$\rho_0^{1,1} = \rho_0^{2,2} = \rho_0^{1,2} = \rho_0^{2,1} = 0,$$

$$\rho_{k+1}^{1,1} = \rho_{k+1}^{2,2} = 1,$$

$$\rho_{k+1}^{1,2} = \rho_{k+1}^{2,1} = q,$$

we have

$$\mathbb{E} y_p^{\ell} y_p^{\ell'} = \xi_{\ell,1,\ell',1}(\rho_p^{\ell,\ell'}) - \xi_{\ell,1,\ell',1}(\rho_p^{\ell,\ell'}).$$
THEOREM 3 (Guerra–Talagrand). Let \((m_p)_{0 \leq p \leq k}\) be a sequence with \(m_0 = 0 < m_1 < \cdots < m_{k-1} < m_k = 1\). Under the assumptions stated above,

\[
F_N(q) := \frac{1}{N} \mathbb{E} \log \sum_{\rho_{1,2} = q} \exp\left( H_N^1(\sigma^1) + H_N^2(\sigma^2) \right)
\]

(19)

\[
\leq 2 \log 2 + Y_0 - \lambda q - \frac{1}{2} \sum_{1 \leq \ell, \ell' \leq 2} \sum_{p=0}^k m_p (\theta_{\ell, \ell'}(\rho_{p+1}) - \theta_{\ell, \ell'}(\rho_p)),
\]

where \(\theta_{\ell, \ell'}(s) := s \xi_{\ell, \ell'}(s) - \xi_{\ell, \ell'}(s)\) and \(Y_0\) is defined as follows. Denote by \(\mathbb{E}_p\) the expectation with respect to \(y_p^{\ell, \ell'}\). Starting with

\[
Y_{k+1} = \log \left( \cosh \left( h_1 + \sum_{p=0}^k y_p^1 \right) \cosh \left( h_2 + \sum_{p=0}^k y_p^2 \right) \cos \lambda \right)
\]

\[
+ \sinh \left( h_1 + \sum_{p=0}^k y_p^1 \right) \sinh \left( h_2 + \sum_{p=0}^k y_p^2 \right) \sin \lambda,
\]

we define decreasingly \(Y_p = m_p^{-1} \mathbb{E}_p \exp m_p Y_{p+1}\) for \(1 \leq p \leq k\). Finally, set \(Y_0 = \mathbb{E}_0 Y_1\).

The inequality (19) is a two-dimensional extension of Guerra’s replica symmetry breaking bound (7). Its proof as well as the higher dimensional extension can be found in [20], Section 15.7. Recall \(q\) from the statement of Theorem 3. Let \(\iota = 1\) if \(q \geq 0\) and \(\iota = -1\) otherwise. For \(1 \leq \ell, \ell' \leq 2\), let \(\rho_{\ell, \ell'}\) be nondecreasing continuous functions on \([0, 1]\) with

\[
\rho_{1,1}(0) = \rho_{1,2}(0) = \rho_{2,1}(0) = \rho_{2,2}(0) = 0,
\]

\[
\rho_{1,1}(1) = \rho_{2,2}(1) = 1, \quad \rho_{1,2}(1) = \rho_{2,1}(1) = |q|.
\]

Assume that these functions are differentiable everywhere except at a finite number of points, at which the right derivatives exist. For any \(s \in [0, 1]\), we define

\[
T(s) = \begin{bmatrix}
\xi_{1,1}(s) & \xi_{1,2}(s) \\
\xi_{2,1}(s) & \xi_{2,2}(s)
\end{bmatrix} := \begin{bmatrix}
\frac{d}{ds} \xi_{1,1}(s) & \frac{d}{ds} \xi_{1,2}(s) \\
\frac{d}{ds} \xi_{2,1}(s) & \frac{d}{ds} \xi_{2,2}(s)
\end{bmatrix}.
\]

In the right-hand side of (21), the derivatives are understood as the ones from the right if one of \(\rho_{\ell, \ell'}\)’s is not differentiable. We suppose that \(T(s)\) is positive semidefinite and its operator norm \(\|T(s)\|\) is uniformly bounded from above by some constant \(K > 0\). For \(\mu \in \mathcal{M}_d\), we consider the classical solution \(\Psi_\mu\) to the two-dimensional Parisi PDE,

\[
\partial_s \Psi_\mu = -\frac{1}{2} \left( [T, \nabla^2 \Psi_\mu] + \alpha_\mu (T \nabla \Psi_\mu, \nabla \Psi_\mu) \right)
\]

(22)
for \((\lambda, s, x) \in \mathbb{R} \times [0, 1) \times \mathbb{R}^2\) with terminal condition

\[
\Psi_{\mu}(\lambda, 1, x) = \log(\cosh x_1 \cosh x_2 \cosh \lambda + \sinh x_1 \sinh x_2 \sinh \lambda).
\]  

(23)

The assumption \(\mu \in \mathcal{M}_d\) guarantees the existence of the solution by a standard application of Hopf–Cole transformation. One may refer to Lemma 3 below for the precise formula of the solution. Our first main result below says that the mapping \(\mu \in \mathcal{M}_d \mapsto \Psi_{\mu}\) is Lipschitz with respect to the metric \(d\) defined by (5).

**Theorem 4.** For any \(\mu, \mu' \in \mathcal{M}_d\),

\[
|\Psi_{\mu}(\lambda, s, x) - \Psi_{\mu'}(\lambda, s, x)| \leq 3Kd(\mu, \mu')
\]

for \((\lambda, s, x) \in \mathbb{R} \times [0, 1) \times \mathbb{R}^2\).

This Lipschitz property allows us to extend \(\Psi_{\mu}\) continuously to all \(\mu \in \mathcal{M}\) using sequences of atomic probability measures. Denote by \(\mathcal{B} = \{\mathcal{B}(r) = (B_1(r), B_2(r)), \mathcal{G}_r, 0 \leq r < \infty\}\) a two-dimensional Brownian motion, where \((\mathcal{G}_r)_{r \geq 0}\) satisfies the usual condition. For \(0 \leq s < t \leq 1\), denote by \(\mathcal{D}[s, t]\) the space of all two-dimensional progressively measurable processes \(v = (v_1, v_2)\) with respect to \((\mathcal{G}_r)_{s \leq r \leq t}\) satisfying \(\sup_{s \leq r \leq t} |v_1(r)| \leq 1\) and \(\sup_{s \leq r \leq t} |v_2(r)| \leq 1\). Endow the space \(\mathcal{D}[s, t]\) with the norm

\[
\|v\|_{s, t} = \left( \mathbb{E} \int_s^t (v_1(w)^2 + v_2(w)^2) \, dw \right)^{1/2}.
\]

Similar to the formulation of (9), we define a functional

\[
\mathcal{F}^{s,t}_{\mu}(\lambda, v, x) = \mathbb{E}[\mathcal{C}^{s,t}_{\mu}(\lambda, v, x) - \mathcal{L}^{s,t}_{\mu}(v)]
\]

for \((\lambda, v, x) \in \mathbb{R} \times \mathcal{D}[s, t] \times \mathbb{R}^2\), where

\[
\mathcal{C}^{s,t}_{\mu}(\lambda, v, x) := \Psi_{\mu}\left(\lambda, t, x + \int_s^t \alpha_{\mu}(w)T(w)v(w) \, dw + \int_s^t T(w)^{1/2} \, dB(w)\right),
\]

\[
\mathcal{L}^{s,t}_{\mu}(v) := \frac{1}{2} \int_s^t \alpha_{\mu}(w)[T(w)v(w), v(w)] \, dw.
\]

The following is an analogue of Theorem 1 for \(\Psi_{\mu}\).

**Theorem 5.** For any \(\mu \in \mathcal{M}\),

\[
\Psi_{\mu}(\lambda, s, x) = \max\{\mathcal{F}^{s,t}_{\mu}(\lambda, v, x) | v \in \mathcal{D}[s, t]\}.
\]

(24)

The maximum of (24) is attained by \(v_{\mu}(r) = \nabla \Psi_{\mu}(\lambda, r, X(r))\), where the two-dimensional stochastic process \((X(r))_{s \leq r \leq t}\) satisfies

\[
X(r) = x + \int_s^r \alpha_{\mu}(w)T(w)\nabla \Psi_{\mu}(\lambda, w, X(w)) \, dw + \int_s^r T(w)^{1/2} \, dB(w).
\]

(25)
Using the notation introduced above, we can now formulate the GT bound in terms of $\Psi_\mu$.

**Theorem 6 (Guerra–Talagrand).** Suppose that $T$ is positive semidefinite for all $s$. Then

$$F_N(q) \leq 2\log 2 + \mathbb{E} \Psi_\mu(\lambda, 0, h_1, h_2) - \lambda q$$

$$(26)$$

Typically, to use this bound, one needs to first find suitable parameters $\lambda$ and $\rho_{\ell, \ell'}$ depending on $q$ such that the right-hand side is less than or equal to $2P(\mu_P)$ for any $q \in [-1, 1]$. In Section 5, we shall see that this can be achieved in the case of $\xi_{1,1} = \xi_{2,2}$ and $h_1 = h_2$, but the general situation remains mysterious.

**2.3. Some properties of the overlap.** Recall the Hamiltonian $H_N$ and the Gibbs measure $G_N$ from (1) and (4). Let $\mu_P$ be the Parisi measure associated to $H_N$ and set $\eta = \min \operatorname{supp} \mu_P$. It is known (see [7]) that

$$\eta = 0 \quad \text{if } h = 0 \quad \text{and} \quad \eta > 0 \quad \text{if } h \neq 0.$$  

$$(27)$$

Recall that as we discussed in the Introduction, the overlap $R_{1,2}$ between two independently sampled spin configurations from $G_N$ is symmetric with respect to the origin if the mixed $p$-spin is even and the external field is absent. The positivity principle of the overlap says that this symmetry will be broken in such a way that the overlap is essentially bounded from below by $\eta$ if the external field is present.

**Theorem 7 (Positivity of the overlap).** Assume that $\xi$ is convex on $[-1, 1]$ and is not identically equal to zero. If $h \neq 0$, then under either of the following two assumptions:

(i) $\xi$ is even,

(ii) $\xi$ is not even and the function below is nondecreasing on $(0, 1]$,

$$\frac{\xi''(s)}{\xi''(s) + \xi''(-s)}.$$  

$$(28)$$

we have that for any $\varepsilon > 0$, there exists a constant $K_0 > 0$ such that

$$\mathbb{E}_{G_N \times G_N}(\sigma^{1, 2} : R_{1,2} \leq \eta - \varepsilon) \leq K_0 \exp\left(-\frac{N}{K_0}\right), \quad \forall N \geq 1.$$  

$$(29)$$

The inequality (29) means that if the external field is present, then the overlap essentially charges weight only in the interval $[\eta, 1] \subseteq (0, 1]$. Positivity of the overlap under the condition (i) was initially established by Talagrand [20], Section 14.10. Our main contribution here is the case (ii), where we allow odd $p$-spin interactions in the Hamiltonian. Below we describe a concrete example of the case (ii).
EXAMPLE 1. Consider $\xi(s) = \beta^2 (\gamma_{2p}^2 s^{2p} + \gamma_{2p+1}^2 s^{2p+1})$ on $[-1, 1]$ with $\gamma_{2p}$ and $\gamma_{2p+1}$ satisfying

$$c := \frac{(2p + 1)\gamma_{2p+1}^2}{(2p - 1)\gamma_{2p}^2} < 1.$$ 

It is easy to verify that this condition ensures the convexity of $\xi$ on $[-1, 1]$. Since

$$\frac{\xi''(s)}{\xi''(s) + \xi''(-s)} = \frac{1 + cs}{2}$$

is nondecreasing on $(0, 1]$, condition (ii) in Theorem 7 is satisfied, from which we obtain (29) for any $\beta > 0$.

Our next result shows that in the absence of the external field $h = 0$, the behavior of the overlap is also drastically influenced by the odd $p$-spin interactions in the Hamiltonian, where the overlap will stay nonnegative.

**THEOREM 8** (Nonnegativity of the overlap). Assume that $\xi$ is convex on $[-1, 1]$ and is not identically equal to zero. If $h = 0$ and the assumption (ii) in Theorem 7 hold, then for any $\epsilon > 0$, there exists a constant $K_0 > 0$ such that

$$\mathbb{E} G_N \times G_N ((\sigma^1, \sigma^2) : R_{1,2} \leq -\epsilon) \leq K_0 \exp \left( -\frac{N}{K_0} \right), \quad \forall N \geq 1. \quad (30)$$

2.4. Chaos in disorder. Recall the Hamiltonians $H_{1N}^1$ and $H_{2N}^2$ from (17). Assume that the Gaussian parts of the Hamiltonians, $X_{1N}^1$ and $X_{2N}^2$, have the covariance structure:

$$\xi_{1,1} = \xi_{2,2} = \xi, \quad \xi_{1,2} = \xi_{2,1} = \xi_0 \quad (31)$$

for some series $\xi_0$ defined in a way similar to $\xi$ and that the external fields satisfy

$$h_1 = h_2 = h. \quad (32)$$

In other words, the two systems have the same distribution and they are coupled through the function $\xi_0$. Denote by $G_N^1$ and $G_N^2$ the Gibbs measures associated to these Hamiltonians in the same fashion as (4). Note that the two systems share the same Parisi measure $\mu_P$ and $\eta := \min \text{supp} \mu_P$ have the property (27). Consider the overlap $R_{1,2}$ between the independently sampled $\sigma^1$ and $\sigma^2$ from $G_N^1$ and $G_N^2$, respectively. We say that there is chaos in disorder between $H_{1N}^1$ and $H_{2N}^2$ if the overlap is concentrated around a constant value. Our main result shows that this behavior holds as long as the two systems are decoupled, $\xi_0 \neq \xi$, for $\xi$ and $\xi_0$ satisfying some mild assumptions.
THEOREM 9 (Disorder chaos). Assume that (31) and (32) hold. If \( \xi \) and \( \xi_0 \) are convex on \([-1, 1]\) and are not identically equal to zero such that

\[
\frac{\xi''(s)}{\xi''(s) + \xi''_0(-s)} \quad \text{and} \quad \frac{\xi''(s)}{\xi''(s) + \xi''_0(s)}
\]

are both nondecreasing on \((0, 1]\) and

\[
\xi_0''(s) < \xi''(|s|)
\]

for \( s \in [-1, 1] \setminus \{0\} \), then there is a constant \( q^* \) such that for any \( \varepsilon > 0 \),

\[
\mathbb{E}G_1^N \times G_2^N((\sigma_1, \sigma_2) : |R_{1,2} - q^*| > \varepsilon) \leq K_0 \exp\left(-\frac{N}{K_0}\right)
\]

for all \( N \geq 1 \), where \( K_0 \) is a constant independent of \( N \). Here, \( q^* = 0 \) if \( h = 0 \) and \( q^* \in (0, \eta) \) if \( h \neq 0 \).

Form this theorem, the overlap is basically concentrated around a constant value \( q^* \) if the two systems are decoupled in an appropriate way (33) and (34). We emphasize that this behavior is completely different from the situation when \( \xi = \xi_0 \), in which case the two systems are indeed identical, \( H_1^N = H_2^N = H_N \), and the overlap typically has nontrivial limiting distribution in the low temperature regime. See, for instance, Examples 1 and 2 in [14]. The following two choices of \( \xi \) and \( \xi_0 \) summarize the previously known results and give new examples of chaos in disorder.

EXAMPLE 2 (mixed even \( p \)-spin models). Assume that the two systems are mixed even \( p \)-spin models and they are correlated through \( \xi_0 = t\xi \) for some \( t \in (0, 1) \). This choice of \((\xi, \xi_0)\) corresponds to (8) and was originally considered in Chatterjee [5], where he provided moments estimates to prove that the overlap is concentrated around 0 when there is no external field \( h = 0 \). Later Chen [6] established (35) in the presence of external field \( h \neq 0 \). One can easily check that \( \xi \) and \( \xi_0 \) are convex functions and (33) and (34) are satisfied. Thus, Theorem 9 proves disorder chaos irrespective of the presence or absence of the external field.

The main merit of Theorem 9 is that it also covers the mixed \( p \)-spin models containing odd \( p \)-spin interactions for properly chosen sequences \((\gamma_p)_{p \geq 2}\).

EXAMPLE 3. Recall \( \xi \) and \( c \) from Example 1. Let \( \xi_{1,1} = \xi_{2,2} = \xi \). For \( t \in [0, 1) \), set \( \xi_0(s) := \beta^2(\gamma_{2p}^2 s^{2p} + t\gamma_{2p+1}^2 s^{2p+1}) \) for \( s \in [-1, 1] \). Since \( c < 1 \) and \( t \in [0, 1) \), one can check that \( \xi_0 \) is convex on \([-1, 1]\). In addition, since

\[
\frac{\xi''(s)}{\xi''(s) + \xi_0''(s)} = \frac{1 + cs}{2(1 + \frac{1+ct}{2}cs)},
\]

\[
\frac{\xi''(s)}{\xi''(s) + \xi_0''(-s)} = \frac{1 + cs}{2(1 + \frac{1+t}{2}cs)},
\]
and $t \in [0, 1)$, a direct differentiation with respect to $s$ implies that they are both nondecreasing on $(0, 1]$. On the other hand,

$$\xi''_0(s) = 2p\beta^2 s^{2p-2}((2p-1)\gamma_{2p}^2 + t(2p+1)\gamma_{2p+1}^2 s)
< 2p\beta^2 |s|^{2p-2}((2p-1)\gamma_{2p}^2 + (2p+1)\gamma_{2p+1}^2 |s|)
= \xi''(|s|)$$

for all $s \in [-1, 1]\{0\}$. Therefore, the conclusion of Theorem 9 holds for all $\beta > 0$.

3. **Directional derivative of the Parisi functional.** In this section, we establish the main results stated in Section 2.1. We will use the variational representation formula (10) for $\Phi_\mu(0, x)$ with $(s, t) = (0, 1)$ throughout this section. Recall the associated maximizer $u_\mu$ from (11). We start by computing the directional derivative of the Parisi functional, which relies on two technical lemmas. The first is the combination of [1], Proposition 2, and [2], Lemma 2.

**Lemma 1.** For any $\mu \in \mathcal{M}$ and $s \in [0, 1]$, $\partial_x \Phi_\mu(s, \cdot)$ is odd, strictly increasing and uniformly bounded by $1$. In addition, the process $u_\mu$ satisfies

$$u_\mu(b) - u_\mu(a) = \int_a^b \zeta(w)^{1/2} \partial_x \Phi_\mu(w, X(w)) dB(w)$$

for all $0 \leq a \leq b \leq 1$.

The second lemma allows us to take derivatives for maximum functions under mild assumptions.

**Lemma 2.** Let $K$ be a metric space and $I$ be an interval with right open edge. Let $f$ be a real-valued function on $K \times I$ and $g(y) = \sup_{a \in K} f(a, y)$. Suppose that there exists a $K$-valued continuous function $a(y)$ on $I$ such that $g(y) = f(a(y), y)$ and $\partial_y f$ is continuous on $K \times I$, then $g$ is right-differentiable with derivative $\partial_y f(a(y), y)$ for all $y \in I$.

**Proof.** Let $y \in I$. Consider any $h > 0$ that satisfies $y + h \in I$. Observe that

$$\frac{g(y + h) - g(y)}{h} = \frac{f(a(y + h), y + h) - f(a(y), y + h)}{h} + \frac{f(a(y), y + h) - f(a(y), y)}{h} \geq \frac{f(a(y), y + h) - f(a(y), y)}{h}.$$
Therefore, \( \lim \inf_{h \downarrow 0} h^{-1}(g(y + h) - g(y)) \geq \partial_y f(a(y), y) \). On the other hand, we also have
\[
\frac{g(y + h) - g(y)}{h} = \frac{f(a(y + h), y + h) - f(a(y + h), y)}{h}
+ \frac{f(a(y + h), y) - f(a(y), y)}{h}
\leq \frac{f(a(y + h), y + h) - f(a(y + h), y)}{h}
= \partial_y f(a(y + h), y(h))
\]
for some \( y(h) \in I \) with \( y(h) \to y \) as \( h \downarrow 0 \), where the last equation used the mean value theorem. Finally, using the continuity of \( \partial_y f \), we obtain
\[
\lim \sup_{h \downarrow 0} h^{-1}(g(y + h) - g(y)) \leq \partial_y f(a(y), y).
\]
This completes our proof. \( \square \)

**Proof of Theorem 2.** Define
\[
f(u, \theta) = \log 2 + F_{\mu_0}^{0,1}(u, h) - \frac{1}{2} \int_0^1 \alpha_{\mu_0}(s) s \xi(s) ds
\]
for \( (u, \theta) \in D[0, 1] \times [0, 1] \). Recall the definition of \( F_{\mu_0}^{0,1} \),
\[
f(u, \theta) = \log 2 + \mathbb{E} \left[ \log \cosh \left( h + \int_0^1 \alpha_{\mu_0}(s) \xi(s) u(s) ds + \int_0^1 \xi(s)^{1/2} dB(s) \right) \right.
\]
\[\left. - \frac{1}{2} \int_0^1 \alpha_{\mu_0}(s) \xi(s)(u(s)^2 + s) ds \right].\]
Its partial derivative with respect to \( \theta \) is clearly continuous on \( D[0, 1] \times [0, 1] \) and a direct computation gives
\[
\partial_0 f(u_{\mu_0}, \theta) = \mathbb{E} \left[ u_{\mu_0}(1) \int_0^1 \xi(s)(\alpha_{\mu}(s) - \alpha_{\mu_0}(s)) u_{\mu_0}(s) ds \right.
\]
\[\left. - \frac{1}{2} \int_0^1 \xi(s)(\alpha_{\mu}(s) - \alpha_{\mu_0}(s))(u_{\mu_0}(s)^2 + s) ds \right].\]
Since \( \{u_{\mu_0}(r)\}_{0 \leq r \leq 1} \) is a martingale from Lemma 1, the first term can be computed as
\[
\int_0^1 \xi(s)(\alpha_{\mu}(s) - \alpha_{\mu_0}(s)) \mathbb{E} u_{\mu_0}(s)^2 ds
\]
and thus,
\[
\partial_0 f(u_{\mu_0}, \theta) = \frac{1}{2} \int_0^1 \xi(s)(\alpha_{\mu}(s) - \alpha_{\mu_0}(s)) (\mathbb{E} u_{\mu_0}(s)^2 - s) ds.
\]
Applying Lemma 2 gives (12). From (12), if \( \mu_0 \) is the Parisi measure, then (ii) clearly holds. Assuming (ii), we note that for any \( \epsilon > 0 \), there exists some \( \delta > 0 \).
such that $P(\mu_\theta) - P(\mu_0) \geq -\varepsilon \theta$ whenever $0 < \theta < \delta$. This and the convexity of $P$ imply

$$P(\mu_\theta) \leq (1 - \theta)P(\mu_0) + \theta P(\mu),$$

so

$$\theta(P(\mu) - P(\mu_0)) = \theta P(\mu) + (1 - \theta)P(\mu_0) - P(\mu_0) \geq -\varepsilon \theta.$$  

Therefore, $P(\mu) \geq P(\mu_0) - \varepsilon$. Since this holds for all $\varepsilon > 0$, we have that $P(\mu) \geq P(\mu_0)$. In other words, $\mu_0$ is the minimizer of the Parisi functional and the uniqueness of the Parisi measure [2] implies that $\mu_0 = \mu_P$. So (ii) implies (i). Finally, we complete our proof by proving that (iii) yields (ii). Let $\mu \in M^k_d$ for some $k \geq 0$. Write $\mu = \sum_{p=0}^{k} a_p \delta_{q_p}$ with $a_p \geq 0$ and $\sum_{p=0}^{k} a_p = 1$. Define $\mu^p = \delta_{q_p}$ and $\mu^{p}_\theta = (1-\theta)\mu_0 + \theta \mu^p$. Now applying (iii) to $\mu^p$, we obtain

$$\frac{d}{d\theta} P(\mu^p_{\theta}) \bigg|_{\theta=0} = \frac{1}{2} \int_0^1 \zeta(s) (\alpha_{\mu^p}(s) - \alpha_{\mu_0}(s)) (\mathbb{E}u_{\mu_0}(s)^2 - s) \, ds \geq 0$$

and thus, using $\sum_{p=0}^{k} a_p = 1$ and $a_p \geq 0$,

$$\frac{d}{d\theta} P(\mu_{\theta}) \bigg|_{\theta=0} = \frac{1}{2} \int_0^1 \zeta(s) (\alpha_{\mu}(s) - \alpha_{\mu_0}(s)) (\mathbb{E}u_{\mu_0}(s)^2 - s) \, ds$$

$$= \sum_{p=0}^{k} a_p \cdot \frac{1}{2} \int_0^1 \zeta(s) (\alpha_{\mu^p}(s) - \alpha_{\mu_0}(s)) (\mathbb{E}u_{\mu_0}(s)^2 - s) \, ds$$

$$= \sum_{p=0}^{k} a_p \frac{d}{d\theta} P(\mu^{p}_{\theta}) \bigg|_{\theta=0} \geq 0.$$  

Here, the second equation used the observation that $\alpha_{\mu} = \sum_{p=0}^{k} a_p \alpha_{\mu^p}$. Since this inequality holds for arbitrary probability measures in $M_d$, an approximation argument using the definition of the right derivative of $P$ implies that $\frac{d}{d\theta} P(\mu_{\theta}) \big|_{\theta=0} \geq 0$ for all $\mu \in M$ and we obtain (ii).

**Proof of Proposition 1.** First, we claim that (13) and (14) hold for $q \in S \cap (0, 1)$. Assume $q \in S \cap (0, 1)$ is isolated. Define $\mu^1, \mu^2 \in M$ such that

$$\alpha_{\mu^1}(w) = \begin{cases} \alpha_{\mu^p}(w), & \text{if } w \in [0, q - \varepsilon) \cup [q, 1], \\ \alpha_{\mu^p}(q), & \text{if } w \in [q - \varepsilon, q), \end{cases}$$

$$\alpha_{\mu^2}(w) = \begin{cases} \alpha_{\mu^p}(w), & \text{if } w \in [0, q) \cup [q + \varepsilon, 1], \\ \alpha_{\mu^p}(q), & \text{if } w \in [q, q + \varepsilon). \end{cases}$$
From Theorem 2, we have
\[
\frac{d}{d\theta} P(\mu_0) \Big|_{\theta=0} = \frac{1}{2} \int_{q-\varepsilon}^{q} \varkappa(r)(\alpha_{\mu_0}(q) - \alpha_{\mu_0}(w))\left(\mathbb{E}u_{\mu_0}(w)^2 - w\right) dw \geq 0,
\]
(36) \[
\frac{d}{d\theta} P(\mu_0) \Big|_{\theta=0} = \frac{1}{2} \int_{q}^{q+\varepsilon} \varkappa(r)(\alpha_{\mu_0}(q) - \alpha_{\mu_0}(w))\left(\mathbb{E}u_{\mu_0}(w)^2 - w\right) dw \geq 0.
\]
Note that \(\alpha_{\mu_0}(q) > \alpha_{\mu_0}(w)\) for \(w \in [q - \varepsilon, q]\) and \(\alpha_{\mu_0}(q - \varepsilon) < \alpha_{\mu_0}(w)\) for \(w \in [q, q + \varepsilon]\). Since \(\mathbb{E}u_{\mu_0}(r)^2\) is a continuous function, the inequalities (36) imply that there exists some \(\varepsilon_0 > 0\) such that \(\mathbb{E}u_{\mu_0}(w)^2 \geq w\) on \([q - \varepsilon_0, q]\) and \(\mathbb{E}u_{\mu_0}(w)^2 \leq w\) on \([q, q + \varepsilon_0]\), which clearly gives (13). Now suppose that \(q\) is an accumulation point of \(S \cap (0, 1)\). Then there exists \((q_n)_{n \geq 1} \subset S \cap (0, 1)\) such that either \(q_n \uparrow q\) or \(q_n \downarrow q\). Assuming the first case, we consider \(\mu_3, \mu_4 \in \mathcal{M}\) defined through
\[
\alpha_{\mu_3}(w) = \begin{cases} 
\alpha_{\mu_0}(w), & \text{if } w \in [0, q - \varepsilon) \cup [q, 1], \\
\alpha_{\mu_0}(q), & \text{if } w \in [q - \varepsilon, q), 
\end{cases}
\]
\[
\alpha_{\mu_4}(w) = \begin{cases} 
\alpha_{\mu_0}(w), & \text{if } w \in [0, q - \varepsilon) \cup [q, 1], \\
\alpha_{\mu_0}(q - \varepsilon), & \text{if } w \in [q - \varepsilon, q). 
\end{cases}
\]
From Theorem 2, we have
\[
\frac{d}{d\theta} P(\mu_0) \Big|_{\theta=0} = \frac{1}{2} \int_{q-\varepsilon}^{q} \varkappa(w)(\alpha_{\mu_0}(q) - \alpha_{\mu_0}(w))\left(\mathbb{E}u_{\mu_0}(w)^2 - w\right) dw \geq 0,
\]
(37)
\[
\frac{d}{d\lambda} P(\mu_0) \Big|_{\lambda=0} = \frac{1}{2} \int_{q-\varepsilon}^{q} \varkappa(w)(\alpha_{\mu_0}(q - \varepsilon) - \alpha_{\mu_0}(w))\left(\mathbb{E}u_{\mu_0}(w)^2 - w\right) dw \geq 0.
\]
(38)
From the condition \(q_n \uparrow q\), we see that \(\alpha_{\mu_0}(q) > \alpha_{\mu_0}(w)\) and \(\alpha_{\mu_0}(q - \varepsilon) < \alpha_{\mu_0}(w)\) for \(w \in [q - \varepsilon, q]\). The inequality (37) then gives \(\mathbb{E}u_{\mu_0}(w)^2 \geq w\) for all \(w\) sufficiently close to \(q\) from the left-hand side. On the other hand, since \(q_n \uparrow q\), the inequality (38) implies that \(\mathbb{E}u_{\mu_0}(w)^2 \leq w\) for all \(w\) sufficiently close to \(q\), again from the left-hand side. Therefore, \(\mathbb{E}u_{\mu_0}(w)^2 = w\) on \([q - \varepsilon_0', q]\) for some \(\varepsilon_0' > 0\). Similarly, the case \(q_n \downarrow q\) also implies \(\mathbb{E}u_{\mu_0}(w)^2 = w\) on \([q, q + \varepsilon_0]\) for some \(\varepsilon_0 > 0\) by using
\[
\alpha_{\mu_5}(w) = \begin{cases} 
\alpha_{\mu_0}(w), & \text{if } w \in [0, q) \cup [q + \varepsilon, 1], \\
\alpha_{\mu_0}(q), & \text{if } w \in [q, q + \varepsilon), 
\end{cases}
\]
\[
\alpha_{\mu_6}(w) = \begin{cases} 
\alpha_{\mu_0}(w), & \text{if } w \in [0, q) \cup [q + \varepsilon, 1], \\
\alpha_{\mu_0}(q + \varepsilon), & \text{if } w \in [q, q + \varepsilon). 
\end{cases}
\]
These yield (13). To show (14), we note that from Lemma 1,

\[ \mathbb{E} u_{\mu_P}(b)^2 - \mathbb{E} u_{\mu_P}(a)^2 = \int_a^b \zeta(r) \mathbb{E} \partial_{xx} \Phi_{\mu_P}(w, X(w))^2 \, dw. \]

From the discussion, above, we see that either \( \mathbb{E} u_{\mu_P}(w)^2 \leq w \) on \([q, q']\) for some \( q' > q \) or \( \mathbb{E} u_{\mu_P}(w)^2 \geq w \) on \([q'', q] \) for some \( q'' < q \). If we are in the first situation, then for all \( s \in [q, q'] \), (39) implies

\[ \int_q^s \zeta(w) \mathbb{E} \partial_{xx} \Phi_{\mu_P}(w, X(w))^2 \, dw = \mathbb{E} u_{\mu_P}(s)^2 - \mathbb{E} u_{\mu_P}(q)^2 \leq s - q = \int_q^s 1 \, dw, \]

and hence (14). In the second situation, the same argument also yields

\[ \int_q^s \zeta(w) \mathbb{E} \partial_{xx} \Phi_{\mu_P}(w, X(w))^2 \, dw = \mathbb{E} u_{\mu_P}(q)^2 - \mathbb{E} u_{\mu_P}(s)^2 \leq q - s = \int_q^s 1 \, dw \]

for all \( s \in [q'', q] \), which concludes (14) and completes the proof of our claim.

Finally, note that Lemma 1 yields \( \mathbb{E} u_{\mu_P}(r)^2 < 1 \) for all \( 0 \leq r \leq 1 \). If \( 1 \in S \), one may take \( \mu = \delta_0 \) and \( \mu_0 = \mu_P \) in (12) to obtain a contradiction since \( \partial_{\theta} \mathcal{P}(\theta_0)|_{\theta=0} < 0 \). Hence, \( 1 \notin S \). If now \( 0 \in S \), then no matter if it is an isolated point or an accumulation point of \( S \), one can argue exactly in the same way as above to obtain \( \mathbb{E} u_{\mu_P}(w)^2 \leq w \) for all \( w \in [0, \varepsilon_0] \) for some \( \varepsilon_0 > 0 \). Consequently, \( \mathbb{E} u_{\mu_P}(0)^2 = 0 \). Since

\[ \int_0^s \zeta(w) \mathbb{E} \partial_{xx} \Phi_{\mu_P}(w, X(w))^2 \, dw = \mathbb{E} u_{\mu_P}(s)^2 - \mathbb{E} u_{\mu_P}(0)^2 \leq s - 0 = \int_0^s 1 \, dw \]

for all \( s \in [0, \varepsilon_0] \), we obtain (14) with \( q = 0 \). This completes our proof. \( \square \)

**Proof of Proposition 2.** For any \( q \in [0, 1] \), define \( \mu_0^q = (1 - \theta) \mu_0 + \theta \delta_q \) for \( 0 \leq \theta \leq 1 \). Since \( \mu_0^q \in \mathcal{M}_{d+1}^0 \), it follows from (16) that \( \frac{d}{d\theta} \mathcal{P}(\theta_0)|_{\theta=0} \geq 0 \). Therefore, \( \mu_0 \) is the Parisi measure by applying Theorem 2(ii). \( \square \)

### 4. The optimal stochastic control problem for \( \Psi_\mu \)

In this section, we will prove Theorems 4, 5 and 6 mostly following the ideas from [2]. Our argument relies on the following calculus lemma, which provides an explicit expression for the function \( \Psi_\mu \) when \( \mu \in \mathcal{M}_d \). As this lemma is not directly related to the core of our arguments, we defer its proof to the Appendix.

**Lemma 3.** Let \( 0 \leq a < b \leq 1 \) and \( 0 \leq m \leq 1 \). Recall \( \rho_{\xi, \nu} \) from (21). Suppose that they are differentiable on \([a, b]\). Let \( A \) be a smooth function on \( \mathbb{R}^2 \) with \( \limsup_{|x| \to \infty} |A(x)|/|x| < \infty \). For \( (s, x) \in [a, b] \times \mathbb{R}^2 \), set

\[ L(s, x) = \frac{1}{m} \log \mathbb{E} \exp \{ A(x_1 + y_1(s), x_2 + y_2(s)) \}, \]
where \((y_1(s), y_2(s))\) is a two-dimensional Gaussian random vector with mean zero and covariance:

\[
\begin{align*}
\mathbb{E}y_1(s)y_1(s) &= \xi'_1,1(\rho_{1,1}(b)) - \xi'_1,1(\rho_{1,1}(s)), \\
\mathbb{E}y_1(s)y_2(s) &= \xi'_1,2(\rho_{1,2}(b)) - \xi'_1,2(\rho_{1,2}(s)), \\
\mathbb{E}y_2(s)y_1(s) &= \xi'_2,1(\rho_{2,1}(b)) - \xi'_2,1(\rho_{2,1}(s)), \\
\mathbb{E}y_2(s)y_2(s) &= \xi'_2,2(\rho_{2,2}(b)) - \xi'_2,2(\rho_{2,2}(s)).
\end{align*}
\]

Then \(L\) satisfies

\[
\partial_s L = -\frac{1}{2}\left( \langle T, \nabla^2 L \rangle + m\langle T \nabla L, \nabla L \rangle \right)
\]

for \((s, x) \in [a, b) \times \mathbb{R}^2\) with terminal condition \(L(b, x) = A(x)\). Moreover, if \(\partial_{x_i} A\) is uniformly bounded by 1, so is \(\partial_{x_i} L\).

**Proof of Theorem 5** for \(\mu \in \mathcal{M}_d\). Suppose that \(\mu\) is atomic with jumps at \(\{q_p\}_{p=1}^k\), where \(q_p < q_{p+1}\) for \(1 \leq p \leq k - 1\). Let \(q_0 = 0, q_{k+1} = 1\) and \(m_p = \alpha_\mu(q_p)\) for \(0 \leq p \leq k\). Without loss of generality, we may assume that the nondifferentiable points of \(\rho_{\ell,\ell'}\) are located at \(\{q_p\}_{p=1}^k\) and in addition, \(q_j = s\) and \(q_{j'} = t\) for some \(0 \leq j < j' \leq k + 1\). Note that since \((y_1(s), y_2(s))\) equals \(\int_{q_p}^{q_{p+1}} T(w)^{1/2} dB(w)\) in distribution for each \(s \in [q_p, q_{p+1}]\) and \(0 \leq p \leq k\), we can write by Lemma 3,

\[
\Psi_\mu(\lambda, q_p, x) = \frac{1}{m_p} \log \mathbb{E} \exp m_p \Psi_\mu \left( \lambda, q_{p+1}, x + \int_{q_p}^{q_{p+1}} T(w)^{1/2} dB(w) \right),
\]

\[(41) \quad \forall j \leq p < j'.\]

We claim that

\[
\Psi_\mu(\lambda, s, x) \geq \max_{v \in \mathcal{D}[s, t]} F_{\mu}^{s, t}(\lambda, v, x).
\]

For \(v \in \mathcal{D}[s, t]\), set

\[
Z_p = \exp \left( -\frac{1}{2} \int_{q_p}^{q_{p+1}} m_p \left( T(w)v(w), v(w) \right) dw - \int_{q_p}^{q_{p+1}} m_p \left( T(w)^{1/2}v(w), dB(w) \right) \right).
\]

Define conditional probability measure \(\tilde{P}(A) = \mathbb{E}[1_A Z_p | \mathcal{F}_{q_p}]\) and set \(\tilde{B}(r) = \int_{q_p}^r m_p T(w)^{1/2} v(w) dw + B(r)\) for \(r \in [q_p, q_{p+1}]\). We use \(\mathbb{E}\) to denote the expectation with respect to \(\tilde{P}\). Since the Girsanov theorem [12], Theorem 5.1, says
that $\tilde{B}$ is a standard Brownian motion starting from $B(q_p)$ under $\tilde{\mathbb{P}}$, we can write
\[
E \exp m_p \Psi_\mu \left( \lambda, q_{p+1}, x + \int_{q_p}^{q_{p+1}} T(w)^{1/2} dB(w) \right)
= E \exp m_p \Psi_\mu \left( \lambda, q_{p+1}, x + \int_{q_p}^{q_{p+1}} T(w)^{1/2} d\tilde{B}(w) \right)
= E \left[ \exp m_p \Psi_\mu \left( \lambda, q_{p+1}, x + \int_{q_p}^{q_{p+1}} m_p T(w)v(w) dw 
+ \int_{q_p}^{q_{p+1}} T(w)^{1/2} dB(w) \right) 
\times \exp \left( - \frac{1}{2} \int_{q_p}^{q_{p+1}} m_p^2 |T(w)v(w), v(w)| dw 
- \int_{q_p}^{q_{p+1}} m_p T(w)^{1/2} v(w) \cdot dB(w) \right) \right] |\mathcal{F}_{q_p}.
\]
From (41) and Jensen’s inequality $m^{-1} \log E[\exp mA] \geq E[A]$, for any measurable $A$ and $m > 0$, it follows
\[
\Psi_\mu(\lambda, q_p, x) \geq E \left[ \Psi_\mu \left( \lambda, q_{p+1}, x + \int_{q_p}^{q_{p+1}} \alpha_\mu(w) T(w)v(w) dw 
+ \int_{q_p}^{q_{p+1}} T(w)^{1/2} dB(w) \right) 
- \frac{1}{2} \int_{q_p}^{q_{p+1}} \alpha_\mu(w) |T(w)v(w), v(w)| dw \right] |\mathcal{F}_{q_p}
\]
for all $j \leq p < j'$. Using this and conditional expectation, a decreasing iteration argument over $p$ from $j'-1$ to $j$ gives
\[
\Psi_\mu(\lambda, s, x) = \Psi_\mu(\lambda, q_j, x)
\geq E \left[ \Psi_\mu \left( \lambda, q_{j'}, x + \sum_{p=j}^{j'-1} \int_{q_p}^{q_{p+1}} \alpha_\mu(w) T(w)v(w) dw 
+ \int_{q_p}^{q_{p+1}} \xi(w)^{1/2} dB(w) \right) 
- \frac{1}{2} \sum_{p=j}^{j'-1} \int_{q_p}^{q_{p+1}} \alpha_\mu(w) |T(w)v(w), v(w)| dw \right] 
= F_{\mu}^{q_j,q_{j'}}(\lambda, v, x)
= F_{\mu}^{s,t}(\lambda, v, x).
\]
Since this is true for arbitrary \( v \in \mathcal{D}[s, t] \), this gives (42).

Note that since \( |\partial x_1 \Psi_\mu(\lambda, 1, \cdot)| \) and \( |\partial x_2 \Psi_\mu(\lambda, 1, \cdot)| \) are uniformly bounded above by 1, Lemma 3 combined with an iteration argument using (41) yields that \( |\partial x_1 \Psi_\mu(\lambda, r, \cdot)| \) and \( |\partial x_2 \Psi_\mu(\lambda, r, \cdot)| \) are also uniformly bounded by 1 for any \( s \leq r \leq t \), which clearly imply that \( v_\mu \in \mathcal{D}[s, t] \). Therefore, to complete the proof, it remains to show that \( \mathcal{F}_{s,t}^{x,t}(\lambda, v_\mu, x) = \Psi_\mu(s, x) \). To this end, we define

\[
Y(r) = \Psi_\mu(\lambda, r, X(r)) - \int_s^r \alpha_\mu(w)[T(w)v_\mu(w), v_\mu(w)]\,dw
- \int_s^r T(w)^{1/2}\,dB(w).
\]

Observe that

\[
\mathbb{E}Y(s) = \mathbb{E}\Psi_\mu(\lambda, s, X(s)) = \Psi_\mu(\lambda, s, x),
\]

\[
\mathbb{E}Y(t) = \mathcal{F}_{s,t}^{x,t}(\lambda, v_\mu, x).
\]

The use of Itô’s formula and (22) implies

\[
d\Psi_\mu = \partial_x \Psi_\mu \, dw + (\nabla \Psi_\mu, dX) + \frac{1}{2} \sum_{i,j=1}^2 \partial_{x_i x_j} \Psi_\mu \, d\langle X_i, X_j \rangle
= -\frac{1}{2} \left( T, \nabla^2 \Psi_\mu \right) + \alpha_\mu \left( T \nabla \Psi_\mu, \nabla \Psi_\mu \right) \, dw
+ \alpha_\mu \zeta \left( T \nabla \Psi_\mu, \nabla \Psi_\mu \right) \, dw + T^{1/2} \left( \nabla \Psi_\mu, dB \right) + \frac{1}{2} \left( T, \nabla^2 \Psi_\mu \right) \, dw
= \frac{1}{2} \alpha_\mu \left( T \nabla \Psi_\mu, \nabla \Psi_\mu \right) \, dw + T^{1/2} \left( \nabla \Psi_\mu, dB \right)
\]

and thus, \( dY = 0 \), which means that \( \mathcal{F}_{s,t}^{x,t}(\lambda, v_\mu, x) = \mathbb{E}Y(t) = \mathbb{E}Y(s) = \Psi_\mu(\lambda, s, x) \). This completes our proof. \( \square \)

**Proof of Theorem 4.** Let \( \mu, \mu' \in \mathcal{M}_d \). Since

\[
\Psi_\mu(\lambda, 1, x) = \Psi_{\mu'}(\lambda, 1, x) = \log(\cosh x_1 \cosh x_2 \cosh \lambda + \sinh x_1 \sinh x_2 \sinh \lambda),
\]

the mean value theorem implies

\[
|\Psi_\mu(\lambda, 1, x) - \Psi_{\mu'}(\lambda, 1, x')| \leq |x - x'|
\]

for \( \lambda \in \mathbb{R} \) and \( x, x' \in \mathbb{R}^2 \). Therefore, for any \( v \in \mathcal{D}[s, 1] \),

\[
|\mathcal{C}_{s,1}^{\lambda, 1}(\lambda, v, x) - \mathcal{C}_{s,1}^{\lambda, 1}(\lambda, v, x')| \leq \int_0^1 |\alpha_\mu(w) - \alpha_{\mu'}(w)| |T(w)v(w)| \,dw
\leq \sqrt{2}Kd(\mu, \mu'),
\]
where the last inequality used \( \| T(w) \| \leq K \) and \( |v(w)| \leq \sqrt{2} \). Also, we know that
\[
| \mathcal{L}_{\mu}^{s,1}(v) - \mathcal{L}_{\mu'}^{s,1}(v) | \leq K d(\mu, \mu')
\]
Combining these two inequalities together leads to
\[
| \mathcal{F}_{\mu}^{s,1}(\lambda, v, x) - \mathcal{F}_{\mu'}^{s,1}(\lambda, v, x) | \leq 3K d(\mu, \mu')
\]
and applying (24) gives the announced inequality. \( \square \)

**Proof of Theorem 5 for arbitrary \( \mu \).** This part of the proof relies on a standard approximation using a sequence of atomic\( \{ \mu_n \}_{n \geq 1} \) with weak limit \( \mu \). Just like the facts that \( \partial x_i \Phi_{1, \mu} \) is uniformly bounded by 1 and \( \lim_{n \to \infty} \partial x_i \Phi_{1, \mu_n} = \partial x_i \Phi_{1, \mu} \) uniformly for \( i = 1, 2 \), one may imitate the same approach as the Appendix in [1] to show that \( | \partial x_i \Psi_{1, \mu} | \leq 1 \), \( \| \nabla^2 \Psi_{1, \mu} \| \leq C \) and \( \lim_{n \to \infty} \nabla^i \Psi_{1, \mu_n} = \nabla^i \Psi_{1, \mu} \) uniformly for \( i = 1, 2 \). These give the existence of the SDE (25) and will lead to (i) and (ii) by using the results for atomic measures we established above and the same argument as in the proof of [2], Theorem 3. As the details are quite routine and follow exactly in the same lines, we will not reproduce them here. \( \square \)

**Proof of Theorem 6.** By the virtue of the Lipschitz property of \( \mu \mapsto \Psi_{1, \mu} \) with respect to the metric \( d \) defined by (5), it suffices to justify (26) for atomic \( \mu \) with jumps at \( \{ q_p \}_{p=1}^k \), where \( q_p < q_{p+1} \) for all \( 1 \leq p \leq k - 1 \). Let \( q_0 = 0 \) and \( q_{k+1} = 1 \). Without loss of generality, we may also assume that the nondifferentiable points of \( \rho_{\ell, \ell'} \) are all at \( \{ q_p \}_{p=1}^k \). Set
\[
\rho_p^{1,1} = \rho_{1,1}(q_p), \quad \rho_p^{2,2} = \rho_{2,2}(q_p), \quad \rho_p^{1,2} = \rho_{1,2}(q_p),
\]
for \( 0 \leq p \leq k + 1 \). Note that (18) follows from (20). Since
\[
\xi_{\ell, \ell'}(\rho_{p+1}) - \xi_{\ell, \ell'}(\rho_p) = \int_{q_p}^{q_{p+1}} \xi_{\ell, \ell'}(s) ds,
\]
the assumption \( T(s) \geq 0 \) implies that
\[
\left\langle \begin{bmatrix}
\xi_{1,1}^{1,1}(\rho_{p+1}^{1,1}) - \xi_{1,1}^{1,1}(\rho_p^{1,1}) & \xi_{1,2}^{1,2}(\rho_{p+1}^{1,2}) - \xi_{1,2}^{1,2}(\rho_p^{1,2}) \\
\xi_{2,1}^{2,2}(\rho_{p+1}^{2,2}) - \xi_{2,1}^{2,2}(\rho_p^{2,2}) & \xi_{2,1}^{2,1}(\rho_{p+1}^{2,1}) - \xi_{2,1}^{2,1}(\rho_p^{2,1})
\end{bmatrix} \begin{pmatrix} x \, x \end{pmatrix} \right\rangle = \int_{q_p}^{q_{p+1}} \langle T(s) x, x \rangle ds \geq 0
\]
for all \( x \in \mathbb{R}^2 \). Thus, the matrix on the left-hand side is positive semidefinite, which ensures that we can construct Gaussian random vectors \( (y_p^{\ell,1}, y_p^{\ell,2}) \) with mean zero and covariance:
\[
\mathbb{E} y_p^{\ell,1} y_p^{\ell',2} = \xi_{\ell, \ell'}(\rho_{p+1}^{\ell,1}) - \xi_{\ell, \ell'}(\rho_p^{\ell,1}).
\]
We now apply Theorem 3 with the choice \( m_p = \mu([0, q_p]) \) for \( 0 \leq p \leq k \) to get (19) as follows. Recall the definition of \( Y_p \) for \( 0 \leq p \leq k + 1 \) from Theorem 3. Define
\[
\begin{align*}
(Z_1^p, Z_2^p) &= \left( h_1 + \sum_{j=0}^{p-1} y_1^j, h_2 + \sum_{j=0}^{p-1} y_2^j \right), \quad \forall 1 \leq p \leq k + 1, \\
(Z_0^1, Z_0^2) &= (h_1, h_2).
\end{align*}
\]
Observe that \( Y_{k+1} = \Psi_1(\lambda, 1, Z_{k+1}^1, Z_{k+1}^2) \). If \( Y_{p+1} = \Psi_1(\lambda, q_{p+1}, Z_{p+1}^1, Z_{p+1}^2) \) for some \( 0 \leq p \leq k \), then Lemma 3 yields
\[
\begin{align*}
Y_p &= \frac{1}{m_p} \log \mathbb{E}_p \exp m_p Y_{p+1} \\
&= \frac{1}{m_p} \log \mathbb{E}_p \exp m_p \Psi_1(\lambda, q_{p+1}, Z_{p+1}^1 + y_p^1, Z_{p+1}^2 + y_p^2) \\
&= \Psi_1(\lambda, q_p, Z_{p+1}^1, Z_{p+1}^2)
\end{align*}
\]
and so \( Y_0 = \Phi_1(\lambda, 0, h_1, h_2) \). On the other hand, since \( \theta'_{\ell, e'}(w) = w \xi_{\ell, e'}''(w) \), we have that for \( \ell = \ell' \),
\[
\theta_{\ell, e'}(\rho_{p+1}^{\ell, \ell'}) - \theta_{\ell, e'}(\rho_{p}^{\ell, \ell'}) = \theta_{\ell, e'}(\rho_{\ell, e'}(q_{p+1})) - \theta_{\ell, e'}(\rho_{\ell, e'}(q_p)) = \int_{q_p}^{q_{p+1}} \rho_{\ell, e'}'(s) \theta_{\ell, e'}(\rho_{\ell, e'}(s)) ds
\]
and for \( \ell \neq \ell' \),
\[
\theta_{\ell, e'}(\rho_{p+1}^{\ell, \ell'}) - \theta_{\ell, e'}(\rho_{p}^{\ell, \ell'}) = \theta_{\ell, e'}(\rho_{\ell, e'}(q_{p+1})) - \theta_{\ell, e'}(\rho_{\ell, e'}(q_p)) = \int_{q_p}^{q_{p+1}} \iota_{\ell, e'}(s) \theta_{\ell, e'}(\rho_{\ell, e'}(s)) ds
\]
Consequently,
\[
\sum_{1 \leq \ell, \ell' \leq 2} \sum_{p=0}^{k} m_p (\theta_{\ell, e'}(\rho_{p+1}^{\ell, \ell'}) - \theta_{\ell, e'}(\rho_{p}^{\ell, \ell'})) = \int_0^1 \alpha_{\mu}(s) \left( \sum_{\ell \neq \ell'} \rho_{\ell, e'}(s) \xi_{\ell, e'}(s) + \iota \sum_{\ell \neq \ell'} \rho_{\ell, e'}(s) \xi_{\ell, e'}(s) \right) ds.
\]
Putting all these together into (19), we obtain (26). \( \square \)
5. The control of the GT bound. This section is devoted to proving Theorems 7 and 9 in Sections 2.3 and 2.4. We assume throughout this section that $X_1^N$ and $X_2^N$ are jointly Gaussian processes with mean zero and covariance,

\[
\begin{align*}
\mathbb{E}X_1^N(\sigma_1)X_1^N(\sigma_2) &= N\xi(R_{1,2}), \\
\mathbb{E}X_2^N(\sigma_1)X_2^N(\sigma_2) &= N\xi(R_{1,2}), \\
\mathbb{E}X_1^N(\sigma_1)X_2^N(\sigma_2) &= N\xi_0(R_{1,2}),
\end{align*}
\]

where $\xi$ and $\xi_0$ are of the form (3). Furthermore, we assume that they are convex on $[-1, 1]$, are not identically equal to zero and

\[
\xi''_0(s) \leq \xi''(|s|), \quad \forall s \in [-1, 1].
\]

Let $h \in \mathbb{R}$. Consider two mixed $p$-spin models,

\[
H_1^N(\sigma^1) = X_1^N(\sigma^1) + h \sum_{i=1}^N \sigma_i^1 \quad \text{and} \quad H_2^N(\sigma^2) = X_2^N(\sigma^2) + h \sum_{i=1}^N \sigma_i^2.
\]

Clearly, they share the same Parisi measure $\mu_P$. Denote by $\eta$ the minimum of the support of $\mu_P$. Recall the formulation of the two-dimensional GT bound from (26). For fixed $q \in S_N$, set

\[
\rho_{1,1}(s) = \rho_{2,2}(s) = s, \\
\rho_{1,2}(s) = \rho_{2,1}(s) = \min(|q|, s)
\]

for $s \in [0, 1]$. From (21), it follows that

\[
T(s) = \begin{bmatrix}
\xi''(s) & \xi_0''(ts) \\
\xi_0''(ts) & \xi''(s)
\end{bmatrix}, \quad \forall s \in [0, |q|) \quad \text{and} \quad T(s) = \begin{bmatrix}
0 & \xi''_0(s) \\
\xi''(s) & 0
\end{bmatrix}, \quad \forall s \in [|q|, 1].
\]

Consequently, from the condition (43), one sees that $T \geq 0$ on $[0, |q|)$; also it is clear that $T \geq 0$ on $[|q|, 1]$.

These allow us to apply Theorem 6 with arbitrary $\mu \in \mathcal{M}$ to get

\[
F_N(q) \leq 2\log 2 + \Psi_\mu(\lambda, 0, h, h) - \lambda q
\]

(45)

\[
-\left( \int_0^{1} \alpha_\mu(s)s\xi''(s) ds + \int_0^{|q|} \alpha_\mu(s)s\xi_0''(ts) ds \right).
\]

Note that the right-hand side of this inequality is indeed well defined for all $q \in [-1, 1]$. We denote this extension by $\Lambda_\mu(\lambda, q)$ and set $\Lambda(q) = \inf_{\lambda \in \mathbb{R}, \mu \in \mathcal{M}} \Lambda_\mu(\lambda, q)$. In the following two subsections, we will control $\Lambda(q)$ using the GT bound in two disjoint regions: $[-\eta, \eta]$ and $[-1, -\eta) \cup (\eta, 1]$. 
5.1. Behavior of $\Lambda$ in $[-\eta, \eta]$. The main result in this subsection is Proposition 3 below. This part of the argument appeared before in [7] and [20], Chapter 14. For completeness, we will give the detailed proof in the terminology of the variational representation (10) and (24). Recall that $\eta$ satisfies (27).

**Proposition 3.** If $h \neq 0$, then there exists some $q^* \in (0, \eta]$ such that

$$
\Lambda(q) < 2P(\mu_p)
$$

for any $q \in [-\eta, \eta] \setminus \{q^*\}$. Here, $q^* = \eta$ if $\xi = \xi_0$ and $q^* < \eta$ if $\xi \neq \xi_0$.

The proof of this proposition relies on the following technical lemma.

**Lemma 4.** Let $s \in [|q|, 1]$ and $x = (x_1, x_2) \in \mathbb{R}^2$. If $\mu, \mu' \in M$ satisfy $\mu = \mu'$ on $[|q|, 1]$, then

$$
\Psi_{\mu}(0, s, x) = \Phi_{\mu'}(s, x_1) + \Phi_{\mu'}(s, x_2),
$$

$$
\partial_x \Psi_{\mu}(0, s, x) = \partial_x \Phi_{\mu'}(s, x_1) \partial_x \Phi_{\mu'}(s, x_2).
$$

**Proof.** For any $|q| \leq s \leq 1$ and $v = (v_1, v_2) \in D[s, 1]$, we write by (44),

$$
x + \int_s^1 \alpha_{\mu}(r)T(r)v(r)dr + \int_s^1 T(r)^{1/2}dB(r)
$$

$$
= \left( x_1 + \int_s^1 \alpha_{\mu'}(r)\xi''(r)v_1(r)dr + \int_s^1 \xi''(r)^{1/2}dB_1(r), 
x_2 + \int_s^1 \alpha_{\mu'}(r)\xi''(r)v_2(r)dr + \int_s^1 \xi''(r)^{1/2}dB_2(r) \right)
$$

and

$$
\int_s^1 \alpha_{\mu}(r)|T(r)v(r), v(r)|dr
$$

$$
= \int_s^1 \alpha_{\mu'}(r)\xi''(r)v_1(r)^2dr + \int_s^1 \alpha_{\mu'}(r)\xi''(r)v_2(r)^2dr.
$$

From the terminal condition of $\Psi_{\mu}$ at (23),

$$
\Psi_{\mu}(0, 1, x) = \log \cosh x_1 + \log \cosh x_2 = \Phi_{\mu'}(1, x_1) + \Phi_{\mu'}(1, x_2),
$$

$$
\partial_x \Psi_{\mu}(0, 1, x) = \tanh x_1 \cdot \tanh x_2 = \partial_x \Phi_{\mu'}(1, x_1) \cdot \partial_x \Phi_{\mu'}(1, x_2).
$$

Using (10) and (24), the equations (48), (49) and (50) yield (46) since

$$
\Psi_{\mu}(0, s, x) = \max_{v = (v_1, v_2) \in D[s, 1]} F^{s, 1}_{\mu}(0, v, x)
$$

$$
= \max_{v_1 \in D[s, 1]} F^{s, 1}_{\mu}(v_1, x_1) + \max_{v_2 \in D[s, 1]} F^{s, 1}_{\mu}(v_2, x_2)
$$

$$
= \Phi_{\mu'}(s, x_1) + \Phi_{\mu'}(s, x_2).$$
To show (47), let \( v_\mu(r) = \nabla \Psi_\mu(0, r, X(r)) \) be the maximizer for \( \Psi_\mu(0, s, x) \), where \( X(r) = (X_1(r), X_2(r)) \) follows (25). The key observation is that the use of (46) leads to

\[
X_i(r) = x_i + \int_r^s \alpha_\mu'(w) \xi''(w) \partial_x \Phi_\mu'(w, X_i(w)) \, dw + \int_s^r \xi''(w)^{1/2} \, dB_i(w)
\]

for \( i = 1, 2 \). Therefore, \( \Phi_\mu'(s, X_i(s)) \) is the maximizer of (10) and \( \partial_x \Phi_\mu'(s, x_i) = \mathbb{E} \partial_x \Phi_\mu'(1, X_i(1)) \) from Lemma 2. Using these and Lemma 2 together with (48), (49) and (51), we obtain (47) since

\[
\partial_\lambda \Psi_\mu(0, s, x) = \partial_\lambda F_{s, 1 \mu}(0, v_\mu, x)
\]

\[
= \mathbb{E} \partial_x \Phi_\mu'(s, X_1(s)) \partial_x \Phi_\mu'(s, X_2(s)).
\]

\[ \square \]

**Proof of Proposition 3.** Assume \( h \neq 0 \). This proof has three major steps:

*Step I.* Define

\[
f(q) = \mathbb{E} \partial_x \Phi_{\mu p}(\eta, h + z_1(q)) \partial_x \Phi_{\mu p}(\eta, h + z_2(q))
\]

for \( q \in [-\eta, \eta] \), where \( z_1(q) \) and \( z_2(q) \) are jointly Gaussian with mean zero and covariance \( \mathbb{E} z_1(q)^2 = \xi'(\eta) = \mathbb{E} z_2(q)^2 \) and \( \mathbb{E} z_1(q) z_2(q) = \xi_0'(q) \). We claim that \( f \) maps \( [-\eta, \eta] \) into itself and has a unique fixed point \( q^* \in (0, \eta] \). Moreover, \( q^* = \eta \) if \( \xi = \xi_0 \) and \( q^* < \eta \) if \( \xi \neq \xi_0 \). To see these, recall from (13) and (14),

\[
\mathbb{E} \partial_x \Phi_{\mu p}(\eta, h + z_1(\eta))^2 = \mathbb{E} \partial_x \Phi_{\mu p}(\eta, h + z_2(\eta))^2
\]

(52)

\[
= \mathbb{E} \partial_x \Phi_{\mu p}(\eta, h + \int_0^{\eta} \xi(r)^{1/2} \, dB(r))^2 = \eta
\]

and

\[
\xi''(\eta) \mathbb{E} \partial_{xx} \Phi_{\mu p}(\eta, h + z_1(\eta))^2 = \xi''(\eta) \mathbb{E} \partial_{xx} \Phi_{\mu p}(\eta, h + z_2(\eta))^2
\]

(53)

\[
= \xi''(\eta) \mathbb{E} \partial_{xx} \Phi_{\mu p}(\eta, h + \int_0^{\eta} \xi(r)^{1/2} \, dB(r))^2 \leq 1.
\]

Using (52) and the Cauchy–Schwarz inequality, \( f \) evidently maps \( [-\eta, \eta] \) into itself, which implies the existence of a fixed point, say \( q^* \). To see its uniqueness, we apply the Gaussian integration by parts to obtain

\[
f'(q) = \xi_0''(q) \mathbb{E} \partial_{xx} \Phi_{\mu p}(\eta, h + z_1(q)) \partial_{xx} \Phi_{\mu p}(\eta, h + z_2(q)).
\]

Since \( \xi_0''(q) \leq \xi''(|q|) \leq \xi''(\eta) \) for \( q \in (-\eta, \eta) \), applying the Cauchy–Schwarz inequality and (53) to this formula leads to \( f' < 1 \) on \( (-\eta, \eta) \), so the fixed point
$q^*$ is unique. Now, since $\partial_x \Phi_{\mu P}$ is odd and strictly increasing (see Lemma 1) and $h \neq 0$, one sees that

$$f(0) = \left( \mathbb{E} \partial_x \Phi_{\mu P}(\eta, h + z) \right)^2 > 0,$$

where $z$ is Gaussian with mean zero and variance $\xi''(\eta)$. So $q^* \in (0, \eta]$. If $\xi = \xi_0$, (52) implies $q^* = \eta$; if $\xi \neq \xi_0$, then the Cauchy–Schwarz inequality and (52) leads to $q^* < \eta$. This ends the proof of our claim.

**Step II.** We check that

\begin{align}
\Lambda_{\mu P}(0, q) &= 2\mathcal{P}(\mu P), \quad (54) \\
\partial_h \Lambda_{\mu P}(0, q) &= f(q) - q \quad (55)
\end{align}

for $|q| \leq \eta$. Consider the variational representation (24) for $\Psi_{\mu P}(\lambda, 0, h, h)$ with $(s, t) = (0, \eta)$. Since $\alpha_{\mu P} = 0$ on $[0, \eta)$,

$$\mathcal{F}_{\mu P}^{0,\eta}(\lambda, v, h, h) = \mathbb{E} \Psi_{\mu P}(\lambda, \eta, (h, h) + \int_0^\eta T(r)^{1/2} dB(r)), \quad \forall v \in \mathcal{D}[0, \eta].$$

Observe that from (44), $\int_0^\eta T(r)^{1/2} dB(r)$ has the covariance structure

$$\int_0^\eta T(r) \, dr = \int_0^{|q|} dr \begin{bmatrix} \xi''(r) & \xi_0''(r) \\ \xi_0''(r) & 0 \end{bmatrix} + \int_{|q|}^\eta dr \begin{bmatrix} 0 & \xi''(r) \\ \xi_0''(r) & \xi'(r) \end{bmatrix}.$$

So we may write

$$\Psi_{\mu P}(\lambda, 0, h, h) = \max_{v \in \mathcal{D}[0, \eta]} \mathcal{F}_{\mu P}^{0,\eta}(\lambda, v, h, h) = \mathbb{E} \Psi_{\mu P}(\lambda, \eta, h + z_1(q), h + z_2(q)),$$

where $(z_1(q), z_2(q))$ is the Gaussian vector defined in Step I. Therefore, using (46),

$$\Psi_{\mu P}(0, 0, h, h) = \mathbb{E} \Psi_{\mu P}(0, \eta, h + z_1(q), h + z_2(q))$$

$$= \mathbb{E} \Phi_{\mu P}(\eta, h + z_1(q)) + \mathbb{E} \Phi_{\eta P}(\eta, h + z_2(q))$$

$$= 2\mathbb{E} \Phi_{\mu P}(\eta, h + z)$$

$$= 2\Phi_{\mu P}(0, h),$$

where $z$ is a Gaussian random variable with mean zero and variance $\mathbb{E} z^2 = \xi''(\eta)^2$, and the last equality used the assumption that $\alpha_{\mu} = 0$ on $[0, |q|)$ and the variational representation (10) for $\Phi_{\mu}(0, h)$ with $(s, t) = (0, \eta)$. In addition, applying (47),

$$\partial_h \Psi_{\mu P}(0, 0, h, h) = \mathbb{E} \partial_h \Psi_{\mu P}(0, \eta, h + z_1(q), h + z_2(q))$$

$$= \mathbb{E} \partial_h \Phi_{\mu P}(\eta, h + z_1(q)) \partial_h \Phi_{\mu P}(\eta, h + z_2(q))$$

$$= f(q).$$
Using again $\alpha_{\mu P} = 0$ on $[0, |q|)$, we then obtain

$$
\Lambda_{\mu P}(0, q) = 2\log 2 + \Psi_{\mu P}(0, 0, h, h) - \int_0^1 \alpha_{\mu P}(s)s\xi''(s)\,ds = 2P(\mu P),
$$

$$
\partial_\lambda \Lambda_{\mu P}(0, q) = \partial_\lambda \Psi_{\mu P}(0, 0, h, h) - q = f(q) - q,
$$

which complete the verification of (54) and (55).

**Step III.** From (55) and Step I, we know $\partial_\lambda \Lambda_{\mu P}(0, q) \neq 0$ for any $q \in [-\eta, \eta] \setminus \{q^*\}$. Depending on the sign of this quantity, we may decrease or increase $\lambda$ slightly to obtain $\Lambda_{\mu P}(\lambda, q) < \Lambda_{\mu P}(0, q)$. As a result, $\Lambda(q) < 2P(\mu P)$ by the definition of $\Lambda(q)$ and (54). This completes our proof. □

5.2. Behavior of $\Lambda$ outside of $[-\eta, \eta]$. For notational convenience, we set $\zeta(s) = \xi''(s)$ and $\zeta_0(s) = \xi''_0(\iota s)$ for $s \in [0, 1]$. Note that since $\xi$ and $\xi_0$ are convex and are not identically equal to zero, the function $\zeta$ is positive on $(0, 1)$ and so is $\zeta_0$ if $\iota = 1$. In addition, $\xi_0 \geq 0$ on $(0, 1]$ and $\zeta_0 = 0$ for at most a finite number of points if $\iota = -1$. The following proposition takes care of the behavior of $\Lambda$ on $[-1, -\eta) \cup (\eta, 1]$.

**Proposition 4.** The following two statements hold:

(i) For $-1 \leq q < -\eta$, if $\xi = \xi_0$ is even and $h \neq 0$, then $\Lambda(q) < 2\Phi_{\mu P}(0, h)$.

(ii) For $|q| > \eta$, if $\xi_0 < \xi$ and $\zeta/(\zeta + \zeta_0)$ is nondecreasing on $(0, 1]$, then $\Lambda(q) < 2\Phi_{\mu P}(0, h)$.

The essential idea to prove this proposition is to construct relevant $\mu \in M$ depending on $q$ and $\mu_P$ such that

$$
\int_0^1 \alpha_\mu(s)s\xi''(s)\,ds + \int_{|q|}^0 \alpha_\mu(s)s\xi''_0(\iota s)\,ds = \int_0^1 \alpha_{\mu P}(s)s\xi''(s)\,ds
$$

and

$$
\Psi_{\mu}(0, 0, h, h) < 2\Phi_{\mu P}(0, h).
$$

Once these are established, it will follow by definition that $\Lambda(q) \leq \Lambda_{\mu P}(0, q) < 2\Phi_{\mu P}(0, h)$. In order to get (56), one natural choice of $\mu$ is via (60) below. The major obstacle here comes from the derivation of (57) for such a choice of $\mu$. This will be handled through the variational representation for $\Psi_{\mu}$ and $\Phi_{\mu_P}$. A key lemma we will need along the line is the global uniqueness of the maximizer for $\Phi_{\mu_P}$.

**Lemma 5.** Let $0 \leq s < t \leq 1$. Suppose that $u^*$ is a maximizer of $\Phi_{\mu}(0, x) = \max_{u \in D[0, t]} F_{\mu, t}^0(u, x)$. If $\alpha_{\mu} > 0$ on $(s, t)$, then $u^*(r) = \partial_x \Phi_{\mu}(r, X(r))$ for $r \leq t$, where

$$
X(r) = x + \int_s^r \alpha_{\mu}(w)\xi(w)\partial_x \Phi_{\mu}(w, X(w))\,dw + \int_s^r \xi(w)^{1/2}dB(w),
$$

$\forall s \leq r \leq t$. 

In other words, the maximizer is unique under the assumption $\alpha_\mu > 0$ on $(s, t)$.

**Proof.** Let $\{a_i\}_{i=0}^n$ be a regular partition of $[s, t]$ with $\int_{a_i}^{a_{i+1}} \alpha_\mu(r) \zeta(r) \, dr < 1$ for $1 \leq i < n$. Define $u_i \in D[a_i, t]$ by $u_i(w) = u^*(w)$ for $a_i \leq w \leq t$ and $v_i \in D[0, a_i]$ by $v_i(w) = u^*(w)$ for $0 \leq w \leq a_i$. Set

$$y_i = x + \int_0^{a_i} \alpha_\mu(w) \zeta(w) v_i(w) \, dw + \int_0^{a_i} \zeta(w)^{1/2} \, dB(w).$$

Using conditional expectation,

$$\Phi_\mu(0, x) = \mathbb{E}(\mathbb{E}[C_{\mu}^{0,1}(u^*, x) - L_{\mu}^{0,1}(u^*) | y_i])$$

$$= \mathbb{E}(\mathbb{E}[C_{\mu}^{a_i,1}(u_i, y_i) - L_{\mu}^{a_i,1}(u_i) | y_i]) - \mathbb{E}L_{\mu}^{0,a_i}(v_i)$$

$$\leq \mathbb{E}\Phi_\mu(a_i, y_i) - \mathbb{E}L_{\mu}^{0,a_i}(v_i)$$

$$= F_{\mu}^{0,a_i}(v_i, x)$$

$$\leq \Phi_\mu(0, x),$$

which implies that $F_{\mu}^{0,a_i}(v_i, x)$'s are the same for all $0 \leq i \leq n$. Using this, we obtain that

$$\mathbb{E}C_{\mu}^{0,a_i}(v_i, x) - \mathbb{E}L_{\mu}^{0,a_i}(v_i) = F_{\mu}^{0,a_i}(v_i, x)$$

$$= F_{\mu}^{0,a_i+1}(v_{i+1}, x)$$

$$= \mathbb{E}C_{\mu}^{0,a_i+1}(v_{i+1}, x) - \mathbb{E}L_{\mu}^{0,a_i+1}(v_{i+1})$$

and thus

$$\mathbb{E}C_{\mu}^{0,a_i}(v_i, x) = \mathbb{E}C_{\mu}^{0,a_i+1}(v_{i+1}, x) - \mathbb{E}L_{\mu}^{a_i,a_i+1}(u'_i)$$

$$= \mathbb{E}C_{\mu}^{a_i,a_i+1}(u'_i, y_i) - \mathbb{E}L_{\mu}^{a_i,a_i+1}(u'_i)$$

$$= \mathbb{E}(\mathbb{E}[C_{\mu}^{a_i,a_i+1}(u'_i, y_i) - \mathbb{E}L_{\mu}^{a_i,a_i+1}(u'_i) | y_i])$$

$$\leq \mathbb{E}\max_{u' \in D[a_i, a_{i+1}]} F_{\mu}^{a_i,a_{i+1}}(u', y_i)$$

$$= \mathbb{E}\Phi_\mu(a_i, y_i),$$

where $u'_i \in D[a_i, a_{i+1}]$ is the restriction of $u^*$ to $[a_i, a_{i+1}]$. Since

$$\max_{u' \in D[a_i, a_{i+1}]} F_{\mu}^{a_i,a_{i+1}}(u', y) = \Phi_\mu(a_i, y), \quad \forall y \in \mathbb{R},$$

and $\mathbb{E}C_{\mu}^{0,a_i}(v_i, x) = \mathbb{E}\Phi_\mu(a_i, y_i)$, these and (58) force that when conditioning on $y_i$, $u'_i$ is the maximizer to the variational problem $\max_{u' \in D[a_i, a_{i+1}]} F_{\mu}^{a_i,a_{i+1}}(u', y_i)$. Therefore, applying the local uniqueness of the maximizer for $(s, t) = (a_i, a_{i+1})$
in Theorem 1 leads to \( u_i'(r) = \partial_x \Phi_\mu(r, X_i(r)) \) on \([a_i, a_{i+1}]\), where \( X_i = (X_i(w))_{a_i \leq w \leq a_{i+1}} \) is the solution to
\[
X_i(r) = y_i + \int_{a_i}^{a_{i+1}} \alpha_\mu(w) \xi(w) \partial_x \Phi_\mu(w, X_i(w)) \, dw + \int_{a_i}^{a_{i+1}} \xi(w)^{1/2} \, dB(w).
\]
Concatenating all these from \( i = 0 \) to \( n - 1 \) together gives the announced result. \( \square \)

The core ingredient of the matter is a quantitative error estimate between the one- and two-dimensional Parisi PDEs for a specific choice of \( \mu \) given in the proposition below.

**PROPOSITION 5.** Assume that \(|q| > \eta \) and

\[
\frac{\xi(s)}{\xi(s) + \xi_0(s)}
\]

is nondecreasing on \((0, 1] \). Define \( \mu \in \mathcal{M} \) by
\[
\alpha_\mu(s) = \begin{cases} 
\frac{\alpha_\mu_P(s) \xi(s)}{\xi(s) + \xi_0(s)}, & \text{if } s \in [0, |q|), \\
\alpha_\mu_P(s), & \text{if } s \in [|q|, 1].
\end{cases}
\]

(i) We have that
\[
\Psi_\mu(0, 0, x) = \Phi_\mu_p(0, x_1) + \Phi_\mu_p(0, x_2)
\]
\[
- \frac{1}{2} \int_0^{|q|} \alpha_\mu_p \xi_0 (\xi - \xi_0) \mathbb{E}(v_1 - \epsilon v_2)^2 \, dw,
\]
where \( v_\mu = (v_1, v_2) \) is the maximizer to the variational problem (5) for \( \Psi_\mu(0, 0, x_1, x_2) \) using \((s, t) = (0, |q|)\).

(ii) Define
\[
(u_1(r), u_2(r)) = \frac{1}{\xi(r) + \xi_0(r)} T(r) v_\mu(r),
\]
\[
(B_1(r), B_2(r)) = \frac{1}{\xi(r)^{1/2}} T(r)^{1/2} B(r)
\]
for \( 0 \leq r \leq |q| \). If
\[
\Psi_\mu(0, 0, x) = \Phi_\mu_p(0, x_1) + \Phi_\mu_p(0, x_2),
\]
then \( u_1 \) and \( u_2 \) are the maximizers for the variational problem (10) of \( \Phi_\mu_p(0, x_1) \) and \( \Phi_\mu_p(0, x_2) \) using \((s, t) = (0, |q|)\) with respect to the standard Brownian motions \( B_1 \) and \( B_2 \), respectively. Moreover, on the interval \([\eta, |q|]\), they are equal
to
\[ u_1(r) = \partial_x \Phi_{\mu_p}(r, X_1(r)), \]
\[ u_2(r) = \partial_x \Phi_{\mu_p}(r, X_2(r)), \]

where \((X_1(r))_{0 \leq r \leq |q|}\) and \((X_2(r))_{0 \leq r \leq |q|}\) satisfy
\[ X_1(r) = x_1 + \int_0^r \alpha_{\mu_p}(w) \zeta(w) \partial_x \Phi_{\mu_p}(w, X_1(w)) \, dw + \int_0^r \zeta(w)^{1/2} \, dB_1(w), \]
\[ X_2(r) = x_2 + \int_0^r \alpha_{\mu_p}(w) \zeta(w) \partial_x \Phi_{\mu_p}(w, X_2(w)) \, dw + \int_0^r \zeta(w)^{1/2} \, dB_2(w). \]

**PROOF.** Note that the well-definedness of \(\mu\) is guaranteed by (59). Let \(v_\mu = (v_1, v_2)\) be the maximizer to the variational problem (5) for \(\Psi_\mu\) with \((s, t) = (0, |q|)\). Set \((u_1, u_2)\) via (62). Here, \(u_1, u_2\) are progressively measurable processes with respect to the filtration \((\mathcal{F}_r)_{r \geq 0}\) and \(B_1, B_2\) are (correlated) standard Brownian motions. Denote by \((C_{\mu_p, 0, |q|}, L_{\mu_p, 0, |q|}, F_{\mu_p, 0, |q|})\) and \((C_{\mu_p, 0, |q|}, L_{\mu_p, 0, |q|}, F_{\mu_p, 0, |q|})\) the functionals defined in the same way as (9) using \(B_1\) and \(B_2\), respectively. Observe that from (46) and the definition of \(u_1, u_2\),
\[ C_{\mu_p, 0, |q|}(0, v_\mu, x) = \Phi_{\mu_p}(|q|, x_1 + \int_0^{|q|} \alpha_{\mu_p}(w) \zeta(w) u_1(w) \, dw + \int_0^{|q|} \zeta(w)^{1/2} \, dB_1(w), \]
\[ + \Phi_{\mu_p}(|q|, x_2 + \int_0^{|q|} \alpha_{\mu_p}(w) \zeta(w) u_2(w) \, dw + \int_0^{|q|} \zeta(w)^{1/2} \, dB_2(w) \]
\[ = C_{\mu_p, 0, |q|}(u_1, x_1) + C_{\mu_p, 0, |q|}(u_2, x_2). \]

In addition, noting that a direct computation gives
\[ (u_1(r), u_2(r)) = \left( \frac{\zeta(r)v_1(r) + \zeta_0(r)v_2(r)}{\zeta(r) + \zeta_0(r)}, \frac{\zeta_0(r)v_1(r) + \zeta(r)v_2(r)}{\zeta(r) + \zeta_0(r)} \right), \]

it follows that
\[ \frac{1}{\zeta + \zeta_0} \langle Tv_\mu, v_\mu \rangle - u_1^2 - u_2^2 \]
\[ = \left( \frac{\zeta}{\zeta + \zeta_0} - \frac{\zeta_0^2 (\zeta + \zeta_0)^2}{(\zeta + \zeta_0)^2} \right) (v_1^2 + v_2^2) + 2\zeta_0 \left( \frac{1}{\zeta + \zeta_0} - \frac{2\zeta}{(\zeta + \zeta_0)^2} \right) v_1 v_2 \]
\[ = \frac{\zeta_0(\zeta - \zeta_0)}{(\zeta + \zeta_0)^2} (v_1 - v_2)^2, \]
which implies
\[
\mathcal{L}_{\mu}^{0,|q|}(v_{\mu}) = \mathcal{L}_{\mu P,1}^{0,|q|}(u_1) + \mathcal{L}_{\mu P,2}^{0,|q|}(u_2) + \frac{1}{2} \int_{0}^{|q|} \frac{\alpha_{\mu P} \xi \xi_0 (\xi - \xi_0)}{(\xi + \xi_0)^2} \mathbb{E}(v_1 - iv_2)^2 \, dw.
\]
Combining these together, the variational representations for \(\Psi_{\mu}(0,0,x)\) and \(\Phi_{\mu P}(0,h)\) yield (61) since
\[
\Psi_{\mu}(0,0,x) = F_{\mu P,1}^{0,|q|}(u_1, x_1) + F_{\mu P,2}^{0,|q|}(u_2, x_2) - \frac{1}{2} \int_{0}^{|q|} \frac{\alpha_{\mu P} \xi \xi_0 (\xi - \xi_0)}{(\xi + \xi_0)^2} \mathbb{E}(v_1 - iv_2)^2 \, dw.
\]
If \(\Psi_{\mu}(0,0,x) = \Phi_{\mu P}(0, x_1) + \Phi_{\mu P}(0, x_2)\), this inequality implies that \(u_1\) and \(u_2\) are the maximizers of the variational representations,
\[
\Phi_{\mu}(0, x_1) = \max_{u \in \mathcal{D}[0,|q|]} F_{\mu P,1}^{0,|q|}(u, x_1) \quad \text{and} \quad \Phi_{\mu}(0, x_2) = \max_{u \in \mathcal{D}[0,|q|]} F_{\mu P,2}^{0,|q|}(u, x_2)
\]
corresponding to the Brownian motions \(B_1\) and \(B_2\), respectively. Since \(\alpha_{\mu} > 0\) on \((\eta, |q|]\), Lemma 5 concludes (ii). □

**Proof of Proposition 4.** First, note that the measure \(\mu\) in (60) is well defined since the function \(\xi / (\xi + \xi_0)\) under both assumptions (i) and (ii) is nondecreasing on \((0,1]\). We plug this \(\mu\) into (45) and let \(\lambda = 0\) to obtain
\[
\Lambda(q) \leq 2 \log 2 + \Psi_{\mu}(0,0,h,h) - \int_{0}^{1} \alpha_{\mu P}(s)s\xi''(s) \, ds.
\]
Thus, to complete the proof, we only need to verify that \(\Psi_{\mu}(0,0,h,h) < 2\Phi_{\mu P}(0,h)\). Suppose the equality holds. Proposition 5(ii) implies that for any \(\eta \leq r \leq |q|\),
\[
\begin{align*}
&u_1(r) = \partial_x \Phi_{\mu P}(r, X_1(r)), \\
&u_2(r) = \partial_x \Phi_{\mu P}(r, X_2(r)),
\end{align*}
\]
where \((X_1(r))_{0 \leq r \leq |q|}\) and \((X_2(r))_{0 \leq r \leq |q|}\) satisfy
\[
\begin{align*}
X_1(r) &= h + \int_{0}^{r} \alpha_{\mu P}(w)\xi(w)\partial_x \Phi_{\mu P}(w, X_1(w)) \, dw + \int_{0}^{r} \xi(w)^{1/2} \, dB_1(w), \\
X_2(r) &= h + \int_{0}^{r} \alpha_{\mu P}(w)\xi(w)\partial_x \Phi_{\mu P}(w, X_2(w)) \, dw + \int_{0}^{r} \xi(w)^{1/2} \, dB_2(w).
\end{align*}
\]
Our proof will clearly be completed by the following two cases.
Case I: $-1 \leq q < -\eta$, $\xi = \xi_0$ is even and $h \neq 0$. Since $\iota = -1$, these assumptions combined with (62) and (63) lead to $u_1 = -u_2$ and $B_1 = -B_2$. Consequently, adding the two equations in (64) together implies $X_1(r) + X_2(r) = 2h$ for $q \leq r \leq -\eta$. On the other hand, since $\partial_x \Phi_{\mu_p}(r, \cdot)$ is odd and strictly increasing from Lemma 1, the equation

$$
\partial_x \Phi_{\mu_p}(r, X_1(r)) = u_1(r) = -u_2(r) = \partial_x \Phi_{\mu_p}(r, -X_2(r))
$$

implies $X_1(r) = -X_2(r)$, which contradicts $X_1(r) + X_2(r) = 2h$ since $h \neq 0$.

Case II: $\xi_0 < \xi$ on $(0, 1]$. Since $\xi > \xi_0 \geq 0$ and $\xi_0 = 0$ for at most a finite number of points, we deduce from (61) and the continuity of $v_1, v_2$ that $v_1 = iv_2$ on $[\eta, |q|]$. From (62), it then follows that $u_1 = u_2$ on $[\eta, |q|]$. Again, using the facts that $\partial_x \Phi_{\mu_p}(r, \cdot)$ is odd and strictly increasing, we conclude $X_1 = iX_2$ on $[\eta, |q|]$ and from (64), for $r \in [\eta, |q|]$,

$$
0 = X_1(r) - iX_2(r) = (1 - i)h + \int_0^r \xi(w)^{1/2}d(B_1(w) - iB_2(w)).
$$

This forces that $B_1 = iB_2$ and, therefore, (63) implies that $\xi(r)^2 - \xi_0(r)^2 = \det T(r) = 0$ for $r \in [\eta, |q|]$. This leads to a contradiction since $\xi > \xi_0 \geq 0$. \square

5.3. Proof of Theorems 7, 8 and 9. Before we start, it is crucial to notice that $\Psi_\mu(\lambda, 0, h, h)$ is a continuous function in $q \in [-1, 1]$ for any $\mu \in \mathcal{M}$ and $\lambda \in \mathbb{R}$. This can be easily shown by following a similar argument as the proof of Theorem 4. Thus, $\Lambda$ is upper semicontinuous on $[-1, 1]$.

Proof of Theorem 7. Note that $H^1_N = H_N = H^2_N$ since $\xi = \xi_0$. Let $\varepsilon > 0$. Denote

$$
q' = \text{Argmax}_{q \in [-1, \eta - \varepsilon]} \Lambda(q).
$$

Here, the existence of $q'$ is guaranteed by the upper semicontinuity of $\Lambda$ on $[-1, \eta - \varepsilon]$. If the assumption (i) holds, then Proposition 3 and the first assertion of Proposition 4 together imply $\Lambda(q) < 2\mathcal{P}(\mu_p)$ for $q \in [-1, \eta - \varepsilon]$, and thus,

$$
\Lambda(q) \leq \Lambda(q') < 2\mathcal{P}(\mu_p), \quad \forall q \in [-1, \eta - \varepsilon].
$$

Now suppose that the condition (ii) is true. Then the series $\xi$ must contain some term $\beta_p s^p$ with $\beta_p \neq 0$ for some odd $p$. This implies that for any $q \in [-1, -\eta)$,

$$
\xi_0(s) = \xi''(-s) < \xi''(s) = \xi(s), \quad \forall s \in (0, 1],
$$

which combined with (28) yields $\Lambda(q) < 2\mathcal{P}(\mu_p)$ for $q \in [-1, -\eta]$ by the second assertion of Proposition 4. Since $h \neq 0$, we can use Proposition 3 to obtain $\Lambda(q) <
$2\mathcal{P}(\mu_p)$ for $q \in [-\eta, \eta - \varepsilon]$, and consequently (65) is valid. In summary, the two assumptions (i) and (ii) lead to

$$\limsup_{N \to \infty} \max_{q \in S_N \cap [-1, \eta - \varepsilon]} \frac{1}{N} \mathbb{E} \log \sum_{R_{1,2}=q} \exp(H_N(\sigma_1) + H_N(\sigma_2)) < 2\mathcal{P}(\mu_p).$$

Finally, from this inequality, (29) can be obtained by using the Gaussian concentration of measure and the Parisi formula. Since this part of the argument is very standard and has appeared in several places, for example, [20], Section 14.12, we omit the details. □

PROOF OF THEOREM 8. Again $H^1_N = H_N = H^2_N$. Note that $\eta = 0$ since $h = 0$. Recall the maximizer $q'$ from the proof of Theorem 7. From the given assumption of Theorem 8, one sees that (66) is also valid. Thus, the second assertion of Proposition 4 implies $\Lambda(q) < 2\mathcal{P}(\mu_p)$ for all $q \in [-\eta, -\varepsilon]$. As a result,

$$\Lambda(q) \leq \Lambda(q') < 2\mathcal{P}(\mu_p), \quad \forall q \in [-1, -\varepsilon],$$

from which it follows that

$$\limsup_{N \to \infty} \max_{q \in S_N \cap [-1, -\varepsilon]} \frac{1}{N} \mathbb{E} \log \sum_{R_{1,2}=q} \exp(H_N(\sigma_1) + H_N(\sigma_2)) < 2\mathcal{P}(\mu_p).$$

The rest of the proof can be completed by an identical argument as the last part of the proof of Theorem 7. □

PROOF OF THEOREM 9. Note that $\xi \neq \xi_0$. Let $q^* \in (0, \eta)$ be the constant stated in Proposition 3 if $h \neq 0$ and set $q^* = 0$ if $h = 0$. From the upper semicontinuity of $\Lambda$, for $\varepsilon > 0$, let $q''$ be the maximizer of

$$\max_{q \in [-1, 1] : |q - q^*| \geq \varepsilon} \Lambda(q).$$

Note that from the assumptions (33) and (34),

(67) $\Lambda(q) < 2\mathcal{P}(\mu_p), \quad \forall q \in [-1, -\eta) \cup (\eta, 1]$ by the second statement of Proposition 4. If $h = 0$, then $\eta = q^* = 0$ and this inequality implies

(68) $\Lambda(q) \leq \Lambda(q'') < 2\mathcal{P}(\mu_p), \quad \forall q \in [-1, 1] \text{ with } |q - q^*| \geq \varepsilon.$

If $h \neq 0$, then Proposition 3 gives $\Lambda(q) < 2\mathcal{P}(\mu_p)$ for $q \in [-\eta, \eta] \setminus \{q^*\}$. This together with (67) gives (68) by the second assertion of Proposition 4. Therefore, we have shown that

$$\limsup_{N \to \infty} \max_{q \in S_N : |q - q^*| \geq \varepsilon} \frac{1}{N} \mathbb{E} \log \sum_{R_{1,2}=q} \exp(H^1_N(\sigma_1) + H^2_N(\sigma_2)) < 2\mathcal{P}(\mu_p).$$

Using this inequality, (35) follows by applying the Gaussian concentration of measure and the Parisi formula. Once again, we skip this part of the argument as it can be found in great detail in the proof of [7], Theorem 7. □
Proof of Lemma 3. We argue by applying the Gaussian integration by parts formula. Define
\[c_1(s) = (\xi_1'(\rho_1,2(b)) - \xi_1'((\rho_1,2(s))) = \int_s^b \rho_1'(l)\xi_1''(\rho_1,2(l)) \, dl \geq 0,\]
\[c_2(s) = (\xi_2'(\rho_2,1(b)) - \xi_2'((\rho_2,1(s))) = \int_s^b \rho_2'(l)\xi_2''(\rho_2,1(l)) \, dl \geq 0,\]
\[d_1(s) = \xi_1'(\rho_1,1(b)) - \xi_1'(\rho_1,1(s)) - c_1(s),\]
\[d_2(s) = \xi_2'(\rho_2,2(b)) - \xi_2'(\rho_2,2(s)) - c_2(s).\]

We parametrize \((y_1(s), y_2(s))\) as
\[\big(y_1(s), y_2(s)\big) = \left(\sqrt{c_1(s)}z_0 + \sqrt{d_1(s)}z_1, \sqrt{c_2(s)}z_0 + \sqrt{d_2(s)}z_2\right),\]
where \(z_0, z_1, z_2\) are i.i.d. standard Gaussian. Note \(c_1 = c_2\) by the symmetry of \(T\). Recall \(\xi_{\ell, e'}\) from (21). Observe that
\[\mathbb{E}y_1'(s)y_1(s) = -\frac{\xi_{1,1}(s)}{2},\]
\[\mathbb{E}y_2'(s)y_2(s) = -\frac{\xi_{2,2}(s)}{2},\]
\[\mathbb{E}y_1'(s)y_2(s) = -\frac{\xi_{1,2}(s)}{2} = -\frac{\xi_{2,1}(s)}{2} = \mathbb{E}y_2'(s)y_1(s),\]
where \(y_1'\) and \(y_2'\) are the derivatives of \(y_1\) and \(y_2\) with respect to \(s\), respectively. From the growth condition of \(A\), it allows us to apply the Gaussian integration by parts to obtain
\[\partial_s L = \frac{1}{\mathbb{E}e^{mA}} \mathbb{E} \left[ y_1'\partial_{x_1} A + y_2'\partial_{x_2} A \right] e^{mA} \]
\[= \frac{1}{\mathbb{E}e^{mA}} \left( \mathbb{E}(y_1'y_1) \mathbb{E}(\partial_{x_1} A + m(\partial_{x_1} A)^2) e^{mA} \right) \]
\[+ \mathbb{E}(y_1'y_2) \mathbb{E}(\partial_{x_2} A + m(\partial_{x_2} A)^2) e^{mA} \]
\[+ \frac{1}{\mathbb{E}e^{mA}} \left( \mathbb{E}(y_2'y_2) \mathbb{E}(\partial_{x_2} A + m(\partial_{x_2} A)^2) e^{mA} \right) \]
\[= -\frac{1}{2\mathbb{E}e^{mA}} \mathbb{E} \left[ \xi_{1,1}(\partial_{x_1} A + m(\partial_{x_1} A)^2) \right. \]
\[+ \xi_{1,2}(\partial_{x_1} A + m(\partial_{x_1} A)(\partial_{x_2} A)) \left. e^{mA} \right].\]
\[-\frac{1}{2}\mathbb{E} e^m A \mathbb{E}\left[\xi_{2, 2}(\partial_{x_2 x_2} A + m(\partial_{x_2} A)^2)\right.\]
\[+ \xi_{2, 1}(\partial_{x_1 x_2} A + m(\partial_{x_1} A)(\partial_{x_2} A))\left.\right] e^m A\]
\[= -\frac{1}{2}\mathbb{E} e^m A \mathbb{E}\left[(T, \nabla^2 A) + m(T \nabla A, \nabla A)\right] e^m A.\]

On the other hand, a direct computation gives
\[
\partial_{x_1} L = \frac{\mathbb{E} \partial_{x_1} A e^m A}{\mathbb{E} e^m A},
\]
\[
\partial_{x_2} L = \frac{\mathbb{E} \partial_{x_2} A e^m A}{\mathbb{E} e^m A}
\]
(69)
and
\[
\partial_{x_1 x_1} L = \frac{\mathbb{E} (\partial_{x_1 x_1} A + m(\partial_{x_1} A)^2)e^m A}{\mathbb{E} e^m A} - m\left(\frac{\mathbb{E} \partial_{x_1} A e^m A}{\mathbb{E} e^m A}\right)^2,
\]
\[
\partial_{x_2 x_2} L = \frac{\mathbb{E} (\partial_{x_2 x_2} A + m(\partial_{x_2} A)^2)e^m A}{\mathbb{E} e^m A} - m\left(\frac{\mathbb{E} \partial_{x_2} A e^m A}{\mathbb{E} e^m A}\right)^2,
\]
\[
\partial_{x_1 x_2} L = \partial_{x_2 x_1} L
\]
\[= \frac{\mathbb{E} (\partial_{x_1 x_2} A + m(\partial_{x_1} A)(\partial_{x_2} A))e^m A}{\mathbb{E} e^m A} - m\left(\frac{\mathbb{E} \partial_{x_1} e^m A}{\mathbb{E} e^m A}\right)\left(\frac{\mathbb{E} \partial_{x_2} A e^m A}{\mathbb{E} e^m A}\right).
\]
Using these, one may easily check that
\[
\langle T, \nabla^2 L \rangle + m\langle T \nabla L, \nabla L \rangle
\]
\[= -\frac{1}{e^m A} \mathbb{E}\left[(T, \nabla^2 A) + m(T \nabla A, \nabla A)\right] e^m A
\]
\[= -2\partial_{x_i} L,
\]
which gives (40). If \(\partial_{x_i} A\) is uniformly bounded by 1, then (69) clearly yields \(|\partial_{x_i} L| \leq 1.\]

Acknowledgments. The author thanks Arnab Sen and the anonymous referee for several suggestions regarding the presentation of the paper.

REFERENCES
[1] AUffinger, A. and CHen, W.-K. (2015). On properties of Parisi measures. Probab. Theory Related Fields 161 817–850. MR3334282
[2] AUffinger, A. and CHen, W.-K. (2015). The Parisi formula has a unique minimizer. Comm. Math. Phys. 335 1429–1444. MR3320318
[3] BOVIER, A. and KLIMOVSKY, A. (2009). The Aizenman–Sims–Starr and Guerra’s schemes for the SK model with multidimensional spins. Electron. J. Probab. 14 161–241. MR2471664
[4] BRAY, A. J. and MORRE, M. A. (1987). Chaotic nature of the spin-glass phase. *Phys. Rev. Lett.* **58** 57–60.

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

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

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

[8] FISHER, D. S. and HUSE, D. A. (1986). Ordered phase of short-range Ising spin-glasses. *Phys. Rev. Lett.* **56** 1601–1604.

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

[10] JAGANNATH, A. and TOBASCO, I. (2015). Some properties of the phase diagram for mixed \( p \)-spin glasses. Preprint. Available at arXiv:1504.02731.

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

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

[13] KRZĄKAŁA, F. and BOUCHAUD, J.-P. (2005). Disorder chaos in spin glasses. *Europhys. Lett.* **72** 472–478.

[14] PANCHENKO, D. (2008). On differentiability of the Parisi formula. *Electron. Commun. Probab.* **13** 241–247. MR2399285

[15] PANCHENKO, D. (2014). The Parisi formula for mixed \( p \)-spin models. *Ann. Probab.* **42** 946–958. MR3189062

[16] PARISI, G. (1979). Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.* **43** 1754–1756.

[17] RIZZO, T. (2009). Chaos in mean-field spin-glass models. In *Spin Glasses: Statics and Dynamics* (A. B. de Monvel and A. Bovier, eds.). *Progress in Probability* **62** 143–157. Birkhäuser, Basel. MR2761983

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

[19] TALAGRAND, M. (2011). *Mean Field Models for Spin Glasses. Volume I: Basic Examples*. Springer, Berlin. MR2731561

[20] TALAGRAND, M. (2011). *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*. Springer, Heidelberg. MR3024566

[21] TONINELLI, F. (2002). About the Almeida–Thouless transition line in the Sherrington–Kirkpatrick mean field spin glass model. *Europhys. Lett.* **60** 764–767.