

VECTOR SPIN GLASSES WITH MATTIS INTERACTION I: THE CONVEX CASE

HONG-BIN CHEN  AND VICTOR ISSA 

ABSTRACT. This paper constitutes the first part of a two-paper series devoted to the systematic study of vector spin glass models whose energy function involves a spin glass part and a general Mattis interaction part.

In this paper, we focus on models whose spin glass part satisfies the usual convexity assumption. We identify the limit free energy via a Parisi-type formula and prove a large deviation principle for the mean magnetization. The proof is remarkably simple and short compared to previous approaches; it relies on treating the Mattis interaction as a parameter of the model.

In the companion paper [14], we establish similar results in the high-temperature regime for models whose spin glass part is not assumed to satisfy the usual convexity assumption.

KEYWORDS AND PHRASES: DISORDERED SYSTEMS, LARGE DEVIATION PRINCIPLES, SPIN GLASSES.

MSC 2020: 82B44, 60F10, 35F21.

CONTENTS

1. Introduction	1
2. Identification of the limit free energy with linear Mattis interaction	9
3. Proofs of main results	13
References	15

1. INTRODUCTION

1.1. Preamble. This paper is the first of a two-part series in which we investigate general mean-field spin glass models with Mattis interaction. In this paper, we focus on models whose spin glass part satisfies the usual convexity assumption (1.2). A simple instance of those models is given by a

(Hong-Bin Chen) NYU-ECNU INSTITUTE OF MATHEMATICAL SCIENCES, NYU SHANGHAI, CHINA

(Victor Issa) DEPARTMENT OF MATHEMATICS, ENS LYON, FRANCE

E-mail addresses: hongbin.chen@nyu.edu, victor.issa@ens-lyon.fr.

system of ± 1 spins interacting through the energy function

$$(1.1) \quad E_N(\sigma) = \frac{\beta}{\sqrt{N}} \sum_{i,j=1}^N W_{ij} \sigma_i \sigma_j + \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \chi_i \sigma_i \right)^2, \quad \forall \sigma \in \{-1, 1\}^N,$$

where $(W_{i,j})_{1 \leq i,j \leq N}$ are independent centered gaussian random variables with variance 1 and $(\chi_i)_{1 \leq i \leq N}$ are centered ± 1 random variables. Our investigation is motivated in part by the study of statistical inference problems with mismatched prior and noise. In modern statistical inference, many high-dimensional estimation problems, such as recovering a low-rank signal from noisy observations, can be formulated in terms of energy-based models. In particular, when the assumed prior on the signal or the noise distribution used by the estimator does not match the true generative model, the resulting posterior landscape behaves like a generalized Sherrington–Kirkpatrick spin glass with additional Mattis interaction [3, 4, 37, 34, 16, 10]. Roughly speaking, the idea is that for inference problems with mismatched prior and noise, one can reduce the computation of the mutual information in the statistical problem to the computation of the free energy of a modified spin glass model similar to (1.1) and derive asymptotic formulas for the mutual information and overlaps (order parameters) that characterize estimator performance. This connection provides rigorous large deviation principles for key quantities, such as overlaps between the estimate and the true signal.

In this series of papers [13, 14], our interest is twofold. First, we prove variational formulas for the large N limit of the free energy defined by

$$f_N = -\frac{1}{N} \log \left(\frac{1}{2^N} \sum_{\sigma \in \{-1, 1\}^N} e^{E_N(\sigma)} \right).$$

Second, we establish large deviation principles for the mean magnetization

$$m_N = \frac{1}{N} \sum_{i=1}^N \chi_i \sigma_i$$

under the Gibbs measure $\langle \cdot \rangle_N$ defined by

$$\langle g(\sigma) \rangle_N \propto \frac{1}{2^N} \sum_{\sigma \in \{-1, 1\}^N} g(\sigma) e^{E_N(\sigma)}.$$

Our main results are stated for very general models in Theorem 1.2 and Theorem 1.4. Here we state a specialized version of those results for the model defined by (1.1), which is an instance of a model satisfying the convexity assumption (1.2) defined below. The definition of the function $\psi = \psi(q; x)$ appearing in this statement is given below in (1.9).

Theorem 1.1. *Almost surely over the randomness of $(W_{i,j})_{i,j}$ and $(\chi_i)_i$, we have*

$$\lim_{N \rightarrow +\infty} f_N = \inf_m \sup_{x,p} \left\{ \psi(\beta^2 p; x) - \frac{\beta^2}{2} \int_0^1 p(s)^2 ds + mx - m^2 - \frac{\beta^2}{2} \right\},$$

where the infimum is taken over $m \in \mathbb{R}$ and the supremum over $x \in \mathbb{R}$ and $p : [0, 1] \rightarrow \mathbb{R}_+$ bounded increasing càdlàg paths. Furthermore, still almost surely over the randomness of $(W_{i,j})_{i,j}$ and $(\chi_i)_i$, the random variable m_N satisfies a large deviation principle under $\langle \cdot \rangle_N$ with rate function

$$J(m) = -m^2 + \phi^*(m) + \sup_{m' \in \mathbb{R}} \{(m')^2 - \phi^*(m')\}.$$

Here $\phi^*(m) = \sup_{x \in \mathbb{R}} \{xm + \phi(x)\}$ and $\phi(x)$ is the almost sure limit as $N \rightarrow +\infty$ of the free energy f_N where the term $\left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \chi_i \sigma_i\right)^2$ in E_N is replaced by $x \sum_{i=1}^N \chi_i \sigma_i$.

Note that as a consequence of Theorem 1.2 below, the quantity $\phi(x)$ used to define the rate function J can be expressed as a variational formula.

For models satisfying the aforementioned convexity assumption (1.2), computing the free energy in the absence of Mattis interaction is a classical problem, and it is now well known that the limit of the free energy can be understood as the supremum of an explicit functional, this is the celebrated Parisi formula [33, 22, 36]. The free energy can also be studied from the perspective of Hamilton–Jacobi equations, which was initiated in [21] and followed by [6, 1, 8, 7, 5, 20]. Recently, a more systematic and general treatment through the prism of Hamilton–Jacobi equations was proposed in [27, 29, 18], the progress of which is summarized in [28].

On the other hand, models with Mattis interaction have not yet received a similar systematic treatment. The identification of their free energy relies on model-dependent proofs. Furthermore, especially in cases where authors are interested in large deviation principles for the mean magnetization, the proofs involve computing a constrained version of the free energy defined by

$$f_N(A) = -\frac{1}{N} \log \left(\frac{1}{2^N} \sum_{\sigma \in A} e^{E_N(\sigma)} \right).$$

The proofs can sometimes be long and technical [23, 19, 10, 32, 16]. Similar techniques have also been employed to prove large deviation principles for the overlap parameter in the context of classical spin glass models [9, 24], also see [25] for related content. For spin glass models with additional Mattis interaction, there are no known general results that allow for a systematic identification of the free energy, nor the verification of large deviation principles for the mean magnetization.

In this series of papers, we resolve this problem and establish systematic results using a new approach that relies on the computation of the following object

$$f_N^G = -\frac{1}{N} \log \left(\frac{1}{2^N} \sum_{\sigma \in \{-1, 1\}^N} \exp \left(\frac{\beta}{\sqrt{N}} \sum_{i,j=1}^N W_{ij} \sigma_i \sigma_j + NG(m_N) \right) \right).$$

The main difference in our approach is that the continuous function G encoding the Mattis interaction is now treated as a *parameter of the model*

instead of being fixed. Furthermore, as opposed to the traditional approach, we do not impose constraints on the spin configuration σ . This allows for smoother, shorter, and less technical computations that one can carry out in a model-independent way. Our main theorems cover many previous model-specific results. For example, this includes [23, Theorems 2.6 & 2.7], [10, Theorem 1], and [16, Theorem 1].

We start with the case $G(m) = xm$, where the sum structure of $NG(m_N) = x \sum_{i=1}^N \chi_i \sigma_i$ allows for the computation of $\phi(x) = \lim_{N \rightarrow +\infty} f_N^{m \rightarrow xm}$ using very classical arguments similar to those needed to establish the Parisi formula with essentially no modification. Then, we verify that ϕ is a continuously differentiable function of x . This ensures, through the Gärtner–Ellis theorem, that m_N satisfies a large deviation principle under $\langle \cdot \rangle_N^0$. Varadhan’s lemma then allows us to compute the large N limit of $\frac{1}{N} \log \langle e^{NG(m_N)} \rangle_N^0$ for all continuous functions G . From this, we deduce a large deviation principle for m_N under $\langle \cdot \rangle_N^G$ and an expression for $\lim_{N \rightarrow +\infty} f_N^G$. For the first part, this is a direct consequence of Bryc’s inverse Varadhan lemma, and the second part follows by observing that

$$\frac{1}{N} \log \langle e^{NG(m_N)} \rangle_N^0 = -f_N^G + f_N^0.$$

Once this is done, one can recover Theorem 1.1 by setting $G(m) = m^2$.

1.2. Setting. In this section, we give a definition of the class of models that we treat.

For $m, n \in \mathbb{N}$, we denote by $\mathbb{R}^{m \times n}$ the space of all $m \times n$ real matrices. For any $a = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathbb{R}^{m \times n}$, we denote its j -th column vector as $a_{\bullet j} = (a_{ij})_{1 \leq i \leq m}$ and its i -th row vector as $a_{i \bullet} = (a_{ij})_{1 \leq j \leq n}$. If not specified, a vector is understood to be a column vector. For a matrix or vector a , we denote by $a^\top$ its transpose. For $a, b \in \mathbb{R}^{m \times n}$, we write $a \cdot b = \sum_{ij} a_{ij} b_{ij}$, $|a| = \sqrt{a \cdot a}$, and similarly for vectors. For $n \in \mathbb{N}$, let S^n be the linear space of $n \times n$ real symmetric matrices and also let S_+^n (resp. S_{++}^n) be the subset consisting of positive semi-definite (resp. definite) matrices. For $a, b \in S^n$, we write $a \geq b$ provided $a - b \in S_+^n$, which gives a natural partial order on S^n .

We start by defining the spin glass part of the model. Let $D \in \mathbb{N}$ be the dimension of a single spin. Let P_1 be a finite positive measure supported on the unit ball $\{\tau \in \mathbb{R}^D : |\tau| \leq 1\}$ of $\mathbb{R}^D$ and we assume that the linear span of vectors in $\text{supp } P_1$ is $\mathbb{R}^D$.

For each $N \in \mathbb{N}$, we denote by $\sigma = (\sigma_{\bullet i})_{i=1}^N = (\sigma_{di})_{1 \leq d \leq D, 1 \leq i \leq N} \in \mathbb{R}^{D \times N}$ the spin configuration consisting of N $\mathbb{R}^D$ -valued spins $\sigma_{\bullet i}$ as column vectors. We view σ as a $D \times N$ matrix. We sample each spin independently from P_1 , and thus the distribution of σ is given by

$$dP_1^{\otimes N}(\sigma) = \otimes_{i=1}^N dP_1(\sigma_{\bullet i}).$$

We let $\xi : \mathbb{R}^{D \times D} \rightarrow \mathbb{R}$ be a smooth function that can be expressed as a power series, and we assume that

$$(1.2) \quad \text{the function } \xi \text{ is convex on } S_+^D.$$

For each N , we consider a centered Gaussian field $(H_N(\sigma))_{\sigma \in \mathbb{R}^{D \times N}}$ with covariance structure given by

$$(1.3) \quad \mathbb{E}[H_N(\sigma)H_N(\sigma')] = N\xi\left(\frac{\sigma\sigma'^\top}{N}\right), \quad \forall \sigma, \sigma' \in \mathbb{R}^{D \times N}.$$

We also define $\theta : \mathbb{R}^{D \times D} \rightarrow \mathbb{R}$ by

$$(1.4) \quad \theta(a) = a \cdot \nabla \xi(a) - \xi(a), \quad \forall a \in \mathbb{R}^{D \times D}.$$

We now describe the Mattis interaction part of the model. Let $L \in \mathbb{N}$ and let $(\chi_i)_{i \in \mathbb{N}}$ be i.i.d. $\mathbb{R}^L$ -valued random vectors, which are, in particular, independent from $H_N(\sigma)$. Let $d \in \mathbb{N}$ and $h : \mathbb{R}^D \times \mathbb{R}^L \rightarrow \mathbb{R}^d$ be a bounded¹ measurable function. For each $i \in \mathbb{N}$, we view $h(\sigma_{\bullet i}, \chi_i)$ as a generalized spin and define the mean magnetization

$$(1.5) \quad m_N = \frac{1}{N} \sum_{i=1}^N h(\sigma_{\bullet i}, \chi_i).$$

Let $G : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuous function, and we add to the system the Curie–Weiss-type interaction $NG(m_N)$.

At this point, we have described the spin glass part and the Mattis interaction part of our energy function. As explained in the introduction, recently a new approach has been put forward to describe the free energy of spin glass models (without Mattis interaction). In this approach, one needs to enrich the model by adding several new parameters. This more sophisticated object plays a key role in our companion paper [14], so we will work with this one in the present paper as well. In order to stick to the classical setting, the reader not interested in this enriched version of the model can throughout the paper choose the parameter q to be the constant path taking the value 0 and set $t = \beta^2/2$.

To define the enriched version of the model, we need to add external fields parameterized by increasing paths and driven by Poisson–Dirichlet cascades. Let $\mathcal{Q}_\infty$ be the collection of bounded increasing càdlàg paths $q : [0, 1] \rightarrow S_+^D$, where the monotonicity is understood as $q(s) - q(s') \in S_+^D$ for $s > s'$. Throughout, given $q \in \mathcal{Q}_\infty$, we will define

$$q(1) = \lim_{s \rightarrow 1^-} q(s),$$

which is well-defined by monotonicity of q .

Throughout, let $\mathfrak{R}$ be the Ruelle probability cascade (RPC) with overlap uniformly distributed over $[0, 1]$ (see [31, Theorem 2.17]). Precisely, $\mathfrak{R}$ is a random probability measure on the unit sphere of a fixed separable

¹Throughout it will be enough to assume that h is bounded on the support of the positive measure $P_1 \otimes \text{Law}(\chi_1)$.

Hilbert space (the exact form of the space is not important), with the inner product denoted by $\alpha \wedge \alpha'$. Let α and α' be independent samples from $\mathfrak{R}$. Then, the law of $\alpha \wedge \alpha'$ under $\mathbb{E}\mathfrak{R}^{\otimes 2}$ is the uniform distribution over $[0, 1]$, where $\mathbb{E}$ integrates the randomness of $\mathfrak{R}$. This overlap distribution uniquely determines $\mathfrak{R}$ (see [31, Theorem 2.13]). Almost surely, the support of $\mathfrak{R}$ is ultrametric in the induced topology. For rigorous definitions and basic properties, we refer to [31, Chapter 2] (also see [18, Chapter 5]). We also refer to [15, Section 4] for the construction and properties of $\mathfrak{R}$ useful in this work.

For $q \in \mathcal{Q}_\infty$ and almost every realization of $\mathfrak{R}$, let $(w^q(\alpha))_{\alpha \in \text{supp } \mathfrak{R}}$ be the $\mathbb{R}^D$ -valued centered Gaussian process with covariance

$$(1.6) \quad \mathbb{E}[w^q(\alpha)w^q(\alpha')^\top] = q(\alpha \wedge \alpha').$$

Its existence and properties are given in [15, Section 4]. Conditioned on $\mathfrak{R}$, let $(w_i^q)_{i \in \mathbb{N}}$ be i.i.d. copies of w^q and also independent of other randomness.

For $N \in \mathbb{N}$, $q \in \mathcal{Q}_\infty$, $t \geq 0$, the full energy function of the system will be chosen to be

$$(1.7) \quad \begin{aligned} H_N^G(\sigma, \alpha) = & NG(m_N) + \sqrt{2t}H_N(\sigma) - Nt\xi\left(\frac{\sigma\sigma^\top}{N}\right) \\ & + \sum_{i=1}^N w_i^q(\alpha) \cdot \sigma_{\bullet i} - \frac{1}{2}q(1) \cdot \sigma\sigma^\top. \end{aligned}$$

We are interested in the infinite-volume limit of the associated free energy defined as

$$(1.8) \quad F_N^G = -\frac{1}{N} \log \iint \exp(H_N^G(\sigma, \alpha)) dP_1^{\otimes N}(\sigma) d\mathfrak{R}(\alpha).$$

In principle, H_N^G and F_N^G depend on the choice of t and q . Throughout this paper, we will keep those parameters fixed so we do not make this dependence explicitly appear in our notations for H_N^G and F_N^G .

As is explained in full detail in Section 1.4, usual models such as (1.1) can be obtained as special cases of the Hamiltonian H_N^G . Briefly, to do so one can set $t = \beta^2/2$ and $q = 0$ and make a wise choice for G in order to absorb $-\frac{N\beta^2}{2}\xi\left(\frac{\sigma\sigma^\top}{N}\right)$ in the Mattis interaction.

Here, we have chosen to introduce the correction terms $Nt\xi(\sigma\sigma^\top/N)$ and $\frac{1}{2}q(1) \cdot \sigma\sigma^\top$ as a convenience to simplify the forthcoming variational formulas. As explained in Section 1.4, they are easily removed. Also, notice that those terms are one-half of the variance of $\sqrt{2t}H_N(\sigma)$ and $\sum_{i=1}^N w_i^q(\alpha) \cdot \sigma_{\bullet i}$, respectively. So they correspond to drifts in exponential martingales.

The large deviation principle we prove holds almost surely with respect to the randomness of $(H_N)_{N \geq 1}$, $(\chi_i)_{i \in \mathbb{N}}$, and $\mathfrak{R}$. Therefore, throughout we will assume that those random variables are defined on a common probability space Ω in order to have a convenient way of stating those results.

1.3. Main results. We let $\overline{F}_N^G = \mathbb{E} F_N^G$, where $\mathbb{E}$ averages over the randomness of H_N , $(\chi_i)_i$, and $\mathfrak{R}$. The random variable F_N^G concentrates strongly around its expectation $\overline{F}_N^G$ in the large N limit [31, Theorem 1.2]. Hence, to describe the limit as $N \rightarrow +\infty$ of F_N^G , it suffices to describe the limit of $\overline{F}_N^G$. To describe this limit, we need to define, for $q \in \mathcal{Q}_\infty$ and $x \in \mathbb{R}^d$,

$$(1.9) \quad \begin{aligned} & \psi(q; x) \\ &= -\mathbb{E} \log \iint \exp(w^q(\alpha) \cdot \tau - \tfrac{1}{2} q(1) \cdot \tau \tau^\top + x \cdot h(\tau, \chi_1)) dP_1(\tau) d\mathfrak{R}(\alpha), \end{aligned}$$

where $\mathbb{E}$ integrates the randomness of $\mathfrak{R}, (\chi_i)_i$ and H_N^G so that ψ is a deterministic function.

Theorem 1.2. *Recall that we assume (1.2). For every continuous function $G : \mathbb{R}^d \rightarrow \mathbb{R}$, we have*

$$(1.10) \quad \lim_{N \rightarrow \infty} \overline{F}_N^G = \inf_m \sup_{x, p} \left\{ \psi(q + t \nabla \xi(p); x) - \int_0^1 \theta(p(s)) ds + m \cdot x - G(m) \right\},$$

where $\inf$ is taken over $m \in \mathbb{R}^d$ and $\sup$ over $x \in \mathbb{R}^d$ and $p \in \mathcal{Q}_\infty$.

When $G = 0$, (1.10) boils down to the classical Parisi formula as written in [15, Proposition 8.4]. Note that here and in [15, Proposition 8.4] the convexity of ξ is assumed on S_+^D and not on the whole space $\mathbb{R}^{D \times D}$. This weaker convexity assumption should not be a surprise. Indeed, in the $D = 1$ case, it can be seen that the Parisi formula still holds assuming only that ξ is convex on $\mathbb{R}_+$ as a consequence of Talagrand's positivity principle, and this weaker convexity assumption is simply a consequence of a generalization of this observation (but with a different proof).

Throughout the paper, the free energy of models with linear Mattis interaction will play a special role. So for every $x \in \mathbb{R}^d$, we let

$$(1.11) \quad \varphi(x) = \lim_{N \rightarrow +\infty} \overline{F}_N^{m \rightarrow x \cdot m},$$

and we let $\varphi^*(m) = \sup_x \{x \cdot m + \varphi(x)\}$ denote the convex dual of $-\varphi$.

Definition 1.3 (Large deviation principle). Let $(X_N)_{N \geq 1}$ be a sequence of random variables in $\mathbb{R}^d$ with associated probability measures $(\mathbb{P}_N)_{N \geq 1}$ and let $I : \mathbb{R}^d \rightarrow [0, \infty]$. We say that the random variables X_N satisfy a large deviation principle under $\mathbb{P}_N$ with rate function I when for every Borel measurable set $A \subseteq \mathbb{R}^d$, we have

$$-\inf_{A^\circ} I \leq \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_N\{X_N \in A\} \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_N\{X_N \in A\} \leq -\inf_{\overline{A}} I,$$

where A° and $\overline{A}$ are the interior and the closure of A , respectively.

In the following, we refer to the large deviation principle as LDP.

Let $\langle \cdot \rangle_N^G$ be the probability measure defined by

$$\langle f(\sigma) \rangle_N^G \propto \int f(\sigma) e^{H_N^G(\sigma)} dP_1^{\otimes N}(\sigma).$$

Observe that $\langle \cdot \rangle_N^G$ is a *random* probability measure. Recall that the random variables $(H_N)_{N \geq 1}$ and $(\chi_i)_{i \in \mathbb{N}}$ are defined on a common probability space Ω . For a random variable $X : \Omega \rightarrow \mathfrak{X}$, we let $X^\omega \in \mathfrak{X}$ denote the realization of the random variable X on the event $\omega \in \Omega$. Recall the mean magnetization m_N defined in (1.5) and φ defined in (1.11). Our next main result is a (quenched) large deviation principle for the random variables m_N under $\langle \cdot \rangle_N^G$.

Theorem 1.4 (LDP for the magnetization). *Recall that we assume (1.2). There exists a full-measure subset $\Omega' \subseteq \Omega$ such that, for every continuous function $G : \mathbb{R}^d \rightarrow \mathbb{R}$ and every $\omega \in \Omega'$, the random variables m_N satisfy a large deviation principle under $\langle \cdot \rangle_N^{G, \omega}$ with rate function I^G defined by*

$$I^G(m) = -G(m) + \varphi^*(m) + \sup_{m' \in \mathbb{R}^d} \{G(m') - \varphi^*(m')\}, \quad \forall m \in \mathbb{R}^d.$$

1.4. Recovering concrete models from the general setting. The main results of this paper are stated for very general models as introduced in Section 1.2. Let us explain how to recover results about more usual models from those.

Given $h : \mathbb{R}^D \times \mathbb{R}^L \rightarrow \mathbb{R}^d$ in (1.8), we define $\tilde{h} : \mathbb{R}^D \times \mathbb{R}^L \rightarrow \mathbb{R}^d \times \mathbb{S}^D$ through

$$\tilde{h}(\tau, \chi) = (h(\tau, \chi), \tau \tau^\top).$$

Henceforth, we isometrically identify $\mathbb{R}^d \times \mathbb{S}^D$ with $\mathbb{R}^{\tilde{d}}$ for some $\tilde{d} \in \mathbb{N}$. Given $G : \mathbb{R}^d \rightarrow \mathbb{R}$ in (1.8), we introduce $\tilde{G} : \mathbb{R}^d \times \mathbb{S}^D \rightarrow \mathbb{R}$ given by

$$\tilde{G}(m, r) = G(m) + \frac{\beta^2}{2} \xi(r).$$

Then, we consider H_N^G in (1.7) with G, h, d , therein substituted with $\tilde{G}, \tilde{h}, \tilde{d}$ introduced above. Observe that setting $t = \beta^2/2$ and $q = 0$, we have

$$\begin{aligned} H_N^{\tilde{G}}(\sigma) &= \beta H_N(\sigma) + N \tilde{G} \left(\frac{1}{N} \sum_{i=1}^N \tilde{h}(\sigma_i, \chi_i) \right) - \frac{N \beta^2}{2} \xi \left(\frac{\sigma \sigma^\top}{N} \right) \\ &= \beta H_N(\sigma) + N G \left(\frac{1}{N} \sum_{i=1}^N h(\sigma_{\bullet i}, \chi_i) \right). \end{aligned}$$

In particular, observe that the model described in (1.1) is recovered by choosing $D = 1$, $d = 1$, $L = 1$, $h(\sigma, \chi) = \text{sgn}(\sigma \chi)$, $G(m) = m^2$, $\xi(x) = x^2$, and P_1 is the uniform measure on $\{-1, 1\}$. In particular observe that thanks to this, Theorem 1.1 can be recovered from Theorem 1.2 and Theorem 1.4.

1.5. Organization of the paper. In Section 2, we compute the limit free energy $\lim_{N \rightarrow +\infty} \overline{F}_N^G$ in the special case where G is of the form $G(m) = x \cdot m$ for some $x \in \mathbb{R}^d$ and prove that this quantity is a continuously differentiable function of x . In Section 3, we use the results of the previous section to give a proof of Theorem 1.2 and Theorem 1.4 by relying only on classical large deviation tools.

1.6. Acknowledgements. We would like to warmly thank Jean-Christophe Mourrat for many useful inputs and interesting discussions during the conception and writing of this paper. HBC acknowledges funding from the NYU Shanghai Start-Up Fund and support from the NYU–ECNU Institute of Mathematical Sciences at NYU Shanghai.

2. IDENTIFICATION OF THE LIMIT FREE ENERGY WITH LINEAR MATTIS INTERACTION

The plan to prove Theorem 1.2 is divided into two steps. First, we consider the free energy with $G(m) = x \cdot m$, namely, $F_N^{m \mapsto x \cdot m}$. In this case, we can think of $F_N^{m \mapsto x \cdot m}$ as adding some linear random external field to F_N^0 . For $F_N^{m \mapsto x \cdot m}$, since the linear field decouples easily in the cavity computation, the standard set of tools is effective with minimal modifications. Thus, the following result holds.

Proposition 2.1. *Recall that we assume (1.2), for every $x \in \mathbb{R}^d$ we have*

$$(2.1) \quad \lim_{N \rightarrow \infty} \overline{F}_N^{m \mapsto x \cdot m} = \sup_{p \in \mathcal{Q}_\infty} \left\{ \psi(q + t \nabla \xi(p); x) - t \int_0^1 \theta(p(s)) ds \right\}.$$

Observe that at $x = 0$, (2.1) reduces to the classical Parisi formula as stated in [15]. At $x \neq 0$ the proof (2.1) follows the same arguments than for the classical Parisi formula. The proof is done in two separate steps, first to prove that the limit free energy is lower-bounded by the right-hand side in (2.1), one can write a variant of the so-called of the Aizenman–Sims–Starr scheme. Roughly speaking, the principle of the computation is comparing $\overline{F}_{N+1}^{m \mapsto x \cdot m}$ with $\overline{F}_N^{m \mapsto x \cdot m}$ by “integrating out” one of the spin variables. This leads to expressions of the form

$$(N+1) \overline{F}_{N+1}^{m \mapsto x \cdot m} - N \overline{F}_N^{m \mapsto x \cdot m} \simeq \psi(q + t \nabla \xi(p); x) - t \int_0^1 \theta(p(s)) ds,$$

with some $p \in \mathcal{Q}_\infty$ possibly depending on N , which yields the desired bound. The second step of the proof consists in verifying that the limit free energy is upper-bounded by the right-hand side in (2.1). To proceed, it is possible to use Gaussian interpolation as in Guerra’s original proof of this bound for the classical Parisi formula [22]. Alternatively, one can also use the more modern point of view in [26, 29] which consists in verifying that the free energy at finite N is, as a function of (t, q) , a supersolution of a partial differential equation of Hamilton–Jacobi type with time parameter t and initial condition $q \mapsto \psi(q; x)$. The comparison principle then yields the desired bound.

Sketch of proof. Since the external field $Nx \cdot m_N$ is linear, the usual strategy — Guerra’s interpolation for the upper bound and the Aizenman–Sims–Starr scheme together with Panchenko’s ultrametricity for the lower bound — works without any significant modification. For completeness, we still sketch the key steps in the said strategy to highlight where the convexity is needed and where the linearity of the external field is used. After these are clarified, we refer to [11] for the detail for the lower and upper bounds below (but in the setting of vector spin glasses without random external field). For brevity, let us write $\overline{F}_N = \overline{F}_N^{m \mapsto x \cdot m}$.

Lower bound via Guerra’s interpolation [22]. Let us display the dependence on t in the Hamiltonian in (1.7) by writing $H_N^G(\sigma, \alpha; t)$. Fix any $p \in \mathcal{Q}_\infty$ and recall θ from (1.4). We consider an independent Gaussian field $(V(\alpha))_{\alpha \in \text{supp } \mathfrak{R}}$ with covariance

$$\mathbb{E}[V(\alpha)V(\alpha')] = \theta(p(\alpha \wedge \alpha')).$$

We also take an i.i.d. sequence of Gaussian field $(w_i^{t\nabla\xi(p)}(\alpha))_{\alpha \in \text{supp } \mathfrak{R}}$ indexed by $i \in \mathbb{N}$ with covariance given as in (1.6) with q therein replaced by $t\nabla\xi(p)$. For $r \in [0, 1]$, we consider the interpolating Hamiltonian

$$\begin{aligned} H_N(\sigma, \alpha; r) &= H_N^G(\sigma, \alpha; rt) + \sqrt{1-r} \sum_{i=1}^N w_i^{t\nabla\xi(p)}(\alpha) \cdot \sigma_{\bullet i} - \frac{1-r}{2} t \nabla\xi(p(1)) \cdot \sigma \sigma^\top \\ &\quad + \sqrt{rNt} V(\alpha) - \frac{rNt}{2} \theta(p(1)). \end{aligned}$$

With this Hamiltonian, we define the interpolating free energy

$$\phi(r) = -\frac{1}{N} \mathbb{E} \log \iint \exp(H_N(\sigma, \alpha; r)) dP_1^{\otimes N}(\sigma) d\mathfrak{R}(\alpha).$$

Then, we have

$$\phi(1) = \overline{F}_N - \frac{1}{N} \mathbb{E} \log \int \exp\left(\sqrt{Nt} V(\alpha) - \frac{Nt}{2} \theta(p(1))\right) d\mathfrak{R}(\alpha),$$

where the term $\frac{1}{N} \mathbb{E} \log \int \cdots d\mathfrak{R}(\alpha)$ can be computed to be $-t \int_0^1 \theta(p(s)) ds$ appearing in (2.1). Here, we used the independence between the Gaussian randomness of $V(\alpha)$ and that of $\sum_{i=1}^N w_i^q(\alpha) \cdot \sigma_{\bullet i}$ in $H_N^G(\sigma, \alpha; rt)$ together with some property of the cascade to decompose $\phi(1)$ into the two terms as in the above display. On the other hand, using $w_i^q(\alpha) + w_i^{t\nabla\xi(p)}(\alpha) \stackrel{d}{=} w_i^{q+t\nabla\xi(p)}(\alpha)$, we have

$$\begin{aligned} \phi(0) &= -\frac{1}{N} \mathbb{E} \log \iint \exp\left(\sum_{i=1}^N x \cdot h(\sigma_{\bullet i}, \chi_i) + \sum_{i=1}^N w_i^{q+t\nabla\xi(p)}(\alpha) \cdot \sigma_{\bullet i} \right. \\ &\quad \left. - \frac{1}{2}(q + t\nabla\xi(p))(1) \cdot \sigma \sigma^\top\right) dP_1^{\otimes N}(\sigma) d\mathfrak{R}(\alpha), \end{aligned}$$

which is equal to $\psi(q; x)$ in (1.9) using the independence of random variables indexed by i . Notice that the linearity of the external field was used here.

Then, the upper bound in (2.1) follows once $\phi(1) \geq \phi(0)$. To show this, we compute the derivative

$$(2.2) \quad \frac{d}{dr} \phi(r) = \frac{1}{2} \mathbb{E} \left\langle \xi(\sigma \sigma'^\top) - \nabla \xi(p(\alpha \wedge \alpha')) \cdot \sigma \sigma'^\top + \theta(p(\alpha \wedge \alpha')) \right\rangle_r,$$

where we used the standard tool of Gaussian integration by parts and $\langle \cdot \rangle_r$ is the Gibbs measure associated with $H_N(\sigma, \alpha; r)$. Notice that in the above display there are no terms involving self-overlaps $\sigma \sigma^\top$ or $\alpha \wedge \alpha = 1$, which might otherwise arise from Gaussian integration by parts. This is because such terms are exactly canceled by the derivatives of $-\frac{rNt}{2} \xi(\frac{\sigma \sigma^\top}{N})$, $-\frac{1-r}{2} t \nabla \xi(p(1)) \cdot \sigma \sigma^\top$, and $-\frac{rNt}{2} \theta(p(1))$ appearing in the Hamiltonian. The definition of θ in (1.4) and the convexity of ξ on $\mathbb{R}^{D \times D}$ ensures that $\xi(a) - \nabla \xi(b) \cdot a - \theta(b) \geq 0$ for every $a, b \in \mathbb{R}^{D \times D}$. Therefore, we get $\frac{d}{dr} \phi(r) \geq 0$ and completes the proof of the lower bound.

Upper bound via Aizenman–Sims–Starr scheme [2] and Panchenko’s ultrametricity [30]. Let $(Z(\sigma))_{\sigma \in \mathbb{R}^{D \times N}}$ and $(Y(\sigma))_{\sigma \in \mathbb{R}^{D \times N}}$ be independent centered $\mathbb{R}^D$ -valued and real-valued Gaussian processes with covariances

$$\mathbb{E}[Z(\sigma)Z(\sigma')^\top] = \nabla \xi\left(\frac{\sigma \sigma'^\top}{N}\right) \quad \text{and} \quad \mathbb{E}[Y(\sigma)Y(\sigma')] = \theta\left(\frac{\sigma \sigma'^\top}{N}\right).$$

We can compute that

$$(N+1)\overline{F}_{N+1} - N\overline{F}_N = A_N^{(1)} - A_N^{(2)} + o(1),$$

with an error term $o(1)$ that vanishes as $N \rightarrow \infty$. Here,

$$\begin{aligned} A_N^{(1)} = & -\mathbb{E} \log \left\langle \int \exp \left(\left(\sqrt{2t}Z(\sigma) + w_{N+1}^q(\alpha) \right) \cdot \tau - \frac{1}{2} \left(2t \nabla \xi \left(\frac{\sigma \sigma^\top}{N} \right) + q(1) \right) \cdot \tau \tau^\top \right. \right. \\ & \left. \left. + h(\tau, \chi_{N+1}) \right) dP_1(\tau) \right\rangle_N \quad \text{and} \quad A_N^{(2)} = -\mathbb{E} \log \left\langle \exp \left(\sqrt{2t}Y(\sigma) - t\theta \left(\frac{\sigma \sigma^\top}{N} \right) \right) \right\rangle_N, \end{aligned}$$

where $\tau = \sigma_{\bullet, N+1}$ is called the cavity spin and $\langle \cdot \rangle_N$ is the Gibbs measure associated with $\overline{F}_N$. Notice that $w_{N+1}^q(\alpha)$ and $h(\tau, \chi_{N+1})$ have randomness independent from that of $\langle \cdot \rangle_N$. It is important to note that we used the linearity $(N+1)x \cdot m_{N+1} = Nx \cdot m_N + h(\tau, \chi_{N+1})$ in this computation, where $Nx \cdot m_N$ is hidden in $\langle \cdot \rangle_N$ and $h(\tau, \chi_{N+1})$ is displayed in the above.

We can see $\limsup_{N \rightarrow \infty} \overline{F}_N \leq \limsup_{N \rightarrow \infty} A_N^{(1)} - A_N^{(2)}$. Heuristically, by passing to a subsequence, we may assume that the distribution of the spin overlap $\frac{\sigma \sigma'^\top}{N}$ under $\mathbb{E} \langle \cdot \rangle_N$ converges to that of $p(U)$ for some $p \in \mathcal{Q}_\infty$ where U is a uniform random variable over $[0, 1]$. Also, we can *synchronize* the spin overlap with the cascade overlap of α as spelled out in [32, Theorem 4]. Panchenko’s ultrametricity result along with Ghirlanda–Guerra’s identities tells us that this is enough to identify the limit of $A_N^{(1)}$ and $A_N^{(2)}$, which are $\psi(q + t \nabla \xi(p); x)$ and $t \int_0^1 \theta(p(s)) ds$, respectively.

This completes the lower bound modulo many technical results. For instance, we need to add some perturbation to $\langle \cdot \rangle_N$ to ensure that the above sketch works.

Relaxation of convexity of ξ . In view of the above proof, the convexity of ξ on $\mathbb{R}^{D \times D}$ is only used to ensure that the right-hand side of (2.2) is non-positive. When $D = 1$, Talagrand's positivity principle allows one to relax this assumption and require ξ to be convex only on $\mathbb{R}_+$. When $D > 1$, one can verify that [15, Proposition 8.1] remains valid in the presence of the random external field $Nx \cdot m_N$, which yields Proposition 2.1 with ξ assumed to be convex only on S_+^D .

Indeed, the proof of [15, Proposition 8.1] is based on cavity computations (similar to those used in the above proof of the upper bound) developed in [15], together with an input from [29] that provides an lower bound (in place of Guerra's bound) via the solution of a relevant Hamilton–Jacobi equation. The variational representation of this solution is precisely the right-hand side of (2.1). In both [15] and [29], the analysis is carried out for vector spin glass models without a random external field. However, as illustrated in the sketch above, the inclusion of such a field does not require any substantial modification of the arguments. $\square$

Recall that in (1.11) we have defined $\varphi(x) = \lim_{N \rightarrow \infty} \overline{F}_N^{m \mapsto x \cdot m}$. It follows from Proposition 2.1 that

$$(2.3) \quad \varphi(x) = \sup_{p \in \mathcal{Q}_\infty} \left\{ \psi(q + t \nabla \xi(p); x) - t \int_0^1 \theta(p(s)) ds \right\}.$$

As explained in the following lemma, φ is a well-behaved function of x . This follows from similar arguments as the ones for [12, Lemma 2.2].

Lemma 2.2. *The function φ is Lipschitz, concave, and continuously differentiable.*

Proof. We denote by $\mathcal{P}(p, x)$ the expression inside $\sup_{p \in \mathcal{Q}_\infty} \{\dots\}$ in (2.3), so that

$$\varphi(x) = \sup_{p \in \mathcal{Q}_\infty} \mathcal{P}(p, x).$$

Furthermore, we have $\nabla_x \overline{F}_N^{m \mapsto x \cdot m} = -\mathbb{E} \langle m_N \rangle_N^{m \mapsto x \cdot m}$, which is bounded uniformly in N . For any $v \in \mathbb{R}^d$, we can compute

$$(v \cdot \nabla_x)^2 \overline{F}_N^{m \mapsto x \cdot m} = -\mathbb{E} \langle (v \cdot m_N)^2 \rangle_N^{m \mapsto x \cdot m} + (\mathbb{E} \langle v \cdot m_N \rangle_N^{m \mapsto x \cdot m})^2 \leq 0.$$

Therefore, φ is Lipschitz and concave.

It is classical that the combination of convexity and differentiability implies continuous differentiability (e.g. [35, Theorem 25.5]). Hence, it only remains to show that φ is differentiable everywhere. Fix any $x \in \mathbb{R}^d$. A vector $a \in \mathbb{R}^d$ is said to be a superdifferential of φ at x when for every $y \in \mathbb{R}^d$,

$$\varphi(y) - \varphi(x) \leq a \cdot (y - x).$$

Since φ is concave, it suffices to show that any superdifferential a of φ at x is unique. For each $\varepsilon > 0$, we choose p_ε to satisfy

$$(2.4) \quad \mathcal{P}(p_\varepsilon, x) \geq \varphi(x) - \varepsilon.$$

Fix any $y \in \mathbb{R}^d$ and let $r \in (0, 1]$. Using the definition of the superdifferential, the fact that φ is a supremum, and (2.4), we have

$$\begin{aligned} y \cdot a &\geq \frac{\varphi(x + ry) - \varphi(x)}{r} \geq \frac{\mathcal{P}(p_\varepsilon, x + ry) - \mathcal{P}(p_\varepsilon, x) - \varepsilon}{r}, \\ y \cdot a &\leq \frac{\varphi(x) - \varphi(x - ry)}{r} \leq \frac{\mathcal{P}(p_\varepsilon, x) - \mathcal{P}(p_\varepsilon, x - ry) + \varepsilon}{r}. \end{aligned}$$

Using the expression in (1.9), we can compute derivatives of $\mathcal{P}(p_\varepsilon, \cdot)$ at x and verify that they are bounded. Hence, we can use Taylor's expansion of $\mathcal{P}(p_\varepsilon, \cdot)$ at x in the above display to get for some constant C independent of r and ε , we have

$$|y \cdot a - y \cdot \nabla \mathcal{P}(p_\varepsilon, x)| \leq Cr + \varepsilon r^{-1}.$$

Setting $r = \sqrt{\varepsilon}$ and sending $\varepsilon \rightarrow 0$, we can see that $y \cdot a$ is uniquely determined. Since y is arbitrary, we conclude that the subdifferential of φ at x is unique, and thus φ is differentiable at x . As explained earlier, this implies that φ is continuously differentiable. $\square$

Remark 2.3. Since φ is Lipschitz we have $\varphi^*(m) = +\infty$ for $|m| > L$ where L is a Lipschitz constant of φ .

3. PROOFS OF MAIN RESULTS

In this section, we provide a proof of our main results Theorem 1.2 and Theorem 1.4. We will rely on the differentiability of $\varphi(x) = \lim_{N \rightarrow +\infty} \overline{F}_N^{m \mapsto x \cdot m}$ as well as classical large deviation tools. Namely we will use the Gärtner–Ellis theorem [17, Theorem 2.3.6], Varadhan's lemma [17, Theorem 4.3.1], and Bryc's inverse Varadhan lemma [17, Theorem 4.4.13].

Lemma 3.1. *There is a full-measure subset $\Omega' \subseteq \Omega$ such that, for every $\omega \in \Omega'$, m_N under $\langle \cdot \rangle_N^{0, \omega}$ satisfies large deviation principle with rate function I defined by*

$$(3.1) \quad I(m) = \varphi^*(m) - \varphi(0), \quad \forall m \in \mathbb{R}^d.$$

Proof. Let $\Omega' \subseteq \Omega$ be the full-measure subset such that for every $\omega \in \Omega'$, we have for all $y \in \mathbb{Q}^d$,

$$\lim_{N \rightarrow \infty} F_N^{m \mapsto m \cdot y, \omega} = \varphi(y).$$

Henceforth, we fix $\omega \in \Omega'$ and define for each $y \in \mathbb{R}^d$,

$$\Lambda_N(y) = \frac{1}{N} \log \langle e^{Ny \cdot m_N} \rangle_N^{0, \omega}.$$

Notice that $\Lambda_N(y) = -F_N^{m \rightarrow y, m, \omega} + F_N^{0, \omega}$. Hence, we have for every $y \in \mathbb{Q}^d$

$$\lim_{N \rightarrow \infty} \Lambda_N(y) = -\varphi(y) + \varphi(0).$$

Since Λ_N is Lipschitz uniformly in N and φ is also Lipschitz, the previous display in fact holds for every $y \in \mathbb{R}^d$. Finally, since by Lemma 2.2, φ is differentiable everywhere, the Gärtner–Ellis theorem (e.g., see [17, Theorem 2.3.6]) guarantees that m_N satisfies a large deviation principle under $\langle \cdot \rangle_N^{0, \omega}$ with rate function

$$I(m) = \sup_{y \in \mathbb{R}^d} \left\{ y \cdot m - \lim_{N \rightarrow +\infty} \Lambda_N(y) \right\} = \varphi^*(m) - \varphi(0).$$

□

Lemma 3.2. *For every continuous function $G : \mathbb{R}^d \rightarrow \mathbb{R}$ and every $\omega \in \Omega'$ (given as in Lemma 3.1), we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \left\langle e^{NG(m_N)} \right\rangle_N^{0, \omega} = \sup_{m \in \mathbb{R}^d} \{G(m) - I(m)\},$$

with I given as in (3.1).

Proof. Since m_N is bounded uniformly in N , so is $G(m_N)$, which allows us to apply Varadhan’s lemma (e.g., see [17, Theorem 4.3.1]) to deduce the desired result from the large deviation principle in Lemma 3.1. □

Using Lemma 3.2 in conjunction with Bryc’s inverse Varadhan lemma, we obtain a proof of Theorem 1.4.

Proof of Theorem 1.4. Fix any G and let $\Gamma : \mathbb{R}^d \rightarrow \mathbb{R}$ be any bounded continuous function. Since

$$\frac{1}{N} \log \left\langle e^{N\Gamma(m_N)} \right\rangle_N^{G, \omega} = \frac{1}{N} \log \left\langle e^{N(G+\Gamma)(m_N)} \right\rangle_N^{0, \omega} - \frac{1}{N} \log \left\langle e^{NG(m_N)} \right\rangle_N^{0, \omega},$$

Lemma 3.2 implies that for every $\omega \in \Omega'$ we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \left\langle e^{N\Gamma(m_N)} \right\rangle_N^{G, \omega} = \sup \{G + \Gamma - \varphi^*\} - \sup \{G - \varphi^*\} = \sup \{\Gamma - I^G\}.$$

Since this holds for any bounded continuous Γ , we can apply Bryc’s inverse Varadhan lemma (e.g., see the “ $\Leftarrow$ ” direction in [17, Theorem 4.4.13]) which guarantees that m_N satisfies a large deviation principle under $\langle \cdot \rangle_N^{G, \omega}$ with rate function I^G . □

From Lemma 3.2, we can derive the Parisi-type formula for $\lim_{N \rightarrow +\infty} F_N^G$.

Proof of Theorem 1.2. Notice that

$$\frac{1}{N} \log \left\langle e^{NG(m_N)} \right\rangle_N^{0, \omega} = -F_N^{G, \omega} + F_N^{0, \omega}.$$

Using Lemma 3.2, we get that for every $\omega \in \Omega'$

$$\lim_{N \rightarrow \infty} F_N^{G, \omega} = \inf_{m \in \mathbb{R}^d} \{\varphi^*(m) - G(m)\}.$$

Finally, using the variational formula for $\varphi(x)$ of Proposition 2.1, we observe that the right-hand side in the previous display is equal to the right-hand side in (1.10). $\square$

REFERENCES

- [1] Elena Agliari, Adriano Barra, Raffaella Burioni, and Aldo Di Biasio. Notes on the p-spin glass studied via Hamilton-Jacobi and smooth-cavity techniques. *J. Math. Phys.*, 53(6):063304, 29, 2012.
- [2] Michael Aizenman, Robert Sims, and Shannon L Starr. Extended variational principle for the Sherrington–Kirkpatrick spin-glass model. *Physical Review B*, 68(21):214403, 2003.
- [3] Jean Barbier, Wei-Kuo Chen, Dmitry Panchenko, and Manuel Sáenz. Performance of Bayesian linear regression in a model with mismatch. *Information and Inference: A Journal of the IMA*, 14(3), 2025.
- [4] Jean Barbier, TianQi Hou, Marco Mondelli, and Manuel Sáenz. The price of ignorance: how much does it cost to forget noise structure in low-rank matrix estimation? In *Advances in Neural Information Processing Systems*, volume 35, pages 36733–36747, 2022.
- [5] Adriano Barra. The mean field Ising model through interpolating techniques. *J. Stat. Phys.*, 132:787–809, 2008.
- [6] Adriano Barra, Gino Del Ferraro, and Daniele Tantari. Mean field spin glasses treated with PDE techniques. *Eur. Phys. J. B*, 86(7):Art. 332, 10, 2013.
- [7] Adriano Barra, Andrea Di Lorenzo, Francesco Guerra, and Antonio Moro. On quantum and relativistic mechanical analogues in mean-field spin models. *Proc. R. Soc. A: Math. Phys. Eng. Sci.*, 470(2172):20140589, 2014.
- [8] Adriano Barra, Giuseppe Genovese, and Francesco Guerra. Equilibrium statistical mechanics of bipartite spin systems. *J. Phys. A*, 44(24):245002, 22, 2011.
- [9] Gérard Ben Arous and Aukosh Jagannath. Spectral gap estimates in mean field spin glasses. *Communications in Mathematical Physics*, 361(1):1–52, 2018.
- [10] Francesco Camilli, Pierluigi Contucci, and Emanuele Mingione. An inference problem in a mismatched setting: a spin-glass model with Mattis interaction. *SciPost Physics*, 12(4):125, 2022.
- [11] Hong-Bin Chen. Self-overlap correction simplifies the Parisi formula for vector spins. *arXiv preprint arXiv:2303.16284*, 2023.
- [12] Hong-Bin Chen. Free energy in spin glass models with conventional order. *Journal of Statistical Physics*, 191(4):49, 2024.
- [13] Hong-Bin Chen and Victor Issa. Vector spin glasses with Mattis interaction I: the convex case. *arXiv preprint*, 2026.
- [14] Hong-Bin Chen and Victor Issa. Vector spin glasses with Mattis interaction II: non-convex high-temperature models. *arXiv preprint*, 2026.
- [15] Hong-Bin Chen and Jean-Christophe Mourrat. On the free energy of vector spin glasses with non-convex interactions. *Prob. Math. Phys.*, to appear, 2023.
- [16] Wei-Kuo Chen. On the mixed even-spin Sherrington–Kirkpatrick model with ferromagnetic interaction. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 50(1):63–83, 2014.
- [17] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38. Springer Science & Business Media, 2009.
- [18] Tomas Dominguez and Jean-Christophe Mourrat. *Statistical mechanics of mean-field disordered systems: a Hamilton-Jacobi approach*. Zurich Lectures in Advanced Mathematics. EMS Press, 2024.
- [19] Silvio Franz and Jacopo Rocchi. Large deviations of glassy effective potentials. *Journal of Physics A: Mathematical and Theoretical*, 53(48):485002, 2020.

- [20] Giuseppe Genovese and Adriano Barra. A mechanical approach to mean field spin models. *Journal of Mathematical Physics*, 50(5):053303, 2009.
- [21] Francesco Guerra. Sum rules for the free energy in the mean field spin glass model. *Fields Institute Communications*, 30(11), 2001.
- [22] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.
- [23] Alice Guionnet, Justin Ko, Florent Krzakala, and Lenka Zdeborová. Estimating rank-one matrices with mismatched prior and noise: universality and large deviations. *Communications in Mathematical Physics*, 406(1):9, 2025.
- [24] Aukosh Jagannath, Patrick Lopatto, and Léo Miolane. Statistical thresholds for tensor PCA. *The Annals of Applied Probability*, 30(4):1910–1933, 2020.
- [25] Justin Ko. Free energy of multiple systems of spherical spin glasses with constrained overlaps. *Electronic Journal of Probability*, 25:1–34, 2020.
- [26] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.
- [27] Jean-Christophe Mourrat. The Parisi formula is a Hamilton-Jacobi equation in Wasserstein space. *Canad. J. Math.*, 74(3):607–629, 2022.
- [28] Jean-Christophe Mourrat. Spin glasses and the Parisi formula. *arXiv preprint arXiv:2510.01054*, 2025.
- [29] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *arXiv preprint arXiv:2010.09114*, 2020.
- [30] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.
- [31] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, 2013.
- [32] Dmitry Panchenko. Free energy in the mixed p -spin models with vector spins. *Ann. Probab.*, 46(2):865–896, 2018.
- [33] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754, 1979.
- [34] Farzad Pourkamali and Nicolas Macris. Mismatched estimation of non-symmetric rank-one matrices under Gaussian noise. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 1288–1293. IEEE, 2022.
- [35] R Tyrrell Rockafellar. *Convex Analysis*, volume 36. Princeton university press, 1970.
- [36] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.
- [37] Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: Thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.