



Article

The Onset of Parisi's Complexity in a Mismatched Inference Problem

Francesco Camilli ^{1,†}, Pierluigi Contucci ^{2,†}  and Emanuele Mingione ^{2,*,†} 

¹ The Abdus Salam International Center for Theoretical Physics, 34151 Trieste, Italy; fcamilli@ictp.it

² Dipartimento di Matematica, Alma Mater Studiorum—Università di Bologna, 40127 Bologna, Italy; pierluigi.contucci@unibo.it

* Correspondence: emanuele.mingione2@unibo.it

† These authors contributed equally to this work.

Abstract: We show that a statistical mechanics model where both the Sherrington–Kirkpatrick and Hopfield Hamiltonians appear, which is equivalent to a high-dimensional mismatched inference problem, is described by a replica symmetry-breaking Parisi solution.

Keywords: statistical mechanics; spin glasses; high-dimensional inference; replica symmetry breaking

1. Introduction

Beginning with Parisi's seminal works on the Sherrington–Kirkpatrick (SK) model [1,2], the ideas and the tools developed in spin glass theory spread across many other research fields such as computer science, probability theory, neural networks and others [3–7]. From a mathematical perspective, efforts to rigorously prove Parisi's theory have yielded powerful techniques, such as interpolation methods [8,9], stochastic stability [10], and synchronization [11], which are currently instrumental in analyzing numerous disordered systems.

In this work, we will consider a family of mean-field spin glasses whose Hamiltonian contains two types of random interactions: the first is the SK type, while the second is a Hopfield model with a finite number of patterns. This class of models can also be interpreted as a high-dimensional inference problem known as a spiked Wigner model in a *mismatched* setting [12–16].

Our main result is a representation of the thermodynamic limit for the quenched pressure per particle in terms of a variational problem of Parisi type. The proof relies on two main ingredients: Guerra's replica symmetry-breaking bound, which allows controlling the SK contribution, and adaptive interpolation, which is employed to linearize the Hopfield interaction. We start with a review of the SK model in Section 2 and then lay the ground for the inferential interpretation of the model under study in Section 3. In Section 4, we define and rigorously identify the exact solution of the model. Finally, we describe some interesting challenges for future investigations.

2. The SK Model

The SK model was introduced in the 1970s by D. Sherrington and S. Kirkpatrick [17] and stands as an explicitly solvable mean-field spin glass. In their work, the authors discovered that the solution obtained through the replica symmetric (RS) approximation was not correct at low temperature. With a groundbreaking approach, Parisi identified a new type of solutions, nowadays called *replica symmetry breaking* (RSB), which proved to be correct at any temperature, thereby revealing a novel mathematical and physical structure [18].



Citation: Camilli, F.; Contucci, P.; Mingione, E. The Onset of Parisi's Complexity in a Mismatched Inference Problem. *Entropy* **2024**, *26*, 42. <https://doi.org/10.3390/e26010042>

Academic Editor: Giuseppe Florio

Received: 27 November 2023

Revised: 21 December 2023

Accepted: 25 December 2023

Published: 30 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

The SK model is defined by its Hamiltonian, that is a function of N spins $\sigma = (\sigma_i)_{i \leq N} \in \{-1, 1\}^N$:

$$H_N^{SK}(\sigma) = -\frac{1}{\sqrt{2N}} \sum_{i,j \leq N} z_{ij} \sigma_i \sigma_j \tag{1}$$

where $\mathbf{z} = (z_{ij})_{i,j \leq N}$ is a collection of i.i.d. standard Gaussian. In physical terms, the couplings between pairs of spins can be ferromagnetic or antiferromagnetic with equal probability. Consider also a random variable ζ with $\mathbb{E}|\zeta| < \infty$ and a collection $\boldsymbol{\zeta} = (\zeta_i)_{i \leq N} \stackrel{\text{iid}}{\sim} \zeta$ representing random external fields acting on the spins. The *Parisi formula* is a representation for the large N limit of the pressure p_N^{SK} defined by

$$p_N^{SK}(\beta, h) = \frac{1}{N} \log \sum_{\sigma \in \{-1, 1\}^N} \exp\left(-\beta H_N^{SK}(\sigma) + h \sigma \cdot \boldsymbol{\zeta}\right) \tag{2}$$

In the definition (2), $(\beta, h) \in \mathbb{R}_{>0} \times \mathbb{R}$ are fixed parameters, and the dependence on the realization of the random collections $\mathbf{z}, \boldsymbol{\zeta}$ is kept implicit. One can prove [5] that p_N^{SK} converges, for almost all realizations of the disorder, to its average $\bar{p}_N^{SK}(\beta, h) = \mathbb{E} p_N^{SK}(\beta, h)$. Notice that \mathbb{E} , taken after the logarithm, averages both the collections of \mathbf{z} and $\boldsymbol{\zeta}$ that are called *quenched variables*. The Hamiltonian (1) can also be regarded as a centered Gaussian process with covariance

$$\mathbb{E} H_N^{SK}(\sigma^1) H_N^{SK}(\sigma^2) = \frac{N}{2} q_N^2(\sigma^1, \sigma^2) \tag{3}$$

where

$$q_N(\sigma^1, \sigma^2) = \frac{1}{N} \sum_{i=1}^N \sigma_i^1 \sigma_i^2 = \frac{1}{N} \sigma^1 \cdot \sigma^2. \tag{4}$$

$q_N(\sigma^1, \sigma^2)$ is the *overlap* between two spin configurations σ^1 and σ^2 .

The Parisi variational principle for the limiting pressure per particle of this model was proved after almost three decades of efforts, and it is mainly due to the works of Guerra [8] and Talagrand [19]. We hereby summarize these milestones in a single theorem.

Theorem 1 (Parisi Formula [8,19]). *Let $\mathcal{M}_{[0,1]}$ be the space of probability measures on $[0, 1]$, $\beta > 0$ and $y \in \mathbb{R}$. Consider the Parisi functional, which is defined as*

$$\chi \in \mathcal{M}_{[0,1]} \mapsto \mathcal{P}(\chi; \beta, y) = \log 2 + \Phi_\chi(0, y, \beta) - \frac{\beta^2}{2} \int_0^1 dq q \chi([0, q]). \tag{5}$$

where $\Phi_\chi(s, y, \beta)$ solves the PDE

$$\begin{cases} \partial_s \Phi_\chi = -\frac{\beta^2}{2} \left(\partial_y^2 \Phi_\chi + \chi([0, s]) (\partial_y \Phi_\chi)^2 \right) \\ \Phi_\chi(1, y, \beta) = \log \cosh y. \end{cases} \tag{6}$$

The following holds

$$\lim_{N \rightarrow \infty} p_N^{SK}(\beta, h) = \inf_{\chi \in \mathcal{M}_{[0,1]}} \mathbb{E}_\zeta \mathcal{P}(\chi; \beta, h \zeta) \quad \text{a.s.} \tag{7}$$

The key tool for the proof is the (Gaussian) *interpolation method*, which is introduced in [9] in order to prove the existence of the large N limit of \bar{p}_N^{SK} .

The thermodynamic equilibrium induced by the pressure p_N^{SK} is called *quenched equilibrium* and is defined as follows. Physical quantities (e.g., energy) are functions of

the disorder variables $\mathbf{z}, \boldsymbol{\zeta}$ and the spin configurations σ . Given a function $f(\mathbf{z}, \boldsymbol{\zeta}, \sigma)$, its equilibrium value is defined as

$$\mathbb{E}\langle f \rangle_N := \mathbb{E} \sum_{\sigma \in \{-1,1\}^N} G_N(\sigma) f(\mathbf{z}, \boldsymbol{\zeta}, \sigma) \tag{8}$$

where G_N is the (random) Boltzmann–Gibbs distribution

$$G_N(\sigma) = \frac{\exp(-\beta H_N^{SK}(\sigma) + h \sigma \cdot \boldsymbol{\zeta})}{Z_N} \tag{9}$$

The measure $\mathbb{E}\langle \cdot \rangle_N$ is called a *quenched measure* and can be viewed as a two-step measuring process. Initially, for a given realization of the disorder variables $\mathbf{z}, \boldsymbol{\zeta}$, one assumes that the system equilibrates according to the canonical Boltzmann–Gibbs distribution G_N defining a (random) measure on the space of spin configurations. The expectation with respect to G_N is denoted by $\langle \cdot \rangle_N$, namely

$$\langle f \rangle_N := \sum_{\sigma \in \{-1,1\}^N} G_N(\sigma) f(\mathbf{z}, \boldsymbol{\zeta}, \sigma) \tag{10}$$

In probabilistic terms, G_N defines a conditional measure given \mathbf{z} and $\boldsymbol{\zeta}$. The remaining degrees of freedom $\mathbf{z}, \boldsymbol{\zeta}$ are then averaged according to their apriori distribution \mathbb{E} .

An important role is played by the concept of *replicas*. Replicas are i.i.d. samples from G_N at fixed disorder. Hence, the equilibrium value of a function $f(\mathbf{z}, \boldsymbol{\zeta}, \sigma^1, \dots, \sigma^n)$ of n replicas and the quenched variables $\mathbf{z}, \boldsymbol{\zeta}$ is defined by

$$\mathbb{E}\langle f \rangle_N = \mathbb{E} \prod_{a \leq n} \sum_{\sigma^a \in \{-1,1\}^N} G_N(\sigma^a) f(\mathbf{z}, \boldsymbol{\zeta}, \sigma^1, \dots, \sigma^n). \tag{11}$$

The computation of derivatives of \bar{p}_N^{SK} shows, using integration by parts, that the SK model is fully characterized by the (joint) distribution of the overlap array $(q_N(\sigma^l, \sigma^{l'}))_{l,l' \leq n} \equiv (q_{l,l'})_{l,l' \leq n}$, namely the overlaps between any finite number n of replicas with respect to the measure (11). The main feature of the Parisi theory is the characterization of the mentioned joint measure by means of two structural properties:

- (i) It is uniquely determined by a one-dimensional marginal, namely the distribution of $q_{1,2}$;
- (ii) The distribution of three replicas has with a probability of one an *ultrametric* support

$$\lim_{N \rightarrow \infty} \mathbb{E}\langle \mathbb{1}(q_{1,2} \geq \min(q_{1,3}, q_{2,3})) \rangle = 1. \tag{12}$$

Despite having a mathematical proof of the Parisi Formula (7) for the SK model, (i) and (ii) have been rigorously proved only in the mixed p -spin model [6,20,21], an extension of the SK model, whose Hamiltonian contains also higher-order interactions (three-body, four-body, etc.).

One of the crucial instruments to achieve a rigorous control of the model is the so-called *Ruelle Probability Cascades* (RPCs), defined by Ruelle [22] when formalizing the properties of the Generalized Random Energy model of Derrida [23]. See also the characterization of RPC in terms of coalescent processes given in [24]. The first direct link between RPC and the SK model appeared in the work of Aizenman–Sims–Starr [25], where the authors found a representation of the thermodynamic limit of quenched pressure per particle in terms of the *cavity fields* distribution. This representation strongly suggested that if the thermodynamic limit of the overlap distribution is described by an RPC, then the Parisi formula is correct.

The first signal that the overlap array is described by an RPC was originally found by Aizenmann and Contucci in [10] with the identification of *stochastic stability* and by

Ghirlanda and Guerra [26]. Both papers show an (infinite) set of identities for the moments of the overlap array distribution. It turns out that these identities actually imply that the support of the joint distribution of the overlap is ultrametric, as proved by Panchenko [27]. It should be noticed that Panchenko’s theorem requires identities for the overlap moments of all orders. The latter do not hold for the bare SK model, but it can be shown that there exists a perturbation of the Hamiltonian that forces the SK model to satisfy them without affecting the limit of the quenched pressure [28].

Once the validity of the Parisi Formula (7) is established, it is natural to ask for the properties of its solution. The uniqueness of the minimizer of (7) has been assessed by Auffinger and Chen [29], and its properties have been investigated for example in [30,31].

A relevant question about the minimizer is the following: for which values of the parameters (β, h) is the solution of (7) a Dirac-delta function δ_q for some $q \in [0, 1]$? In this case, we say that the model is *replica symmetric* and the Parisi Formula (7) reads

$$\lim_{N \rightarrow \infty} p_N^{\text{SK}}(\beta, h) = \inf_{q \in [0,1]} \left\{ \log 2 + \mathbb{E}_{z, \xi} \log \cosh(\beta z \sqrt{q} + h \xi) + \frac{\beta^2}{4} (1 - q)^2 \right\}. \quad (13)$$

The replica symmetric region can be identified [6,32] with the region of parameters (β, h) where the overlap is a self-averaging quantity, namely

$$\lim_{N \rightarrow \infty} \mathbb{E} \langle (q_{1,2} - q^*)^2 \rangle_N = 0 \quad (14)$$

where q^* is exactly the value that realizes the infimum in (13). The physics conjecture is that the replica symmetric region can be identified by the so called Almeida–Thouless [33]

$$\beta^2 \mathbb{E}_{z, \xi} \cosh^{-4}(\beta z \sqrt{q^*} + h \xi) \leq 1. \quad (15)$$

The above conjecture is proved only in the case of Gaussian external field $\xi_i \sim \mathcal{N}(0, 1)$ [34]. An alternative characterization of the replica symmetric region has been obtained in [6,35]. If the minimizer corresponds to a non-trivial distribution (i.e., with non-zero variance), we say that *replica symmetry breaking* occurs, and the overlap is not a self-averaging quantity.

The Parisi formula has been extended to other mean field models with centered Gaussian interactions: vector spins [36], multispecies models [11,37,38], multiscale models [39,40]. Finally, we mention that the SK model fulfills a remarkable universality property: as long as z_{ij} ’s are independent, centered, and with unit variance, the thermodynamic limit is still described by the Parisi solution [41].

In this work, we show that a class of non-centered Gaussian spin glasses admits an interpretation of high-dimensional inference that extends the celebrated correspondence between the spiked Wigner model and the SK model in the Nishimori line where replica symmetry is always fulfilled [3]. We show that the addition of an SK Hamiltonian to a Hopfield with a finite number of patterns can be mapped into a high-dimensional mismatched inference problem, where the statistician ignores the correct apriori distribution on the signal components they have to reconstruct. We shall see that even this slight mismatch may lead to the emergence of complexity, namely to the breakdown of replica symmetry, which is instead guaranteed under very mild hypotheses for *optimal* statisticians.

3. High-Dimensional Inference and Statistical Physics

High-dimensional inference aims at recovering a *ground truth* signal, ξ in the following, that is usually a vector with a very large number of components from some noisy observations of it, which is denoted by \mathbf{Y} . The main feature of this setting is that the dimension of the signal, i.e., the number of real parameters to reconstruct, and the number of observations at disposal are a function of one another, typically a polynomial. For instance, for our purposes, ξ will be a vector of \mathbb{R}^N and \mathbf{Y} will be an $N \times N$ matrix for a total of N^2

noisy observations. Hence, if the number of observations becomes large, the number of parameters to retrieve also does. Contrary to what happens in typical low-dimensional settings, where max-likelihood, or Maximum A Posteriori (MAP) approaches yield provably satisfactory reconstruction performances, in a high-dimensional setting, this is not always the case. In particular, one needs to devise another kind of more refined estimators that exploit the marginal posterior probabilities for each signal component.

Both approaches described above are Bayesian, and the knowledge of a prior distribution on the signal components can play a key role especially for high-dimensional problems. Furthermore, to compose the posterior measure for the entire signal, one needs the *likelihood* of the data, which is the probability of an outcome \mathbf{y} of the variable \mathbf{Y} given a certain ground truth realization $\boldsymbol{\zeta} = \mathbf{x}$. As we shall discuss soon, under certain hypotheses, the Bayesian approach highlights the correspondence of relevant information theoretic quantities with thermodynamic ones. Among the others, a key quantity is the *mutual information* between the signal $\boldsymbol{\zeta}$ and the observations \mathbf{Y} , which quantifies the residual amount of information left in \mathbf{Y} about $\boldsymbol{\zeta}$ after the noise corruption. As intuition may suggest, the mutual information gives access to the best reconstruction error that is information theoretically achievable.

Finally, we stress that the high dimensionality of the problem can induce phase transition in some parameters of the model, like the so-called *signal-to-noise ratio* (SNR), that tunes the strength of the signal with respect to that of the noise in the observations.

3.1. Bayes-Optimality and Nishimori Identities

For the sake of simplicity, we start by considering a signal $\boldsymbol{\zeta} = (\zeta_i)_{i \leq N} \in \mathbb{R}^N$ of *i.i.d.* (independently and identically distributed) components $\zeta_i \stackrel{\text{iid}}{\sim} P_{\zeta}$, where P_{ζ} has a finite fourth moment. The observations at the disposal of a statistician can be modeled as a stochastic function of the ground-truth signal: $\mathbf{Y} = \mathbf{F}(\boldsymbol{\zeta}; \mathbf{z})$, where \mathbf{z} is the source of randomness or simply the noise. Knowing the function \mathbf{F} , from a Bayesian perspective, translates directly into having the likelihood of the model, namely the conditional distribution $dP_{\mathbf{Y}|\boldsymbol{\zeta}=\mathbf{x}}(\mathbf{y}) = p_{\mathbf{Y}|\boldsymbol{\zeta}=\mathbf{x}}(\mathbf{y})d\mathbf{y}$, which we assume to have a density $p_{\mathbf{Y}|\boldsymbol{\zeta}=\mathbf{x}}(\mathbf{y})$ over the Lebesgue measure. Observe that the likelihood is strongly affected by the nature of the noise.

According to Bayes' rule, the posterior distribution of $\boldsymbol{\zeta}$ given the data is:

$$dP_{\boldsymbol{\zeta}|\mathbf{Y}=\mathbf{y}}(\mathbf{x}) = \frac{p_{\mathbf{Y}|\boldsymbol{\zeta}=\mathbf{x}}(\mathbf{y})dP_{\boldsymbol{\zeta}}(\mathbf{x})}{\mathcal{Z}(\mathbf{y})}, \quad \mathcal{Z}(\mathbf{y}) = \int p_{\mathbf{Y}|\boldsymbol{\zeta}=\mathbf{x}}(\mathbf{y})dP_{\boldsymbol{\zeta}}(\mathbf{x}), \quad (16)$$

where $dP_{\boldsymbol{\zeta}}(\mathbf{x}) = \prod_{i \leq N} dP_{\zeta}(x_i)$, and $\mathcal{Z}(\mathbf{y})$ is the probability of a given realization of the data, which is sometimes also called *evidence*. In practice, the above posterior, which would be ideal to perform inference, is rarely available, and the statistician is not aware either of the likelihood or of the correct prior distribution for the signal, or even both. This motivates the following definition of a special inference setting:

Definition 1 (Bayes optimality). *The statistician is said to be Bayes optimal, or in the Bayes-optimal setting, if they are aware both of P_{ζ} and $\mathbf{F}(\cdot; \mathbf{z})$; namely, they have access to the posterior (16).*

The above is saying that an optimal statistician knows everything about the model except for the ground truth $\boldsymbol{\zeta}$ itself. The Bayes-optimal setting is thus often used as a theoretical framework to establish the *information theoretical limits*. Indeed, it is known that the mean square error between the ground truth and an estimator $\hat{\boldsymbol{\zeta}}(\mathbf{y})$

$$\text{MSE}(\hat{\boldsymbol{\zeta}}) = \mathbb{E} \|\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}(\mathbf{y})\|^2 \quad (17)$$

is minimized by an optimal statistician that can use the posterior mean as an estimator, yielding the *minimum mean square error* (MMSE)

$$\text{MMSE} = \mathbb{E} \|\boldsymbol{\zeta} - \mathbb{E}_{\boldsymbol{\zeta}|\mathbf{Y}}\boldsymbol{\zeta}\|^2. \quad (18)$$

In the following, we shall denote averages with respect to the posterior as $\langle \cdot \rangle_{\mathbf{Y}}$.

Another important consequence of this setting is the so-called *Nishimori identities*, which can be stated as follows. Given any continuous bounded function f of the data \mathbf{Y} , the ground truth ξ and $n - 1$ i.i.d. samples from the posterior $(\mathbf{x}^{(k)})_{k=2}^n$, one has

$$\mathbb{E}\langle f(\mathbf{Y}, \xi, (\mathbf{x}^{(k)})_{k=2}^n) \rangle_{\mathbf{Y}} = \mathbb{E}\langle f(\mathbf{Y}, \mathbf{x}^{(1)}, (\mathbf{x}^{(k)})_{k=2}^n) \rangle_{\mathbf{Y}}, \tag{19}$$

where $\mathbf{x}^{(1)} \sim P_{\xi|\mathbf{Y}}$. An elementary proof can be found in [42]. These identities are enforcing a symmetry between replicas drawn from the posterior and the ground truth. For instance, a direct application of the Nishimori identities yields

$$\text{MMSE} = \mathbb{E}\|\xi - \langle \mathbf{x} \rangle_{\mathbf{Y}}\|^2 = \mathbb{E}\langle \|\mathbf{x} - \langle \mathbf{x} \rangle_{\mathbf{Y}}\|^2 \rangle_{\mathbf{Y}}. \tag{20}$$

It is important to stress that, as it can be seen from the above equation, an optimal statistician is actually able to compute the minimum mean square error using their posterior.

At this point, the reader will have noticed a similarity with the Statistical Mechanics formalism. In fact, it is possible to interpret $\mathcal{Z}(\mathbf{y})$ as the partition function of a model with Hamiltonian $-\log p_{\mathbf{Y}|\xi=\mathbf{x}}(\mathbf{y})$ and unit inverse absolute temperature. The pressure per particle of such a model would thus be

$$p_N(\mathbf{y}) = \frac{1}{N} \log \mathcal{Z}(\mathbf{y}), \quad \bar{p}_N = \mathbb{E}_{\mathbf{Y}} p_N(\mathbf{Y}) = -\frac{1}{N} \mathcal{H}(\mathbf{Y}), \tag{21}$$

namely minus the Shannon entropy of the data per signal component, which is related to the mutual information

$$\frac{1}{N} I_N(\mathbf{Y}; \xi) = \frac{1}{N} \mathcal{H}(\mathbf{Y}) - \frac{1}{N} \mathcal{H}(\mathbf{Y} | \xi). \tag{22}$$

The contribution coming from the conditional entropy $\mathcal{H}(\mathbf{Y} | \xi)$ can be regarded as due only to the noise, since for fixed ξ , the only randomness in \mathbf{Y} is due to \mathbf{Z} .

We stress here that Bayes optimality, and the Nishimori identities, under rather mild hypotheses [43] are enough to grant *replica symmetry* in the model, i.e., concentration of the order parameters in the model. For the models we are interested in, the latter can be shown to imply finite-dimensional variational principles for the limiting mutual information.

3.2. The Spiked Wigner Model

The spiked Wigner model (WSM) was first introduced in [44] as a model for Principal Component Analysis (PCA), and since then, it was widely studied in recent literature. Without pretension of being exhaustive, we refer the interested reader to [42,45–51]. For our purposes, we restrict ourselves to the case where the signal is an N -dimensional vector of ± 1 s, drawn from a Rademacher distribution $\xi_i \stackrel{\text{iid}}{\sim} P_{\xi} = (\delta_{-1} + \delta_1)/2$. The function $\mathbf{F}(\cdot; \mathbf{z})$ is a *Gaussian channel*, namely

$$Y_{ij} = \sqrt{\frac{\mu}{2N}} \xi_i \xi_j + z_{ij} \tag{23}$$

where $z_{ij} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$, and μ is a positive parameter called the *signal-to-noise ratio*. The statistician is tasked with the recovery of ξ given the observations \mathbf{Y} . The Bayes-optimal posterior measure for this inference problem can be written directly as a Boltzmann–Gibbs random measure thanks to the Gaussian nature of the likelihood:

$$G_N(\sigma) = \frac{1}{\mathcal{Z}(\mathbf{z}, \xi)} e^{-H_N(\sigma; \mathbf{z}, \xi)}, \quad -H_N(\sigma; \mathbf{z}, \xi) = \sum_{i,j=1}^N \sqrt{\frac{\mu}{2N}} z_{ij} \sigma_i \sigma_j + \frac{\mu}{2N} \sigma_i \xi_i \sigma_j \xi_j \tag{24}$$

where we have already exploited the fact that $(\xi_i)^2 = \sigma_i^2 = 1$. We are denoting the posterior samples with σ . Since the quantity we are interested in is the quenched pressure of this model

$$\bar{p}_N = \frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \{-1,1\}^N} \exp [- H_N(\sigma; \mathbf{z}, \xi)] \tag{25}$$

that is connected to the mutual information $I_N(\mathbf{Y}; \xi) / N$ by a simple shift with an additive constant, we are allowed to perform a gauge transformation without altering its value:

$$z_{ij} \mapsto z_{ij} \tilde{\xi}_i \tilde{\xi}_j, \quad \sigma_i \sigma_j \mapsto \sigma_i \tilde{\xi}_i \tilde{\xi}_j \sigma_j. \tag{26}$$

This results in a Hamiltonian that is now independent of the original ground-truth signal

$$-H'_N(\sigma; \mathbf{z}) = \sum_{i,j=1}^N \left(\sqrt{\frac{\mu}{2N}} z_{ij} + \frac{\mu}{2N} \right) \sigma_i \sigma_j \tag{27}$$

and the coupling between spins are Gaussian random variables with a mean equal to their variance. This condition identifies a peculiar region of the phase space of a spin-glass model, which is called *Nishimori line*. In fact, the Nishimori identities were first discovered and studied in the context of gauge spin-glasses. Despite looking simpler, the above model retains most of the features we need for our study.

For inference models with additive Gaussian noise, like the one above, it is possible to prove the so-called *I-MMSE* relation:

$$\frac{d}{d\mu} \frac{I_N(\mathbf{Y}; \xi)}{N} = \frac{1}{4N^2} \mathbb{E} \| \xi \xi^T - \langle \sigma \sigma^T \rangle \|_F^2, \tag{28}$$

where $\| \cdot \|_F$ is the Frobenius norm and $\langle \cdot \rangle$ denotes the expectation with respect to the Boltzmann–Gibbs measure induced by (27). Hence, once the mutual information is known, the MMSE can be accessed through a derivative with respect to the signal-to-noise ratio. A clarification is in order here: the above is the MMSE on the reconstruction of the rank-one matrix $\xi \xi^T$, because, due to flip symmetry, here we do not have any actual information on the single vector ξ , but only on the *spike* $\xi \xi^T$.

3.3. Sub-Optimality and Replica Symmetry Breaking

There are several ways to break Bayes optimality. Some examples are that the statistician does not know the signal-to-noise ratio μ [13,52]; the statistician adopts a likelihood different from that of the true model [14]; the statistician adopts a wrong prior [12,53]; combinations of the previous and many others. We will focus on the mismatching priors case, where the statistician not only adopts a wrong prior on the ground-truth elements, but they are not aware of the rank of the spiked matrix hidden inside the noise, which is denoted by M . The rest is assumed to be known. The channel of the inference problem is

$$Y_{ij} = \sqrt{\frac{\mu}{2N}} \sum_{k=1}^M \xi_i^{(k)} \xi_j^{(k)} + z_{ij}. \tag{29}$$

If the statistician assumes a Rademacher prior to the signal components and a rank-one hidden matrix, they will write a posterior in the form

$$\mathcal{P}_{\xi|\mathbf{z},\xi}(\sigma) = \frac{1}{\mathcal{Z}(\mathbf{z}, \xi)} \exp (- H_N(\sigma; \mathbf{z}, \xi)) \tag{30}$$

where

$$-H_N(\sigma; \mathbf{z}, \boldsymbol{\zeta}) = \sum_{i,j=1}^N \left(\sqrt{\frac{\mu}{2N}} z_{ij} + \frac{\mu}{2N} \sum_{k=1}^P \zeta_i^{(k)} \zeta_j^{(k)} \right) \sigma_i \sigma_j \tag{31}$$

The slash on quantities emphasizes that they are not the Bayes-optimal ones. In this setting, one can no longer rely on the Nishimori identities, and in principle, replica symmetry is no longer guaranteed. On the contrary, as we shall argue later on, a mismatch in the prior only is already sufficient to cause *replica symmetry breaking*.

4. The Model

Let M be a fixed integer and $k \in \{1, \dots, M\}$. Consider two independent random collections $(z_{ij})_{i,j \leq N} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$ and $(\zeta_i^{(k)})_{i \leq N}^{k \leq M} \stackrel{\text{iid}}{\sim} P_\zeta$ where P_ζ is such that $\mathbb{E}[\zeta^4] < \infty$. The above random collections play the role of quenched disorder in the model. Consider N Ising spins $\sigma = (\sigma_i)_{i \leq N} \in \{+1, -1\}^N$ and the Hamiltonian function

$$\begin{aligned} H_N(\sigma; \mu, \nu, \lambda) &\equiv H_N(\sigma) = H_N^{\text{int}}(\sigma) - \mathbf{h} \cdot \sigma \\ &= - \sum_{i,j=1}^N \left(\sqrt{\frac{\mu}{2N}} z_{ij} + \frac{\nu}{2N} \sum_{k \leq M} \zeta_i^{(k)} \zeta_j^{(k)} \right) \sigma_i \sigma_j - \sum_{k \leq M} \lambda_k \sum_{i=1}^N \zeta_i^{(k)} \sigma_i \end{aligned} \tag{32}$$

with $\mu, \nu \geq 0$, $\lambda = (\lambda_k)_{k \leq M} \in \mathbb{R}^M$. Here, H_N^{int} is the interacting part while

$$\mathbf{h} = (h_i)_{i \leq N} \equiv \mathbf{h}(\lambda, \boldsymbol{\zeta}) = \left(\sum_{k \leq M} \lambda_k \zeta_i^{(k)} \right)_{i \leq N} \tag{33}$$

denotes the random external field acting on the spins. The Hamiltonian (32) is determined by the choice of M, μ, ν, λ and P_ζ . For $\mu = \nu$, the interaction term H_N^{int} coincides with the Hamiltonian (31). Note that for some special choices of the parameters, we recover some well-known spin glass models:

- $\nu = 0$ gives the SK model (1) at $\beta = \sqrt{\mu}$ and random external field \mathbf{h} .
- $\mu = 0$ gives the Hopfield model [6,7,18] with a finite number of patterns $(\boldsymbol{\zeta}^{(k)})_{k \leq M}$.
- $M = 1, \nu = \mu, \lambda_1 = 0$ and $P_\zeta = \frac{1}{2}(\delta_{-1} + \delta_{+1})$ gives the SK model on the Nishimori line (27). As we have seen in Section 3, the latter can be also viewed as a spiked Wigner model in the Bayesian-optimal setting.

Notice that the entire model can be interpreted as a *Hopfield model* where the traditional Hebbian matrix $\sum_{k \leq M} \zeta_i^{(k)} \zeta_j^{(k)}$ is corrupted by Gaussian noise. Furthermore, if the Hebbian coupling is replaced by a constant matrix, the model reduces to an SK model with the addition of a ferromagnetic interaction, and it was studied in [54].

Our main result is the computation of the thermodynamic limit of the pressure per particle

$$p_N(\mu, \nu, \lambda) = \frac{1}{N} \log \sum_{\sigma \in \{-1, 1\}^N} e^{-H_N(\sigma)} \tag{34}$$

whose variance can be shown to converge to 0 as an $O(N^{-1})$, namely:

Lemma 1. Assume $\mathbb{E}\zeta_1^4 < \infty$. Then, for any $\mu, \nu \in \mathbb{R}$ and $\lambda \in \mathbb{R}^K$

$$\mathbb{E}[p_N(\mu, \nu, \lambda) - \mathbb{E}p_N(\mu, \nu, \lambda)]^2 \leq \frac{K}{N} \tag{35}$$

where K is a suitable positive constant.

We thus focus on $\bar{p}_N(\mu, \nu, \lambda) = \mathbb{E} p_N(\mu, \nu, \lambda)$. The proof of this lemma makes use of the Efron–Stein concentration inequality to bound the variance, and it is simple but tedious. It follows closely that of ([12], Lemma 9). We are now in a position to state our main theorem:

Theorem 2 (Variational solution). *If $\mathbb{E}[\zeta^4] < +\infty$ then*

$$\lim_{N \rightarrow \infty} \bar{p}_N(\mu, \nu, \lambda) = \sup_{x \in \mathbb{R}^M} \varphi(x; \mu, \nu, \lambda) \tag{36}$$

where

$$\varphi(x; \mu, \nu, \lambda) := -\frac{\nu|x|^2}{2} + \mathbb{E} \mathcal{P}(\sqrt{\mu}, h(x, \lambda, \zeta)), \quad x = (x_k)_{k \leq M} \in \mathbb{R}^M \tag{37}$$

and \mathcal{P} is the Parisi functional (5) with a random external field

$$h(x, \lambda, \zeta) = \sum_{k \leq M} (\lambda_k + \nu x_k) \zeta^{(k)} \tag{38}$$

and \mathbb{E} denotes the expectation with respect to $\zeta^{(k)} \stackrel{\text{iid}}{\sim} P_\zeta$. The consistency equations are

$$x_k = \frac{\partial}{\partial x_k} \mathbb{E} \mathcal{P}(\sqrt{\mu}, h(x, \lambda, \zeta)), \quad k \leq M \tag{39}$$

Moreover, there exists $C > 0$ such that for any $k \leq M$, one has $|x_k| \leq C$ and the supremum in (36) can be restricted to $[-C, C]^M$.

The proof of the theorem is based on the concentration of the Mattis magnetization, which is the normalized scalar product between a spin-configuration (or sample from the wrong posterior measure) and one of the $\zeta^{(k)}$:

$$m_N(\sigma | \zeta^{(k)}) = \frac{1}{N} \sum_{i=1}^N \sigma_i \zeta_i^{(k)} = \frac{1}{N} \sigma \cdot \zeta^{(k)}. \tag{40}$$

The Hamiltonian can thus be rewritten using (40) in the following form:

$$H_N(\sigma) = \sqrt{\mu} H_N^{SK}(\sigma) - N \sum_{k \leq M} \left[\frac{\nu}{2} m_N^2(\sigma | \zeta^{(k)}) + \lambda_k m_N(\sigma | \zeta^{(k)}) \right] \tag{41}$$

The Mattis magnetization, in fact, plays the role of an order parameter for this model. The concentration we can prove is only an integral average over some suitably small magnetic fields, which is still sufficient for our purposes:

Proposition 1 (Concentration of Mattis Magnetizations). *Consider a k such that $k \leq M$. Let $\epsilon_k \in [s_N, 2s_N]$ with $s_N = \frac{1}{2}N^{-\alpha}$, $\alpha \in (0, 1/(2M))$ for all $k \leq M$. For any $y \in \mathbb{R}$, we denote by $\langle \cdot \rangle_{N,y}$ the Boltzmann–Gibbs measure induced by the Hamiltonian $H_{N,y}(\sigma) = H_N(\sigma) - y \sigma \cdot \zeta^{(k)}$. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{s_N^M} \int_{s_N}^{2s_N} \prod_{\ell \leq M} d\epsilon_\ell \mathbb{E} \left\langle \left(m_N(\sigma | \zeta^{(k)}) - \mathbb{E} \langle m_N(\sigma | \zeta^{(k)}) \rangle_{N,\epsilon} \right)^2 \right\rangle_{N,\epsilon} = 0, \tag{42}$$

for all $\mu, \nu \geq 0$ and $\lambda \in \mathbb{R}^M$.

We shall omit the proof of the above result as it is completely analogous to the one in [12]. We will need an intermediate lemma that leads to it (see Lemma 2 later) together with a second key ingredient: the *adaptive interpolation* technique [48] combined with

Guerra’s replica symmetry-breaking upper bound for the quenched pressure of the SK model [8].

Proof of Theorem 2. Here, we outline the main steps of the proof of the variational principle for the thermodynamic limit. The proof is achieved via two bounds that match in the $N \rightarrow \infty$ limit. Let us start by defining the interpolating Hamiltonian

$$\begin{aligned}
 H_N(t; \sigma) &:= H_N(\sigma; \mu, (1-t)\nu, \lambda + R_\epsilon(t)) = \\
 &= \sqrt{\mu} H_N^{SK}(\sigma) - (1-t) \frac{N\nu}{2} \sum_{k \leq M} m_N^2(\sigma | \zeta^{(k)}) - N \sum_{k \leq M} (\lambda_k + R_\epsilon^{(k)}(t)) m_N(\sigma | \zeta^{(k)})
 \end{aligned} \tag{43}$$

where $R_\epsilon = (R_\epsilon^{(k)})_{k \leq M}$ and

$$R_\epsilon^{(k)}(t) = \epsilon_k + \nu \int_0^t ds r_\epsilon^{(k)}(s), \quad \epsilon_k \in [s_N, 2s_N], \quad s_N = \frac{N^{-\alpha}}{2} \tag{44}$$

with $\alpha \in (0, 1/(2M))$ and where the interpolating functions $r_\epsilon^{(k)}$, that must be continuously differentiable in $[0, 1]$ and non negative, will be suitably chosen. With this interpolation, one is able to prove the following sum rule:

Proposition 2. *The following sum rule holds:*

$$\bar{p}_N(\mu, \nu, \lambda) = \bar{p}_N^{SK}(\sqrt{\mu}, \lambda + R_\epsilon(1)) - \frac{\nu}{2} \int_0^1 dt \sum_{k \leq M} [r_\epsilon^{(k)2}(t) - \Delta_\epsilon^{(k)}(t)] + O(s_N) \tag{45}$$

where

$$\Delta_\epsilon^{(k)}(t) := \mathbb{E} \left\langle \left(m_N(\sigma | \zeta^{(k)}) - r_\epsilon^{(k)}(t) \right)^2 \right\rangle_{N, R_\epsilon(t)}. \tag{46}$$

The proof consists of the computation of the derivative of the interpolating pressure related to the model (43). It follows closely that of ([12], Proposition 7), to which we refer the interested reader. Since the remainder $\Delta_\epsilon^{(k)}$ is non-negative, the above proposition already yields a bound for the quenched pressure of our model when we choose $r_\epsilon^{(k)} = x_k \in \mathbb{R}$ constant:

$$\liminf_{N \rightarrow \infty} \bar{p}_N(\mu, \nu, \lambda) \geq \sup_{x \in \mathbb{R}^M} \varphi(x; \mu, \nu, \lambda) \tag{47}$$

where we used Lipschitz continuity of the SK pressure in the magnetic fields.

The upper bound requires more attention. First, we notice that $\bar{p}_N^{SK}(\sqrt{\mu}, \lambda + R_\epsilon(1))$ is convex in the magnetic fields and that $R_\epsilon^{(k)}(1) = \int_0^1 dt (\epsilon_k + \nu r_\epsilon^{(k)}(t))$. Hence, we can use Jensen’s inequality and Lipschitz continuity of \bar{p}^{SK} to obtain:

$$\bar{p}_N(\mu, \nu, \lambda) \leq \int_0^1 dt \left[\bar{p}_N^{SK}(\sqrt{\mu}, \lambda + r_\epsilon(t)) - \nu \sum_{k \leq M} \frac{r_\epsilon^{(k)2}(t)}{2} \right] + \frac{\nu}{2} \sum_{k \leq M} \int_0^1 \Delta_\epsilon^{(k)}(t) dt + O(s_N). \tag{48}$$

Now, we use Guerra’s bound for the SK pressure, that, importantly, is uniform in N , and we average over ϵ on both sides

$$\begin{aligned}
 \bar{p}_N(\mu, \nu, \lambda) &\leq \mathbb{E}_\epsilon \int_0^1 \varphi(r_\epsilon(t); \mu, \nu, \lambda) dt + \frac{\nu}{2} \sum_{k \leq M} \mathbb{E}_\epsilon \int_0^1 \Delta_\epsilon^{(k)}(t) dt + O(s_N) \leq \\
 &\leq \sup_{x \in \mathbb{R}^M} \varphi(x; \mu, \nu, \lambda) + \frac{\nu}{2} \sum_{k \leq M} \mathbb{E}_\epsilon \int_0^1 \Delta_\epsilon^{(k)}(t) dt + O(s_N).
 \end{aligned} \tag{49}$$

What remains to do is to prove that $\mathbb{E}_{\epsilon} \Delta_{\epsilon}^{(k)}(t) \xrightarrow{N \rightarrow \infty} 0$ for a proper choice of the interpolating functions R_{ϵ} . The choice is made through a system of coupled ODEs

$$\dot{R}_{\epsilon}^{(k)}(t) \equiv \nu r_{\epsilon}^{(k)}(t) = \nu \mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, R_{\epsilon}(t)}, \quad R_{\epsilon}^{(k)}(0) = \epsilon_k, \quad \forall k \leq M. \quad (50)$$

One can easily check that the above system is regular enough to admit a unique solution on the interval $t \in [0, 1]$. In this case, the remainder to push to 0 would appear as

$$\frac{1}{s_N^M} \int_{s_N}^{2s_N} \prod_{\ell \leq M} d\epsilon_{\ell} \mathbb{E} \left\langle \left(m_N(\sigma | \xi^{(k)}) - \mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, R_{\epsilon}(t)} \right)^2 \right\rangle_{N, R_{\epsilon}(t)}. \quad (51)$$

The goal is now to apply a concentration lemma here:

Lemma 2. *Let $y \in [y_1, y_2]$, $\delta \in (0, 1)$ and denote by $\langle \cdot \rangle_{N, y}$ the Boltzmann–Gibbs expectation associated to the Hamiltonian $H_N(\sigma; \mu, \nu, \lambda + y \mathbf{e}_k)$ where $k \leq M$ and \mathbf{e}_k is the k -th canonical basis vector of \mathbb{R}^M . Then*

$$\mathbb{E} \left\langle \left(m_N(\sigma | \xi^{(k)}) - \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, y} \right)^2 \right\rangle_{N, y} = \frac{1}{N} \frac{d^2}{dy^2} \bar{p}_N(\mu, \nu, \lambda + y \mathbf{e}_k) \quad (52)$$

$$\begin{aligned} \mathbb{E} \left[\left(\langle m_N(\sigma | \xi) \rangle_{N, y} - \mathbb{E} \langle m_N(\sigma | \xi) \rangle_{N, y} \right)^2 \right] &\leq \frac{12K}{\delta^2 N} + \\ &+ 8\sqrt{a} \frac{d}{dy} [\bar{p}_N(\mu, \nu, \lambda + (y + \delta) \mathbf{e}_k) - \bar{p}_N(\mu, \nu, \lambda + (y - \delta) \mathbf{e}_k)] \end{aligned} \quad (53)$$

with K a positive constant.

Notice that the integral in (51) is over ϵ and not over the effective magnetic field of the model, which is instead $R_{\epsilon}(t)$. Nevertheless, we can integrate over the magnetic fields $R_{\epsilon}(t)$ with a change of variables. This involves a Jacobian that is larger than 1. In fact, thanks to Liouville’s theorem ([55], Corollary 3.1, Chapter V), one can prove that

$$\begin{aligned} \det \frac{\partial R_{\epsilon}(t)}{\partial \epsilon} &= \exp \left[\int_0^t \nu \sum_{k \leq M} \frac{\partial}{\partial R_{\epsilon}^{(k)}(s)} \mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, R_{\epsilon}(s)} ds \right] = \\ &\exp \left[\int_0^t N \nu \sum_{k \leq M} \mathbb{E} \left\langle \left(m_N(\sigma | \xi^{(k)}) - \mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, R_{\epsilon}(s)} \right)^2 \right\rangle_{N, R_{\epsilon}(s)} ds \right] \geq 1, \end{aligned} \quad (54)$$

when $\nu \geq 0$.

This allows us to bound the thermal fluctuations in (51) using (52) and then Liouville’s theorem:

$$\begin{aligned} \Delta_T^{(k)} &:= \frac{1}{s_N^M} \int_{s_N}^{2s_N} \prod_{\ell \leq M} d\epsilon_{\ell} \mathbb{E} \left\langle \left(m_N(\sigma | \xi^{(k)}) - \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, R_{\epsilon}(t)} \right)^2 \right\rangle_{N, R_{\epsilon}(t)} \leq \\ &\leq \frac{1}{s_N^M} \prod_{\ell \leq M} \int_{R_{s_N}^{(\ell)}(t)}^{R_{2s_N}^{(\ell)}(t)} dh_{\ell} \frac{1}{N} \frac{d^2}{dh_k^2} \bar{p}_N(\mu, \nu, \lambda + \mathbf{h}) = \\ &= \frac{1}{Ns_N^M} \prod_{\ell \leq M, \ell \neq k} \int_{R_{s_N}^{(\ell)}(t)}^{R_{2s_N}^{(\ell)}(t)} dh_{\ell} \left[\mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, \mathbf{h}; h_k = R_{2s_N}^{(k)}} - \mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, \mathbf{h}; h_k = R_{s_N}^{(k)}} \right] \end{aligned} \quad (55)$$

Since ξ_i has a bounded second moment, using Cauchy–Schwarz inequality, one can show that $|\mathbb{E} \langle m_N(\sigma | \xi^{(k)}) \rangle_{N, \mathbf{h}}|$ is uniformly bounded by a constant C . Hence, $|R_{\epsilon}^{(k)}(t)| \leq \epsilon_k + Ct\nu \leq 1 + C\nu$ for any $k \leq M$ by construction (recall (44) and (50)). Therefore, $\Delta_T = O\left(\frac{1}{Ns_N^M}\right)$.

The fluctuations induced by the disorder can be bounded in a very similar fashion using (53):

$$\Delta_D^{(k)} := \frac{1}{s_N^M} \int_{s_N}^{2s_N} \prod_{\ell \leq M} d\epsilon_\ell \mathbb{E} \left(\langle m_N(\sigma|\zeta^{(k)}) \rangle_{N, R_\epsilon(t)} - \mathbb{E} \langle m_N(\sigma|\zeta^{(k)}) \rangle_{N, R_\epsilon(t)} \right)^2 = O\left(\frac{1}{N\delta^2} + \frac{\delta}{s_N^M}\right) \quad (56)$$

Hence, overall (51), that equals $\Delta_T^{(k)} + \Delta_D^{(k)}$ is a $O\left(\frac{1}{Ns_N^M} + \frac{1}{N\delta^2} + \frac{\delta}{s_N^M}\right)$. δ can be chosen as a function of N in order to optimize the convergence rate: $\delta = s_N^{2M/3} N^{-1/3}$. Using Fubini's theorem in (49) to exchange the t and ϵ averages and then Dominated Convergence, one concludes the proof. \square

From the the variational problem (36), we can deduce also the differentiability properties of the limiting pressure obtaining the average values of the relevant thermodynamic quantities of the model:

Corollary 1. *Let $p(\mu, \nu, \lambda) = \lim_{N \rightarrow \infty} \bar{p}_N(\mu, \nu, \lambda)$, and $\Omega_{\mu, \nu, \lambda} = \operatorname{argmax}_{x \in [-C, C]^M} \varphi(x; \mu, \nu, \lambda)$. Then*

- *If $\Omega_{\mu, \nu, \lambda} = \{\bar{x}\}$ is a singleton with $\bar{x} = (\bar{x}_k)_{k \leq M}$ then for any $k \leq M$, one has*

$$\lim_{N \rightarrow \infty} \mathbb{E} \langle m_N(\sigma|\zeta^k) \rangle_N = \frac{\partial}{\partial \lambda_k} p(\mu, \nu, \lambda) = \bar{x}_k \quad (57)$$

and

$$\lim_{N \rightarrow \infty} \mathbb{E} \langle q_{12}^2 \rangle_N = 1 - 4 \frac{\partial}{\partial \mu} p(\mu, \nu, \lambda) = \int q^2 d\chi^*(q). \quad (58)$$

- *where $\chi^*(q)$ denotes the unique measure solving the Parisi variational principle in Theorem 1. If $\{|x|^2, x \in \Omega_{\mu, \nu, \lambda}\} = \{|\bar{x}|^2\}$ is a singleton then*

$$\frac{\partial}{\partial \nu} p(\mu, \nu, \lambda) = \frac{|\bar{x}|^2}{2}. \quad (59)$$

More generally, let y be one of the variables $\sqrt{\mu}, \nu, \lambda_1, \dots, \lambda_M$, then the function $y \mapsto p(\mu, \nu, \lambda)$ is convex. By Danskin theorem (see [56]), $y \mapsto p(\mu, \nu, \lambda)$ is differentiable if and only if the set $\left\{ \frac{\partial \varphi(x; \mu, \nu, \lambda)}{\partial y}, x \in \Omega_{\mu, \nu, \lambda} \right\}$ is a singleton.

5. Conclusions and Perspectives

In this paper, we offer an overview of the Parisi formula from a mathematical physics perspective, emphasizing its potential applications, particularly in addressing the mismatched inference problem outlined earlier. Building upon our previous work [12], we investigate a scenario where a statistician, tasked with reconstructing a finite-rank matrix, lacks knowledge about the underlying matrix generation process, including both the matrix elements and its rank. We consider the case in which the statistician assumes a rank-one matrix, leading to a mismatch between the "true" Bayes posterior and the one used for inference. Our key contribution is the proof that, contrary to what happens in the Bayes-optimal setting, this Bayesian mismatch induces replica symmetry breaking in the model. Consequently, we express the pressure of the corresponding spin glass as an infinite-dimensional variational principle over the space of distributions on $[0, 1]$.

The chosen mismatch scenario shares some similarities with those studied in [57,58] with the fundamental difference being that here the rank of the hidden matrix is finite. In a recent work [59], the authors consider a general case of mismatch, which includes mismatching priors and likelihoods. The mentioned paper proves a universality property with respect to the likelihood assumed by the statistician provided that observations remain independent given the ground truth.

Despite these advancements, all the proofs available so far in the literature break down when considering a high-rank hidden matrix. To rigorously comprehend this scenario, addressing the solution of the Hopfield model is of crucial importance. However, to the best of our knowledge, its complete solution remains elusive [5,6].

Author Contributions: All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: Plerluigi Contucci and Emanuele Mingione were partially supported by the EU H2020 ICT48 project Humane AI Net contract number 952026; by the Italian Extended Partnership PE01—FAIR Future Artificial Intelligence Research—Proposal code PE00000013 under the MUR National Recovery and Resilience Plan; by the project PRIN 2022—Proposal code: J53D23003690006.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors thank Jorge Kurchan, Nicolas Macris and Farzad Pourkamali for fruitful interactions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Parisi, G. An Infinite Number of Order Parameters for Spin Glasses. *Phys. Rev. Lett.* **1979**, *43*, 1754–1756. [[CrossRef](#)]
2. Parisi, G. A Sequence of Approximated Solutions to the S-K Model for Spin Glasses. *J. Phys. A* **1980**, *13*, L115. [[CrossRef](#)]
3. Nishimori, H. *Statistical Physics of Spin Glasses and Information Processing: An Introduction*; Oxford University Press: Oxford, NY, USA, 2001.
4. Mézard, M.; Montanari, A. *Information, Physics, and Computation*; Oxford Academic: Oxford, UK, 2009.
5. Talagrand, M. *Mean Field Models for Spin Glasses: Volume I: Basic Examples*; Springer: Berlin/Heidelberg, Germany, 2010.
6. Talagrand, M. *Mean Field Models for Spin Glasses: Volume II: Advanced Replica-Symmetry and Low Temperature*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 55. [[CrossRef](#)]
7. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Spin-glass models of neural networks. *Phys. Rev. A* **1985**, *32*, 1007–1018. [[CrossRef](#)] [[PubMed](#)]
8. Guerra, F. Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model. *Commun. Math. Phys.* **2003**, *233*, 1–12. [[CrossRef](#)]
9. Guerra, F.; Toninelli, F.L. The Thermodynamic Limit in Mean Field Spin Glass Models. *Commun. Math. Phys.* **2002**, *230*, 71–79. [[CrossRef](#)]
10. Aizenman, M.; Contucci, P. On the Stability of the Quenched State in Mean Field Spin Glass Models. *J. Stat. Phys.* **1998**, *92*, 765–783. [[CrossRef](#)]
11. Panchenko, D. The free energy in a multi-species Sherrington–Kirkpatrick model. *Ann. Probab.* **2015**, *43*, 3494–3513. [[CrossRef](#)]
12. Camilli, F.; Contucci, P.; Mingione, E. An inference problem in a mismatched setting: A spin-glass model with Mattis interaction. *SciPost Phys.* **2022**, *12*, 125. [[CrossRef](#)]
13. Pourkamali, F.; Macris, N. Mismatched Estimation of rank-one symmetric matrices under Gaussian noise. *arXiv* **2021**, arXiv:cs.IT/2107.08927.
14. Barbier, J.; Hou, T.; Mondelli, M.; Sáenz, M. The price of ignorance: How much does it cost to forget noise structure in low-rank matrix estimation? *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 36733–36747.
15. Fu, T.; Liu, Y.; Barbier, J.; Mondelli, M.; Liang, S.; Hou, T. Mismatched estimation of non-symmetric rank-one matrices corrupted by structured noise. *arXiv* **2023**, arXiv:2302.03306.
16. Barbier, J.; Chen, W.K.; Panchenko, D.; Sáenz, M. Performance of Bayesian linear regression in a model with mismatch. *arXiv* **2021**, arXiv:2107.06936.
17. Sherrington, D.; Kirkpatrick, S. Solvable Model of a Spin-Glass. *Phys. Rev. Lett.* **1975**, *35*, 1792–1796. [[CrossRef](#)]
18. Mezard, M.; Parisi, G.; Virasoro, M. *Spin Glass Theory and Beyond*; Lecture Notes in Physics Series; World Scientific: Singapore, 1987.
19. Talagrand, M. The Parisi Formula. *Ann. Math.* **2006**, *163*, 221–263. [[CrossRef](#)]
20. Panchenko, D. *The Sherrington-Kirkpatrick Model*; Springer: Berlin/Heidelberg, Germany, 2015.
21. Auffinger, A.; Jagannath, A. Thouless–Anderson–Palmer equations for generic p -spin glasses. *Ann. Probab.* **2019**, *47*, 2230–2256. [[CrossRef](#)]
22. Ruelle, D. A mathematical reformulation of Derrida’s REM and GREM. *Commun. Math. Phys.* **1987**, *108*, 225–239. [[CrossRef](#)]
23. Derrida, B. A generalization of the Random Energy Model which includes correlations between energies. *J. Phys. Lett.* **1985**, *46*, 401–407. [[CrossRef](#)]

24. Bolthausen, E.; Sznitman, A.S. On Ruelle's Probability Cascades and an Abstract Cavity Method. *Commun. Math. Phys.* **1998**, *197*, 247–276. [[CrossRef](#)]
25. Aizenman, M.; Sims, R.; Starr, S. Extended variational principle for the Sherrington-Kirkpatrick spin-glass model. *Phys. Rev. B* **2003**, *68*, 214403. [[CrossRef](#)]
26. Ghirlanda, S.; Guerra, F. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A Math. Gen.* **1998**, *31*, 9149. [[CrossRef](#)]
27. Panchenko, D. The Parisi ultrametricity conjecture. *Ann. Math.* **2011**, *177*, 383–393. [[CrossRef](#)]
28. Contucci, P.; Mingione, E.; Starr, S. Factorization Properties in d-Dimensional Spin Glasses. Rigorous Results and Some Perspectives. *J. Stat. Phys.* **2012**, *151*, 809–829. [[CrossRef](#)]
29. Auffinger, A.; Chen, W.K. The Parisi Formula has a Unique Minimzer. *Commun. Math. Phys.* **2015**, *335*, 1429–1444. [[CrossRef](#)]
30. Auffinger, A.; Chen, W.K. On properties of Parisi measures. *Probab. Theory Relat. Fields* **2013**, *161*, 817–850. [[CrossRef](#)]
31. Jagannath, A.; Tobasco, I. Some Properties of the Phase Diagram for Mixed p -Spin Glasses. *Probab. Theory Relat. Fields* **2017**, *167*, 615–672. [[CrossRef](#)]
32. Pastur, L.A.; Shcherbina, M. Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model. *J. Stat. Phys.* **1991**, *62*, 1–19. [[CrossRef](#)]
33. de Almeida, J.R.L.; Thouless, D.J. Stability of the Sherrington-Kirkpatrick solution of a spin glass model. *J. Phys. A Math. Gen.* **1978**, *11*, 983. [[CrossRef](#)]
34. Chen, W.K. On the Almeida-Thouless transition line in the SK model with centered Gaussian external field. *arXiv* **2021**, arXiv:2103.04802.
35. Guerra, F.; Toninelli, F.L. Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model. *J. Math. Phys.* **2002**, *43*, 3704–3716. [[CrossRef](#)]
36. Panchenko, D. Free energy in the mixed p -spin models with vector spins. *Ann. Probab.* **2018**, *46*, 865–896. [[CrossRef](#)]
37. Barra, A.; Contucci, P.; Mingione, E.; Tantari, D. Multi-Species Mean Field Spin Glasses. Rigorous Results. *Ann. Inst. Henri Poincaré* **2013**, *16*, 691–708. [[CrossRef](#)]
38. Subag, E. TAP approach for multispecies spherical spin glasses II: The free energy of the pure models. *Ann. Probab.* **2023**, *51*, 1004–1024. [[CrossRef](#)]
39. Contucci, P.; Mingione, E. A Multi-scale Spin-Glass Mean-Field Model. *Commun. Math. Phys.* **2019**, *368*, 1323–1344. [[CrossRef](#)]
40. Mourrat, J.C.; Panchenko, D. Extending the Parisi formula along a Hamilton-Jacobi equation. *Electron. J. Probab.* **2020**, *25*, 1–17. [[CrossRef](#)]
41. Carmona, P.; Hu, Y. Universality in Sherrington-Kirkpatrick's spin glass model. In *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*; Elsevier: Amsterdam, The Netherlands, 2006; Volume 42, pp. 215–222.
42. Lelarge, M.; Miolane, L. Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Relat. Fields* **2017**, *173*, 859–929. [[CrossRef](#)]
43. Barbier, J. Overlap matrix concentration in optimal Bayesian inference. *Inf. Inference A J. IMA* **2020**, *10*, 597–623. [[CrossRef](#)]
44. Johnstone, I.M. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Stat.* **2001**, *29*, 295–327. [[CrossRef](#)]
45. Krzakala, F.; Xu, J.; Zdeborová, L. Mutual information in rank-one matrix estimation. In Proceedings of the 2016 IEEE Information Theory Workshop (ITW), Cambridge, UK, 11–14 September 2016; pp. 71–75. [[CrossRef](#)]
46. Barbier, J.; Macris, N.; Miolane, L. The layered structure of tensor estimation and its mutual information. In Proceedings of the 55th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 3–6 October 2017.
47. Barbier, J.; Macris, N. The adaptive interpolation method for proving replica formulas. Applications to the Curie-Weiss and Wigner spike models. *J. Phys. A Math. Theor.* **2019**, *52*, 294002. [[CrossRef](#)]
48. Barbier, J.; Macris, N. The adaptive interpolation method: A simple scheme to prove replica formulas in Bayesian inference. *Probab. Theory Relat. Fields* **2019**, *174*, 1133–1185. [[CrossRef](#)]
49. Barbier, J.; Dia, M.; Macris, N.; Krzakala, F.; Zdeborová, L. Rank-one matrix estimation: Analysis of algorithmic and information theoretic limits by the spatial coupling method. *arXiv* **2018**, arXiv:1812.02537.
50. Alaoui, A.E.; Krzakala, F.; Jordan, M. Fundamental limits of detection in the spiked Wigner model. *Ann. Stat.* **2020**, *48*, 863–885. [[CrossRef](#)]
51. Lesieur, T.; Krzakala, F.; Zdeborová, L. Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *J. Stat. Mech. Theory Exp.* **2017**, *2017*, 073403. [[CrossRef](#)]
52. Pourkamali, F.; Macris, N. Mismatched Estimation of Non-Symmetric Rank-One Matrices Under Gaussian Noise. In Proceedings of the 2022 IEEE International Symposium on Information Theory (ISIT), Espoo, Finland, 26 June–1 July 2022; pp. 1288–1293. [[CrossRef](#)]
53. Verdú, S. Mismatched Estimation and Relative Entropy. *IEEE Trans. Inf. Theory* **2010**, *56*, 3712–3720. [[CrossRef](#)]
54. Chen, W. On the mixed even-spin Sherrington-Kirkpatrick model with ferromagnetic interaction. *Ann. Inst. Henri Poincaré* **2014**, *50*, 63–83. [[CrossRef](#)]
55. Hartman, P. *Ordinary Differential Equations*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1964.
56. Bertsekas, D. *Nonlinear Programming*; Athena Scientific: Nashua, NH, USA, 2003.
57. Camilli, F.; Mézard, M. Matrix factorization with neural networks. *Phys. Rev. E* **2023**, *107*, 064308. [[CrossRef](#)]

-
58. Camilli, F.; Mézard, M. The Decimation Scheme for Symmetric Matrix Factorization. *arXiv* **2023**, arXiv:2307.16564.
 59. Guionnet, A.; Ko, J.; Krzakala, F.; Zdeborová, L. Estimating rank-one matrices with mismatched prior and noise: Universality and large deviations. *arXiv* **2023**, arXiv:2306.09283.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.