

HAMILTON-JACOBI EQUATIONS FOR MEAN-FIELD DISORDERED SYSTEMS

J.-C. MOURRAT

ABSTRACT. We argue that Hamilton-Jacobi equations provide a convenient and intuitive approach for studying the large-scale behavior of mean-field disordered systems. This point of view is illustrated on the problem of inference of a rank-one matrix. We compute the large-scale limit of the free energy by showing that it satisfies an approximate Hamilton-Jacobi equation with asymptotically vanishing viscosity parameter and error term.

1. MOTIVATION

The goal of this paper is to propose a new approach to the computation of the large-scale limit of the free energy of mean-field disordered systems. The new method is based on showing that the finite-volume free energy satisfies an approximate Hamilton-Jacobi equation, with viscosity parameter and error term that vanish in the large-scale limit.

The paper grew out of my attempt to build some intuition for the celebrated Parisi formula for such systems, see [19, 15, 22, 24, 21]. The classical variational formulation of the free energy allows to write this quantity as the supremum of an energy and an entropy terms. For the Sherrington-Kirkpatrick model, the Parisi formula identifies the limit as an infimum instead, defying intuition. The plot thickens further when one considers more complicated systems such as the perceptron and the Hopfield models, which are expected to have limit free energies given by saddle-point variational problems [23].

In this paper, I propose a change of viewpoint that puts the main emphasis on the fact that the free energy satisfies a Hamilton-Jacobi equation, up to a small error. In this new point of view, this is the fundamental property that should be the center of attention and should receive an explanation. As is well-known, if the nonlinearity in the Hamilton-Jacobi equation is convex, then the solution can be expressed as an inf-sup variational problem. This suggests a Hamilton-Jacobi interpretation with convex nonlinearity for the Sherrington-Kirkpatrick model. However, it is unclear why one should expect the nonlinearity to always be convex. In fact, in the model that will be the focus of our attention here, the nonlinearity is concave, not convex. This still allows for a variational representation, but as a sup-inf instead of an inf-sup. More importantly, this suggests that the Hamilton-Jacobi point of view may be more robust and transparent than the variational representations.

Date: April 29, 2020.

2010 Mathematics Subject Classification. 82B44, 82D30.

Key words and phrases. spin glass, statistical inference, Hamilton-Jacobi equation.

The observation that the finite-volume free energy satisfies an approximate Hamilton-Jacobi equation is already present in the influential paper [14] which provided a first step towards the proof of the Parisi formula. However, the idea of relying on this equation to prove the convergence of the free energy in the infinite volume limit has been left largely unexplored, except for a few papers in the physics literature [6, 5]. As was explained there as well as in [20, 8], this idea can easily be made rigorous in the case of the Curie-Weiss model. Although the interactions in this model are not disordered, it is illustrative to explain the main ideas in this simple case.

For the Curie-Weiss model, we would like to compute, for each $t \geq 0$, the large- N limit of the free energy

$$F_N^\circ(t) := \frac{1}{N} \log \sum_{\sigma \in \{\pm 1\}^N} \exp\left(\frac{t}{N} \sum_{i,j=1}^N \sigma_i \sigma_j\right).$$

We aim to do so by identifying a PDE satisfied by $F_N^\circ(t)$, possibly up to error terms that vanish in the large- N limit. However, at this stage we can only calculate derivatives with respect to t and infer information about the distribution of $\sum_{i,j=1}^N \sigma_i \sigma_j$ under the associated Gibbs measure. (For instance, the first and second derivatives are related to the mean and variance of this variable.) In order to find a closed set of equations, we need to “enrich” our free energy by introducing another quantity into the problem. This additional quantity should hopefully be simpler than $\sum_{i,j} \sigma_i \sigma_j$, and display some nontrivial correlations with the latter. In the present case, this quantity is very easy to guess: it is simply the average magnetization $\sum_{i=1}^N \sigma_i$. (In more complicated settings, our intuition can be guided e.g. by cavity calculations.) Although we may a priori only care about calculating $F_N^\circ(t)$, it is thus natural to introduce, for each $t \geq 0$ and $h \in \mathbb{R}$, the enriched free energy

$$F_N(t, h) := \frac{1}{N} \log \sum_{\sigma \in \{\pm 1\}^N} \exp\left(\frac{t}{N} \sum_{i,j=1}^N \sigma_i \sigma_j + h \sum_{i=1}^N \sigma_i\right).$$

Denoting by $\langle \cdot \rangle$ the associated Gibbs measure, we then observe that

$$(1.1) \quad \partial_t F_N = \frac{1}{N^2} \left\langle \sum_{i,j=1}^N \sigma_i \sigma_j \right\rangle \quad \text{and} \quad \partial_h F_N = \frac{1}{N} \left\langle \sum_{i=1}^N \sigma_i \right\rangle,$$

so that

$$\partial_t F_N - (\partial_h F_N)^2 = \frac{1}{N^2} \left\langle \left(\sum_{i=1}^N \sigma_i - \sum_{i=1}^N \langle \sigma_i \rangle \right)^2 \right\rangle.$$

Since the right side in the identity above is a variance, we may expect it to be small. Moreover, since $F_N(t, h)$ encodes complete information on the law of $\sum \sigma_i$, it should be possible to find an expression for this variance in terms of F_N . We find indeed that

$$\partial_t F_N - (\partial_h F_N)^2 = \frac{1}{N} \partial_h^2 F_N.$$

On this simple example, the free energy thus solves an exact Hamilton-Jacobi equation with viscosity term equal to N^{-1} . After observing that the value of

$F_N(0, h)$ does not depend on N , we have completely identified the limit F_∞ of F_N as the viscosity solution to

$$\partial_t F_\infty - (\partial_h F_\infty)^2 = 0.$$

In a nutshell, due to the mean-field character of the model, we expect to be able to identify a handful of quantities whose statistics are related to one another. These relations will produce non-trivial identities between the first derivatives of the free energy: a Hamilton-Jacobi equation. There will be error terms, which one may expect to control by second-order derivatives, since these second-order derivatives are equal to the variances of the quantities of interest.

We aim to carry an argument that has a similar structure for disordered mean-field models¹. However, for the Sherrington-Kirkpatrick and similar models, an important difficulty arises: the number of informative quantities one needs to add to the “enriched” free energy is infinite. In physicists’ language, the system has a functional order parameter. As is well-known, this is bound to create very important technical difficulties. We will thus focus on the simpler setting provided by an inference problem. In this context, an additional symmetry forces the system to be replica-symmetric for every choice of parameters, and thus a simpler argument based on the addition of a single quantity suffices to “close the equation”. We define the model on which we will focus and state our main results in the next section.

2. RANK-ONE MATRIX ESTIMATION, MAIN RESULTS

We consider the problem of estimating a vector $\bar{x} = (\bar{x}_1, \dots, \bar{x}_N) \in \mathbb{R}^N$ of independent entries distributed according to a bounded measure P , given the observations of

$$Y := \sqrt{\frac{t}{N}} \bar{x} \bar{x}^t + W,$$

where $W = (W_{ij})_{1 \leq i, j \leq N}$ are independent standard Gaussian random variables, independent of the vector \bar{x} . We denote the joint law of \bar{x} and W by \mathbb{P} , with associated expectation \mathbb{E} . Note that we seek to recover N parameters from N^2 observations, each with a signal-to-noise ratio of the order of N^{-1} ; this should therefore be the critical scaling for the inference of \bar{x} .

By Bayes’ rule, the posterior distribution of \bar{x} given the observation of Y is the probability measure

$$(2.1) \quad \frac{e^{H_N(t, x)} dP_N(x)}{\int_{\mathbb{R}^N} e^{H_N(t, x')} dP_N(x')},$$

¹It is amusing to note that starting from (1.1), the argument carries verbatim to the case when one adds a random field to the Curie-Weiss model.

where we use the shorthand notation P_N for the product measure $P^{\otimes N}$, and where $H_N(t, x)$ is defined by

$$\begin{aligned}
 H_N(t, x) &:= \sqrt{\frac{t}{N}} \sum_{i,j=1}^N Y_{ij} x_i x_j - \frac{t}{2N} \sum_{i,j=1}^N x_i^2 x_j^2 \\
 &= \sum_{i,j=1}^N \left(\sqrt{\frac{t}{N}} W_{ij} x_i x_j + \frac{t}{N} x_i x_j \bar{x}_i \bar{x}_j - \frac{t}{2N} x_i^2 x_j^2 \right) \\
 (2.2) \quad &= \sqrt{\frac{t}{N}} x \cdot W x + \frac{t}{N} (x \cdot \bar{x})^2 - \frac{t}{2N} |x|^4.
 \end{aligned}$$

This will be explained in more details and in a slightly more general context in Appendix A. Note that although we suppress it from the notation, the quantity $H_N(t, x)$ is random in that it depends on the realization of \bar{x} and W . Throughout the paper, we write $|x|$ to denote the ℓ^2 norm of the vector $x \in \mathbb{R}^N$.

Our goal is to understand the large- N behavior of the normalizing constant in (2.1). The asymptotic behavior of this quantity has already been obtained multiple times in the literature; we refer to [11, 17, 2, 16, 3, 4, 12] for references. As was explained above, the point of the present paper is to devise yet another proof of this result, which centers on the identification of an appropriate Hamilton-Jacobi equation. Naturally, several elements of the proof presented here can also be found in these previous works; the main difference is the global structure of the argument.

In the spirit of the previous section, we start by introducing an “enriched” system. Let $z = (z_i)_{1 \leq i \leq N}$ be a vector of independent standard Gaussian random variables, independent of \bar{x} and W under \mathbb{P} . For every $t, h \geq 0$ and $x \in \mathbb{R}^N$, we define

$$H_N(t, h, x) := \sqrt{\frac{t}{N}} x \cdot W x + \frac{t}{N} (x \cdot \bar{x})^2 - \frac{t}{2N} |x|^4 + \sqrt{h} z \cdot x + h x \cdot \bar{x} - \frac{h}{2} |x|^2.$$

The difference between the quantity above and that in (2.2), namely

$$\sqrt{h} z \cdot x + h x \cdot \bar{x} - \frac{h}{2} |x|^2 = \sum_{i=1}^N \left(\sqrt{h} z_i x_i + h x_i \bar{x}_i - \frac{h}{2} x_i^2 \right),$$

is the energy associated with the much simpler inference problem in which we try to recover $\bar{x} \in \mathbb{R}^N$ from the observation of $\sqrt{h} \bar{x} + z \in \mathbb{R}^N$. We define the free energy

$$(2.3) \quad F_N(t, h) := \frac{1}{N} \log \left(\int_{\mathbb{R}^N} e^{H_N(t, h, x)} dP_N(x) \right),$$

as well as its expectation (with respect to the variables \bar{x} , W and z)

$$(2.4) \quad \bar{F}_N(t, h) := \mathbb{E}[F_N(t, h)].$$

For every $h \geq 0$, we set

$$(2.5) \quad \psi(h) := \mathbb{E} \log \int_{\mathbb{R}} \exp \left(\sqrt{h} z_1 x + h x \bar{x}_1 - \frac{h}{2} x^2 \right) dP(x) = \bar{F}_1(0, h).$$

In this last expression, all the variables are scalar. Observe that $\bar{F}_N(0, h) = \psi(h)$ does not depend on N . Our main goal is to prove the following result.

Theorem 2.1 (Convergence to HJ). *For every $M \geq 1$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\sup_{[0, M]^2} (F_N - f)^2 \right] = 0,$$

where $f(t, h)$ is the viscosity solution of the Hamilton-Jacobi equation

$$(2.6) \quad \begin{cases} \partial_t f - 2(\partial_h f)^2 = 0 & \text{in } (0, +\infty)^2, \\ -\partial_h f = 0 & \text{on } (0, +\infty) \times \{0\}, \end{cases}$$

with initial condition $f(0, h) = \psi(h)$.

The next proposition is the main ingredient for the proof of Theorem 2.1. It states that the averaged free energy satisfies an approximate Hamilton-Jacobi equation with asymptotically vanishing viscosity parameter.

Proposition 2.2 (Approximate HJ in finite volume). *There exists $C < \infty$ such that for every $N \geq 1$ and uniformly over $[0, \infty)^2$,*

$$0 \leq \partial_t \bar{F}_N - 2(\partial_h \bar{F}_N)^2 \leq \frac{2}{N} \partial_h^2 \bar{F}_N + 2\mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2 \right] + \frac{C}{Nh},$$

and moreover,

$$(2.7) \quad \partial_h \bar{F}_N \geq 0.$$

In Proposition 2.2, we kept the variables (t, h) implicit for notational convenience. A more precise statement would be that for every $(t, h) \in [0, \infty)^2$, we have

$$0 \leq \partial_t \bar{F}_N(t, h) - 2(\partial_h \bar{F}_N(t, h))^2 \leq \dots$$

The right side of this inequality is interpreted as $+\infty$ when $h = 0$.

The next section is devoted to the proof of Proposition 2.2. We will also give some basic estimates on the derivatives of F_N and show that F_N is concentrated around its expectation \bar{F}_N . Section 4 starts with the definitions relevant to the notion of viscosity solutions. We then prove Theorem 2.1 using the results of Section 3. The argument is similar to more standard situations for vanishing viscosity limits, although some additional difficulties appear. We close the section by discussing a variational representation for f given by the Hopf-Lax formula. A generalization to tensors of arbitrary order is then obtained in Section 5. In order to make the paper fully self-contained, two appendices are included. In Appendix A, we recall the proof of the Nishimori identity, which is a property of inference problems and is the main technical mechanism that allows to “close the equation” and remain in the replica-symmetric phase. In Appendix B, we prove the comparison principle and the Hopf-Lax formula for viscosity solutions of (2.6).

3. APPROXIMATE HAMILTON-JACOBI EQUATION AND BASIC ESTIMATES

The main purpose of this section is to prove Proposition 2.2. We will also record basic estimates on the derivatives of the free energy and its concentration properties that will be useful in the next section.

We denote by $\langle \cdot \rangle$ the Gibbs measure associated with the energy $H_N(t, h, \cdot)$. That is, for each bounded measurable function $f : \mathbb{R}^N \rightarrow \mathbb{R}$, we set

$$(3.1) \quad \langle f(x) \rangle := \frac{1}{Z_N(t, h)} \int_{\mathbb{R}^N} f(x) e^{H_N(t, h, x)} dP_N(x),$$

where

$$Z_N(t, h) := \int_{\mathbb{R}^N} e^{H_N(t, h, x)} dP_N(x).$$

Note that although the notation does not display it, this random probability measure depends on t, h , as well as on the realization of the random variables \bar{x} , W and z . We will also consider “replicated” (or tensorized) versions of this measure, and write x, x', x'' , etc. for the canonical “replicated” random variables. Conditionally on \bar{x}, W and z , these random variables are independent and each is distributed according to the Gibbs measure $\langle \cdot \rangle$. Abusing notation slightly, we still denote this tensorized measure by $\langle \cdot \rangle$. An important ingredient for the proof of Proposition 2.2 is the Nishimori identity, which is a feature of inference problems whose proof is recalled in Appendix A below. For simplicity of notation, we only state this identity in the case of two or three replicas, since this will be sufficient for our purpose: for every bounded measurable function $f : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$, we have

$$(3.2) \quad \mathbb{E} \langle f(x, x') \rangle = \mathbb{E} \langle f(x, \bar{x}) \rangle,$$

and for every bounded measurable function $f : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$,

$$(3.3) \quad \mathbb{E} \langle f(x, x', x'') \rangle = \mathbb{E} \langle f(x, x', \bar{x}) \rangle.$$

Proof of Proposition 2.2. We decompose the proof into three steps.

Step 1. In this step, we compute the first derivatives of \bar{F}_N . Starting with the derivative with respect to t , we have

$$(3.4) \quad \partial_t F_N(t, h) = \frac{1}{N} \left\langle \frac{1}{2\sqrt{tN}} x \cdot W x + \frac{1}{N} (x \cdot \bar{x})^2 - \frac{1}{2N} |x|^4 \right\rangle.$$

By Gaussian integration by parts, we have for every $i, j \in \{1, \dots, N\}$ that

$$\mathbb{E} \langle W_{ij} x_i x_j \rangle = \mathbb{E} [\partial_{W_{ij}} \langle x_i x_j \rangle] = \sqrt{\frac{t}{N}} \mathbb{E} \langle x_i^2 x_j^2 - x_i x_j x'_i x'_j \rangle,$$

and thus, taking the expectation in (3.4), we get

$$\partial_t \bar{F}_N(t, h) = \frac{1}{2N^2} \mathbb{E} \langle -(x \cdot x')^2 + 2(x \cdot \bar{x})^2 \rangle.$$

Using also the Nishimori identity (3.2), we conclude that

$$(3.5) \quad \partial_t \bar{F}_N(t, h) = \frac{1}{2N^2} \mathbb{E} \langle (x \cdot \bar{x})^2 \rangle.$$

Similarly, since

$$(3.6) \quad \mathbb{E} \langle z_i x_i \rangle = \mathbb{E} [\partial_{z_i} \langle x_i \rangle] = \sqrt{h} \mathbb{E} \langle x_i^2 - x_i x_i' \rangle,$$

we have

$$(3.7) \quad \begin{aligned} \partial_h \bar{F}_N(t, h) &= \frac{1}{N} \mathbb{E} \left\langle \frac{1}{2\sqrt{h}} z \cdot x + x \cdot \bar{x} - \frac{1}{2} |x|^2 \right\rangle \\ &= \frac{1}{N} \mathbb{E} \left\langle \frac{-1}{2} x \cdot x' + x \cdot \bar{x} \right\rangle \\ (3.8) \quad &= \frac{1}{2N} \mathbb{E} \langle x \cdot \bar{x} \rangle. \end{aligned}$$

We thus deduce that

$$(3.9) \quad \partial_t \bar{F}_N - 2(\partial_h \bar{F}_N)^2 = \frac{1}{2N^2} \mathbb{E} \langle (x \cdot \bar{x} - \mathbb{E} \langle x \cdot \bar{x} \rangle)^2 \rangle.$$

In particular, this quantity is non-negative. Note also that

$$(3.10) \quad \partial_h \bar{F}_N(t, h) = \frac{1}{2N} \mathbb{E} \langle x \cdot x' \rangle = \frac{1}{2N} \mathbb{E} [|\langle x \rangle|^2] \geq 0,$$

so that property (2.7) holds.

Step 2. In the remaining two steps, we will control the right side of (3.9) in terms of the quantities $\partial_h^2 \bar{F}_N$ and $\mathbb{E} [(\partial_h F_N - \partial_h \bar{F}_N)^2]$. In this step, we show that these quantities allow for a control of the fluctuations of the variable $H'_N(h, x)$ defined by

$$H'_N(h, x) := \frac{1}{2\sqrt{h}} z \cdot x + x \cdot \bar{x} - \frac{1}{2} |x|^2.$$

More precisely, we show that

$$(3.11) \quad \begin{aligned} &\mathbb{E} \langle (H'_N(h, x) - \mathbb{E} \langle H'_N(h, x) \rangle)^2 \rangle \\ &\leq N \partial_h^2 \bar{F}_N(t, h) + N^2 \mathbb{E} \left[(\partial_h F_N(t, h) - \partial_h \bar{F}_N(t, h))^2 \right] + CNh^{-1}. \end{aligned}$$

Our starting point is the variance decomposition

$$\begin{aligned} &\mathbb{E} \langle (H'_N(h, x) - \mathbb{E} \langle H'_N(h, x) \rangle)^2 \rangle \\ &= \mathbb{E} \langle (H'_N(h, x) - \langle H'_N(h, x) \rangle)^2 \rangle + \mathbb{E} [(\langle H'_N(h, x) \rangle - \mathbb{E} \langle H'_N(h, x) \rangle)^2]. \end{aligned}$$

Since

$$(3.12) \quad \partial_h F_N(t, h) = \frac{1}{N} \langle H'_N(h, x) \rangle,$$

and $\bar{F}_N = \mathbb{E}[F_N]$, we readily have that

$$\mathbb{E} [(\langle H'_N(h, x) \rangle - \mathbb{E} \langle H'_N(h, x) \rangle)^2] = N^2 \mathbb{E} [(\partial_h F_N(t, h) - \partial_h \bar{F}_N(t, h))^2].$$

We also have that

$$(3.13) \quad \partial_h^2 F_N(t, h) = \frac{1}{N} (\langle (H'_N(h, x))^2 \rangle - \langle H'_N(h, x) \rangle^2) - \frac{1}{4Nh^{\frac{3}{2}}} \langle z \cdot x \rangle,$$

and thus, taking expectations and using (3.6), we get

$$\partial_h^2 \bar{F}_N(t, h) = \frac{1}{N} \mathbb{E} \left[\langle (H'_N(h, x))^2 \rangle - \langle H'_N(h, x) \rangle^2 \right] - \frac{1}{4Nh} \mathbb{E} \langle |x|^2 - x \cdot x' \rangle.$$

Recall that we assume that the measure P has bounded support. This implies that the last term in the display above is bounded by Ch^{-1} , and thus yields (3.11).

Step 3. In order to conclude, there remains to show that the variance of $x \cdot \bar{x}$ is controlled by that of $H'_N(h, x)$. We show that

$$(3.14) \quad \mathbb{E} \langle (x \cdot \bar{x} - \mathbb{E} \langle x \cdot \bar{x} \rangle)^2 \rangle \leq 4 \mathbb{E} \langle (H'_N(h, x) - \mathbb{E} \langle H'_N(h, x) \rangle)^2 \rangle.$$

In view of (3.7) and (3.8), it suffices to show that

$$\mathbb{E} \langle (x \cdot \bar{x})^2 \rangle \leq 4 \mathbb{E} \langle H'_N(h, x)^2 \rangle.$$

For every $i \neq j$, we have, using Gaussian integration by parts and the Nishimori identity,

$$\mathbb{E} \langle z_i z_j x_i x_j \rangle = \sqrt{h} \mathbb{E} \langle z_j x_i x_j (x_i - x'_i) \rangle = h \mathbb{E} \langle x_i x_j (x_i - x'_i) (x_j + x'_j - 2\bar{x}_j) \rangle,$$

while for $i = j$,

$$\mathbb{E} \langle z_i^2 x_i^2 \rangle = \sqrt{h} \mathbb{E} \langle z_i x_i^2 (x_i - x'_i) \rangle + \mathbb{E} \langle x_i^2 \rangle = h \mathbb{E} \langle x_i^2 (x_i - x'_i) (x_i + x'_i - 2\bar{x}_i) \rangle + \mathbb{E} \langle x_i^2 \rangle.$$

As a consequence,

$$\begin{aligned} & \mathbb{E} \left\langle \left(\frac{1}{2\sqrt{h}} z \cdot x \right)^2 \right\rangle \\ &= \frac{1}{4} \left(\mathbb{E} \langle |x|^4 - 2|x|^2(x \cdot \bar{x}) - (x \cdot \bar{x})^2 + 2(x \cdot \bar{x})(x \cdot x') \rangle + \frac{1}{h} \mathbb{E} \langle |x|^2 \rangle \right). \end{aligned}$$

Similarly,

$$\begin{aligned} & \mathbb{E} \left\langle \frac{2}{2\sqrt{h}} z \cdot x \left(x \cdot \bar{x} - \frac{1}{2}|x|^2 \right) \right\rangle \\ &= \mathbb{E} \left\langle |x|^2 \left(x \cdot \bar{x} - \frac{1}{2}|x|^2 \right) \right\rangle - \mathbb{E} \left\langle (x \cdot x') \left(x \cdot \bar{x} - \frac{1}{2}|x|^2 \right) \right\rangle. \end{aligned}$$

We therefore obtain that

$$\begin{aligned} \mathbb{E} \langle H'_N(h, x)^2 \rangle &= \mathbb{E} \left\langle \left(\frac{1}{2\sqrt{h}} z \cdot x \right)^2 \right\rangle + \mathbb{E} \left\langle \frac{2}{2\sqrt{h}} z \cdot x \left(x \cdot \bar{x} - \frac{1}{2}|x|^2 \right) \right\rangle \\ &\quad + \mathbb{E} \langle (x \cdot \bar{x})^2 \rangle + \frac{1}{4} \mathbb{E} \langle |x|^4 \rangle - \mathbb{E} \langle (x \cdot \bar{x}) |x|^2 \rangle \\ &= \frac{3}{4} \mathbb{E} \langle (x \cdot \bar{x})^2 \rangle - \frac{1}{2} \mathbb{E} \langle (x \cdot \bar{x})(x \cdot x') \rangle + \frac{1}{4h} \mathbb{E} \langle |x|^2 \rangle. \end{aligned}$$

By the Cauchy-Schwarz inequality and the Nishimori identity, we have

$$|\mathbb{E} \langle (x \cdot \bar{x})(x \cdot x') \rangle| \leq \mathbb{E} \langle (x \cdot \bar{x})^2 \rangle,$$

and thus (3.14) is proved. \square

Before turning to the proof of Theorem 2.1, we record simple derivative and concentration estimates in the next two lemmas. We use the notation

$$|W| := \sup\{|Wx| : x \in \mathbb{R}^N, |x| \leq 1\}.$$

Of course, this quantity depends on N , and as we will see in the proof of Lemma 3.2, it grows like \sqrt{N} . The notation may be slightly misleading, in that it does not display the dependency on N . A similar convention is already in place when we write $|x|$ to denote the ℓ^2 norm of the vector $x \in \mathbb{R}^N$, a quantity which is typically of the order of \sqrt{N} .

Lemma 3.1 (Derivative estimates). *There exists a constant $C < \infty$ such that the following estimates hold uniformly over $[0, \infty)^2$:*

$$(3.15) \quad |\partial_t \bar{F}_N| + |\partial_h \bar{F}_N| \leq C,$$

$$(3.16) \quad |\partial_t F_N| \leq C + \frac{C|W|}{\sqrt{Nt}},$$

$$(3.17) \quad |\partial_h F_N| \leq C + \frac{C|z|}{\sqrt{Nh}},$$

$$(3.18) \quad \partial_h^2 F_N \geq -\frac{C|z|}{N^{\frac{1}{2}}h^{\frac{3}{2}}}.$$

Proof. Recall that the measure P has bounded support. The estimates in (3.15), (3.16), (3.17), and (3.18) are thus immediate consequences of (3.5)-(3.8), (3.4), (3.12), and (3.13) respectively. \square

We now turn to a concentration estimate. Since this is sufficient for our purposes, we simply state an L^2 bound in the probability space, and prove it using the elementary Efron-Stein inequality. The statement could be strengthened to a Gaussian-type integrability using concentration results such as [7, Theorems 5.5 and 6.2] (and this also allows to improve the rate of decay to 0 as N tends to infinity).

Lemma 3.2 (Concentration of free energy). *There exists $C < \infty$ such that for every $M \geq 1$ and $N \in \mathbb{N}$,*

$$\mathbb{E} \left[\sup_{[0, M]^2} (F_N - \bar{F}_N)^2 \right] \leq CM^{\frac{4}{3}} N^{-\frac{1}{3}}.$$

Proof. We recall that \mathbb{E} denotes the expectation with respect to the random variables \bar{x} , W and z . For each integer i , we write $\mathbb{E}[\cdot | \{\bar{x}_i\}^c]$ to denote the expectation with respect to all the variables except \bar{x}_i . In other words, $\mathbb{E}[\cdot | \{\bar{x}_i\}^c]$ denotes the conditional expectation with respect to the σ -algebra generated by the random variables $(\bar{x}_j)_{j \neq i}$, W and z . We define $\mathbb{E}[\cdot | \{z_i\}^c]$

and $\mathbb{E}[\cdot | \{W_{ij}\}^c]$ similarly. The Efron-Stein inequality and (2.4) give us that

$$\begin{aligned} \mathbb{E}\left[(F_N - \bar{F}_N)^2\right] &\leq \sum_{1 \leq i, j \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{W_{ij}\}^c])^2\right] \\ &\quad + \sum_{1 \leq i \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{z_i\}^c])^2\right] \\ &\quad + \sum_{1 \leq i \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{\bar{x}_i\}^c])^2\right]. \end{aligned}$$

By the Gaussian Poincaré inequality (see e.g. [9, (2.5)] or [1]), we have

$$\mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{W_{ij}\}^c])^2\right] \leq \mathbb{E}\left[(\partial_{W_{ij}} F_N)^2\right],$$

and

$$\mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{z_i\}^c])^2\right] \leq \mathbb{E}\left[(\partial_{z_i} F_N)^2\right].$$

Moreover,

$$\partial_{W_{ij}} F_N = t^{\frac{1}{2}} N^{-\frac{3}{2}} \langle x_i x_j \rangle,$$

and

$$\partial_{z_i} F_N = h^{\frac{1}{2}} N^{-1} \langle x_i \rangle,$$

so that

$$\sum_{1 \leq i, j \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{W_{ij}\}^c])^2\right] + \sum_{1 \leq i \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{z_i\}^c])^2\right] \leq \frac{C(t+h)}{N}.$$

Since

$$\partial_{\bar{x}_i} F_N = \frac{1}{N} \left\langle \frac{2t}{N} x_i (x \cdot \bar{x}) + h x_i \right\rangle$$

is bounded by $C(t+h)/N$, and the support of the law of \bar{x}_i is bounded, we also have that

$$\sum_{1 \leq i \leq N} \mathbb{E}\left[(F_N - \mathbb{E}[F_N | \{\bar{x}_i\}^c])^2\right] \leq \frac{C}{N}(t+h)^2.$$

We have thus shown that there exists $C < \infty$ such that for every $(t, h) \in [0, \infty)^2$,

$$(3.19) \quad \mathbb{E}\left[(F_N - \bar{F}_N)^2(t, h)\right] \leq \frac{C}{N}(t + t^2 + h + h^2).$$

In order to complete the proof, there remains to use a regularity estimate for $F_N - \bar{F}_N$. By Lemma 3.1, for every $t, t', h, h' \geq 0$ satisfying

$$|t - t'| + |h - h'| \leq 1,$$

we have

$$|F_N(t, h) - F_N(t', h')| \leq C \left(1 + \frac{|W|}{\sqrt{N}} + \frac{|z|}{\sqrt{N}}\right) (|t - t'|^{\frac{1}{2}} + |h - h'|^{\frac{1}{2}}).$$

On the other hand, it is clear from Lemma 3.1 that \bar{F}_N is uniformly Lipschitz continuous, so in particular the estimate above also holds if F_N is replaced by \bar{F}_N . Hence, for any $\varepsilon \in (0, 1]$, if we set

$$A_\varepsilon := \varepsilon \mathbb{N}^2 = \{0, \varepsilon, 2\varepsilon, \dots\}^2,$$

then

$$\sup_{[0, \infty)^2} |F_N - \bar{F}_N| - \sup_{A_\varepsilon} |F_N - \bar{F}_N| \leq C \left(1 + \frac{|W|}{\sqrt{N}} + \frac{|z|}{\sqrt{N}}\right) \sqrt{\varepsilon}.$$

Moreover, for every $M \in [1, \infty)$, we have by (3.19) that

$$\begin{aligned} \mathbb{E} \left[\sup_{A_\varepsilon \cap [0, M]^2} (F_N - \bar{F}_N)^2 \right] &\leq \sum_{(t, h) \in A_\varepsilon \cap [0, M]^2} \mathbb{E} [(F_N - \bar{F}_N)^2(t, h)] \\ &\leq CM^4 \varepsilon^{-2} N^{-1}. \end{aligned}$$

Combining the two previous displays yields

$$\mathbb{E} \left[\sup_{[0, M]^2} (F_N - \bar{F}_N)^2 \right] \leq C\varepsilon \mathbb{E} \left[1 + \frac{|W|^2}{N} + \frac{|z|^2}{N} \right] + \frac{CM^4}{N\varepsilon^2},$$

and we clearly have $\mathbb{E}[|z|^2] = N$. In order to conclude, there remains to verify that $\mathbb{E}[|W|^2] \leq CN$ (and then choose $\varepsilon = M^{\frac{4}{3}} N^{-\frac{1}{3}}$). For every fixed $x \in \mathbb{R}^N$ satisfying $|x| \leq 1$ and every $i \in \{1, \dots, N\}$, we have that $(Wx)_i$ is a centered Gaussian random variable with variance $|x|^2 \leq 1$, and moreover, the random variables $((Wx)_i)_{1 \leq i \leq N}$ are independent. We deduce that there exists $C < \infty$ such that for every $|x| \leq 1$,

$$\mathbb{E} [\exp(C^{-1}|Wx|^2)] \leq C \exp(CN),$$

and thus by the Chebyshev inequality, after enlarging $C < \infty$ if necessary, we have that for every $a \geq C$,

$$\mathbb{P} [|Wx|^2 \geq aN] \leq \exp\left(-\frac{aN}{C}\right).$$

Now, let $A \subseteq \mathbb{R}^N$ be a finite set such that any two points in A are at distance at least $1/2$ from one another, and no point of $\{|x| \leq 1\}$ can be added to A without violating this property. By this property of maximality, it must be that for every x satisfying $|x| \leq 1$, there exists $y \in A$ such that $|x - y| \leq 1/2$. Since for any $x, y \in \mathbb{R}^N$, we have

$$|Wx - Wy| \leq |W| |x - y|,$$

it follows that

$$|W| = \sup_{|x| \leq 1} |Wx| \leq \sup_{x \in A} |Wx| + \frac{|W|}{2},$$

and thus

$$|W| \leq 2 \sup_{x \in A} |Wx|.$$

In order to construct such a set A , we simply pick points in $\{|x| \leq 1\}$ in some arbitrary manner, until the maximality property is reached. Note that the balls centered at each of the points in A and of radius $1/4$ are disjoint; they are also contained in the ball of radius $5/4$. Computing the volume of these sets, we infer that $|A| \leq 5^N$, and thus, by a union bound, we have for every $a \geq C$ that

$$\mathbb{P} [|W|^2 \geq aN] \leq 2 \exp\left(-N \left(\frac{a}{C} - \log 5\right)\right).$$

This implies in particular that $\mathbb{E}[|W|^2] \leq CN$, as desired. \square

4. CONVERGENCE TO VISCOSITY SOLUTION

The main goal of this section is to show that Proposition 2.2 implies Theorem 2.1. We also comment on variational representations for solutions of Hamilton-Jacobi equations at the end of the section. To start with, we recall the definition of viscosity solutions.

Definition 4.1. We say that a function $f \in C([0, \infty)^2)$ is a *viscosity subsolution* of (2.6) if for every $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ such that (t, h) is a local maximum of $f - \phi$, we have

$$\begin{cases} (\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \leq 0 & \text{if } h > 0, \\ \min(-\partial_h \phi, \partial_t \phi - 2(\partial_h \phi)^2)(t, h) \leq 0 & \text{if } h = 0. \end{cases}$$

We say that a function $f \in C([0, \infty)^2)$ is a *viscosity supersolution* of (2.6) if for every $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ such that (t, h) is a local minimum of $f - \phi$, we have

$$\begin{cases} (\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \geq 0 & \text{if } h > 0, \\ \max(-\partial_h \phi, \partial_t \phi - 2(\partial_h \phi)^2)(t, h) \geq 0 & \text{if } h = 0. \end{cases}$$

We say that a function $f \in C([0, \infty)^2)$ is a *viscosity solution* of (2.6) if it is a viscosity sub- and supersolution. We may also say that a function $f \in C([0, \infty)^2)$ is a viscosity solution of

$$(4.1) \quad \begin{cases} \partial_t f - 2(\partial_h f)^2 \leq 0 & \text{in } (0, +\infty)^2, \\ -\partial_h f \leq 0 & \text{on } (0, +\infty) \times \{0\}, \end{cases}$$

if it is a viscosity subsolution of (2.6). Similarly, we may say that a function $f \in C([0, \infty)^2)$ is a viscosity solution of (4.1) with the inequalities reversed if it is a viscosity supersolution of (2.6).

The mechanism allowing to identify uniquely the viscosity solution to (2.6) subject to appropriate initial condition relies on the following classical comparison principle.

Proposition 4.2 (Comparison principle). *Let u be a subsolution and v be a supersolution of (2.6) such that both u and v are uniformly Lipschitz continuous in the variable h . We have*

$$\sup_{[0, \infty)^2} (u - v) = \sup_{\{0\} \times [0, \infty)} (u - v).$$

The proof of Proposition 4.2 is given in Appendix B. (Besides the inconvenience that the domain under consideration is unbounded, the proof is classical.) In the statement of Proposition 4.2, we assume a uniform Lipschitz continuity property in the variable h . As will be clear from the proof, this assumption can be weakened, and possibly be removed. This assumption is meant to allow for a simpler proof, and is not causing additional difficulties elsewhere since it is very easy to check that our candidate solutions satisfy it.

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. By Lemma 3.2, it suffices to study the convergence of \overline{F}_N as N tends to infinity. Recall that $\overline{F}_N(h, 0) = \psi(h)$ does not depend on N . Moreover, it is clear from (3.4) and (3.8) that \overline{F}_N is uniformly Lipschitz in both variables. Hence, by the Arzelá-Ascoli theorem, the sequence (\overline{F}_N) is precompact for the topology of local uniform convergence. Let f be such that \overline{F}_N converges to f locally uniformly as N tends to infinity along a subsequence. For notational convenience, we will omit to refer to the particular subsequence along which this convergence holds. Our goal is to show that f is a viscosity solution of (2.6). By the comparison principle (Proposition 4.2), this would identify f uniquely, and thus prove the theorem.

We decompose the rest of the proof into six steps.

Step 1. We show that f is a viscosity supersolution of (2.6). It is easy to show that in the definition of viscosity supersolution, replacing the phrase “local minimum” by “strict local minimum” yields an equivalent definition. Let $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ be such that $f - \phi$ has a strict local minimum at the point (t, h) . Since \overline{F}_N converges to f locally uniformly, there exists a sequence $(t_N, h_N) \in (0, \infty) \times [0, \infty)$ converging to (t, h) as N tends to infinity and such that $\overline{F}_N - \phi$ has a strict local minimum at (t_N, h_N) . If $h_N > 0$ infinitely often, then along a subsequence on which this property holds, we have that the first derivatives of \overline{F}_N and ϕ at (t_N, h_N) coincide, and thus by Proposition 2.2 that

$$(\partial_t \phi - 2(\partial_h \phi)^2)(t_N, h_N) \geq 0.$$

By continuity, this implies that

$$(\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \geq 0,$$

as desired. There remains to consider the case when $h_N = 0$ infinitely often. In this case, we must have $h = 0$. We can also assert that

$$(4.2) \quad \partial_h(\overline{F}_N - \phi)(t_N, h_N) \geq 0, \quad \partial_t(\overline{F}_N - \phi)(t_N, h_N) = 0,$$

and we recall that, by Proposition 2.2,

$$(4.3) \quad (\partial_t \overline{F}_N - 2(\partial_h \overline{F}_N)^2)(t_N, h_N) \geq 0.$$

If $-\partial_h \phi(t, h) \geq 0$, then there is nothing to show. Otherwise, using the first statement in (4.2), we find that

$$(\partial_h \phi)^2(t, h) = \lim_{N \rightarrow \infty} (\partial_h \phi)^2(t_N, h_N) \leq \liminf_{N \rightarrow \infty} (\partial_h \overline{F}_N)^2(t_N, h_N),$$

and thus, using also the second statement in (4.2) and (4.3),

$$(\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \geq \limsup_{N \rightarrow \infty} (\partial_t \overline{F}_N - 2(\partial_h \overline{F}_N)^2)(t_N, h_N) \geq 0.$$

This completes the proof of the fact that f is a supersolution.

Step 2. We next show that f is a subsolution of (2.6). In this step, we focus on contact points of the form $(t, 0)$; that is, we give ourselves $t > 0$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ such that $f - \phi$ has a strict local maximum at the point $(t, 0)$. In this case, there exists a sequence $(t_N, h_N) \in (0, \infty) \times [0, \infty)$ converging

to $(t, 0)$ and such that $\overline{F}_N - \phi$ has a local maximum at (t_N, h_N) . If $h_N = 0$, then we must have that

$$\partial_h(\overline{F}_N - \phi)(t_N, h_N) \leq 0.$$

This inequality still holds, and is in fact an equality, if $h_N > 0$. In view of (2.7), we thus deduce that

$$-\partial_h \phi(t_N, h_N) \leq 0.$$

Letting N tend to infinity, we obtain that $-\partial_h \phi(t, 0) \leq 0$, as desired.

Step 3. We now consider the remaining possible contact points. Let $t, h > 0$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ be such that $f - \phi$ has a local maximum at the point (t, h) . For the remainder of this proof, we allow the value of the constant $C < \infty$ to change from place to place, and to depend on t, h, f and ϕ , without further notice. For convenience, we introduce the notation

$$(4.4) \quad \delta_N := \|\overline{F}_N - f\|_{L^\infty([0, t+1] \times [0, h+1])}^{\frac{1}{4}} + N^{-\frac{1}{12}},$$

and

$$\tilde{\phi}(t', h') := \phi(t', h') + (t - t')^2 + (h - h')^2.$$

We clearly have that $f - \tilde{\phi}$ has a strict local maximum at (t, h) . We also have that for every $(t', h') \in [0, t+1] \times [0, h+1]$,

$$(\overline{F}_N - \tilde{\phi})(t', h') \leq (f - \phi)(t', h') - (t - t')^2 - (h - h')^2 + \delta_N^4,$$

while

$$(\overline{F}_N - \tilde{\phi})(t, h) \geq (f - \phi)(t, h) - \delta_N^4.$$

Since $f - \phi$ has a local maximum at (t, h) , we infer that for N sufficiently large, the function $\overline{F}_N - \tilde{\phi}$ has a local maximum at (t_N, h_N) satisfying

$$(4.5) \quad (t - t_N)^2 + (h - h_N)^2 \leq 2\delta_N^4.$$

The point of replacing ϕ by $\tilde{\phi}$ was precisely to obtain such an explicit estimate. We have

$$(4.6) \quad \partial_h(\overline{F}_N - \tilde{\phi})(t_N, h_N) = 0, \quad \partial_t(\overline{F}_N - \tilde{\phi})(t_N, h_N) = 0.$$

We next wish to use Proposition 2.2 to conclude. However, since the concentration result in Lemma 3.2 applies to $F_N - \overline{F}_N$ rather than its derivatives in h , we will want to take a small local average in the h variable to control the term involving $\partial_h(F_N - \overline{F}_N)$. In preparation for this, we show in this step that there exists a constant $C < \infty$ such that for every $h' \in \mathbb{R}$ satisfying $|h' - h_N| \leq C^{-1}$, we have

$$(4.7) \quad |\overline{F}_N(t_N, h') - \overline{F}_N(t_N, h_N) - (h' - h_N)\partial_h \overline{F}_N(t_N, h_N)| \leq C(h' - h_N)^2.$$

We start by writing Taylor's formula

$$(4.8) \quad \begin{aligned} & \overline{F}_N(t_N, h') - \overline{F}_N(t_N, h_N) \\ &= (h' - h_N)\partial_h \overline{F}_N(t_N, h_N) + \int_{h_N}^{h'} (h' - h'')\partial_h^2 \overline{F}_N(t_N, h'') dh''. \end{aligned}$$

The same identity also holds with \bar{F}_N replaced by $\tilde{\phi}$. Since $\bar{F}_N - \tilde{\phi}$ has a local maximum at (t_N, h_N) , and in view of (4.6), we get that for $|h' - h_N| \leq C^{-1}$,

$$\int_{h_N}^{h'} (h' - h'') \partial_h^2 \bar{F}_N(t_N, h'') dh'' \leq \int_{h_N}^{h'} (h' - h'') \partial_h^2 \tilde{\phi}(t_N, h'') dh''.$$

Moreover, the integral on the right side is bounded by $C(h' - h_N)^2$, since $\tilde{\phi}$ is assumed to be smooth. By (3.18), we also have that $\partial_h^2 \bar{F}_N \geq -C$, and thus

$$\int_{h_N}^{h'} (h' - h'') |\partial_h^2 \bar{F}_N(t_N, h'')| dh'' \leq C(h' - h_N)^2.$$

Inequality (4.7) then follows using (4.8) once more.

Step 4. We set

$$G_N(t', h') := \delta_N^{-1} \int_{h'}^{h'+\delta_N} \bar{F}_N(t', h'') dh''.$$

It is clear that the function G_N converges to f locally uniformly as N tends to infinity. Hence, there exists a sequence $t'_N, h'_N > 0$ such that for every N sufficiently large, the function $G_N - \tilde{\phi}$ has a local maximum at (t'_N, h'_N) . Repeating the argument of the previous step, we also obtain that

$$(4.9) \quad (t - t'_N)^2 + (h - h'_N)^2 \leq 2\delta_N^4.$$

We note that

$$(4.10) \quad \partial_h(G_N - \tilde{\phi})(t'_N, h'_N) = 0, \quad \partial_t(G_N - \tilde{\phi})(t'_N, h'_N) = 0,$$

and

$$(4.11) \quad \partial_h^2(G_N - \tilde{\phi})(t'_N, h'_N) \leq 0.$$

Recall from Proposition 2.2 that for every $h' > 0$,

$$(4.12) \quad (\partial_t \bar{F}_N - 2(\partial_h \bar{F}_N)^2)(t'_N, h') \leq \frac{2}{N} \partial_h^2 \bar{F}_N(t'_N, h') \\ + 2\mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2(t'_N, h') \right] + \frac{C}{Nh'}.$$

In the next two steps, we will show the following estimates:

$$(4.13) \quad \int_{h'_N}^{h'_N+\delta_N} (\partial_h \bar{F}_N(t'_N, h') - \partial_h G_N(t'_N, h'_N))^2 dh' \leq C\delta_N^2,$$

and

$$(4.14) \quad \int_{h'_N}^{h'_N+\delta_N} \mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2(t'_N, h') \right] dh' \leq C\delta_N^2.$$

For now, we assume that these estimates hold and show how to conclude. Using the fact that $\partial_h \bar{F}_N$ is bounded, Jensen's inequality, and (4.13), we obtain that

$$(4.15) \quad \left| (\partial_h G_N)^2(t'_N, h'_N) - \delta_N^{-1} \int_{h'_N}^{h'_N+\delta_N} (\partial_h \bar{F}_N)^2(t'_N, h') dh' \right| \\ \leq C\delta_N^{-1} \int_{h'_N}^{h'_N+\delta_N} |\partial_h \bar{F}_N(t'_N, h') - \partial_h G_N(t'_N, h'_N)| dh' \leq C\sqrt{\delta_N}.$$

Averaging over $h' \in [h'_N, h'_N + \delta_N]$ in (4.12), using the estimate above and (4.14), we get

$$(\partial_t G_N - 2(\partial_h G_N)^2)(t'_N, h'_N) \leq \frac{2}{N} \partial_h^2 G_N(t'_N, h'_N) + C\sqrt{\delta_N}.$$

Appealing to (4.10)-(4.11), passing to the limit $N \rightarrow \infty$ and recalling that the first derivatives of ϕ and $\tilde{\phi}$ coincide at (t, h) , we conclude that

$$(4.16) \quad (\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \leq 0,$$

as desired.

Step 5. In order to complete the proof, there remains to show (4.13) and (4.14). In this step, we prove (4.14). The argument relies on the fact that, by integration by parts, we have for any smooth function $g \in C^\infty([a, b], \mathbb{R})$ that

$$(4.17) \quad \begin{aligned} \|g'\|_{L^2(a,b)}^2 &= g(b)g'(b) - g(a)g'(a) - \int_a^b gg'' \\ &\leq g(b)g'(b) - g(a)g'(a) + \|g\|_{L^\infty(a,b)} \|g''\|_{L^1(a,b)}. \end{aligned}$$

Applying (4.17) with $g = F_N - \bar{F}_N$, using that $\partial_h F_N$ is locally bounded and the Cauchy-Schwarz inequality, we get that the left side of (4.14) is bounded by

$$\begin{aligned} &\mathbb{E} \left[\sup_{h' \in [h'_N, h'_N + \delta_N]} (F_N(t'_N, h') - \bar{F}_N(t'_N, h'))^2 \right]^{\frac{1}{2}} \\ &\quad \times \left(C + \mathbb{E} \left[\left(\int_{h'_N}^{h'_N + \delta_N} |\partial_h^2 (F_N - \bar{F}_N)(t'_N, h')| dh' \right)^2 \right]^{\frac{1}{2}} \right). \end{aligned}$$

By Lemma 3.2, the first term in this product is bounded by $CN^{-\frac{1}{6}}$. For the second term, we use (3.18) to observe that, for the constant $C = C_0$ identified there,

$$\begin{aligned} |\partial_h^2 F_N(t'_N, h')| &\leq \left| \partial_h^2 F_N(t'_N, h') + \frac{C_0|z|}{N^{\frac{1}{2}}h^{\frac{3}{2}}} \right| + \frac{C_0|z|}{N^{\frac{1}{2}}h^{\frac{3}{2}}} \\ &= \partial_h^2 F_N(t'_N, h') + \frac{2C_0|z|}{N^{\frac{1}{2}}h^{\frac{3}{2}}}, \end{aligned}$$

and thus, using again that $\partial_h F_N$ is locally bounded, we conclude that

$$\mathbb{E} \left[\left(\int_{h'_N}^{h'_N + \delta_N} |\partial_h^2 F_N(t'_N, h')| dh' \right)^2 \right]^{\frac{1}{2}} \leq C + \frac{C\delta_N}{h^{\frac{3}{2}}} \leq C.$$

Since $N^{-\frac{1}{6}} \leq \delta_N^2$, this completes the proof of (4.14).

Step 6. We now prove (4.13). Observe that

$$\partial_h G_N(t'_N, h'_N) = \delta_N^{-1} (\bar{F}_N(t'_N, h'_N + \delta_N) - \bar{F}_N(t'_N, h'_N)).$$

We use (4.17) with g replaced by the function

$$g_N : h' \mapsto \bar{F}_N(t'_N, h') - \bar{F}_N(t'_N, h'_N) - (h' - h'_N) \partial_h G_N(t'_N, h'_N)$$

to get that

$$\begin{aligned} \int_{h'_N}^{h'_N+\delta_N} (\partial_h \bar{F}_N(t'_N, h') - \partial_h G_N(t'_N, h'_N))^2 dh' \\ \leq \|g_N\|_{L^\infty(h'_N, h'_N+\delta_N)} \int_{h'_N}^{h'_N+\delta_N} |\partial_h^2 \bar{F}_N(t'_N, h')| dh'. \end{aligned}$$

Using also (3.18) and (3.15), we obtain that

$$\int_{h'_N}^{h'_N+\delta_N} (\partial_h \bar{F}_N(t'_N, h') - \partial_h G_N(t'_N, h'_N))^2 dh' \leq C \|g_N\|_{L^\infty(h'_N, h'_N+\delta_N)}.$$

Since \bar{F}_N is Lipschitz continuous in the variable t , we also have that

$$\begin{aligned} \|g_N\|_{L^\infty(h'_N, h'_N+\delta_N)} \leq C |t_N - t'_N| \\ + \sup_{h' \in [h'_N, h'_N+\delta_N]} |\bar{F}_N(t_N, h') - \bar{F}_N(t_N, h'_N) - (h' - h'_N) \partial_h G_N(t_N, h'_N)|. \end{aligned}$$

The estimate (4.13) then follows using (4.5), (4.9) and (4.7). \square

We conclude this section with some remarks on variational representations for the function f appearing in Theorem 2.1. Solutions to Hamilton-Jacobi equations of the form $\partial_t f - \mathbf{H}(\nabla f)$ with convex (resp. concave) \mathbf{H} have a variational representation given by the Hopf-Lax formula, in which the convex (resp. concave) dual of \mathbf{H} appears (see e.g. [13, Theorem 10.3.4.3]). It is usually under this variational presentation that the limit free energy of mean-field statistical mechanics models is identified. In our case, the function \mathbf{H} is simply $p \mapsto 2p^2$, whose convex dual is $q \mapsto \frac{q^2}{8}$.

Proposition 4.3 (Hopf-Lax formula). *For every $t \geq 0$ and $h \geq 0$, we set*

$$(4.18) \quad f(t, h) := \sup_{h' \geq 0} \left(\psi(h') - \frac{(h' - h)^2}{8t} \right),$$

with the understanding that $f(0, h) = \psi(h)$. The function f is the unique viscosity solution of (2.6) that satisfies $f(0, h) = \psi(h)$ and is globally Lipschitz continuous in the variable h .

For completeness, we provide a proof of this classical result in Appendix B. Denoting $\mathbf{H}(p) := 2p^2$ and $\mathbf{H}^*(q) := \frac{q^2}{8}$, we have the following equivalent expressions for f which may be of interest:

$$\begin{aligned} f(t, h) &= \sup_{h' \geq 0} \left(\psi(h') - t \mathbf{H}^* \left(\frac{h' - h}{t} \right) \right) \\ &= \sup_{h' \geq 0} \inf_{p \in \mathbb{R}} \left(\psi(h') - t \left(p \frac{h' - h}{t} - \mathbf{H}(p) \right) \right) \\ &= \sup_{h' \geq 0} \inf_{p \in \mathbb{R}} (\psi(h') + p(h - h') + t \mathbf{H}(p)). \end{aligned}$$

We stress that the proof of Theorem 2.1 does not require that f be identified by such a variational presentation. Moreover, the analysis of f itself does not necessarily require explicit usage of this formula. For instance, assuming for convenience that \bar{x}_1 has mean zero, if one wants to observe that $\partial_h f(t, 0) = 0$ for small values of $t \geq 0$, which at least on a heuristic level corresponds to a

regime where there is no correlation between x and \bar{x} , see (3.8), then we may proceed as follows. First, we check that there exists a constant $C < \infty$ such that for every $h \geq 0$, we have $\psi(h) \leq Ch^2$. (We see from (B.17) that $\psi'(0) = 0$, and next verify that ψ'' is bounded using Gaussian integration by parts.) We next observe that the function

$$(t, h) \mapsto \frac{Ch^2}{1 - 8Ct}$$

is a supersolution of (2.6) on $(0, (8C)^{-1}) \times [0, \infty)$, and thus, by the comparison principle, the solution f to (2.6) remains below this supersolution. Since the null function is a subsolution, we deduce that $\partial_h f(t, 0) = 0$ for every $t < (8C)^{-1}$.

5. EXTENSION TO TENSORS

We now explain how to adapt the method to tensors of arbitrary order. In this setting, the result was obtained in [18, 3]. One motivation for exploring this generalization is that some methods, such as that used in [12], do not seem to generalize well to tensors of odd order.

We fix an integer $p \geq 1$. Generalizing the previous setting, we consider the problem of estimating the vector $\bar{x} = (\bar{x}_1, \dots, \bar{x}_N) \in \mathbb{R}^N$ given the observation of

$$\frac{\sqrt{t}}{N^{\frac{p-1}{2}}} \bar{x}^{\otimes p} + W,$$

where $W = (W_{i_1 \dots i_p})_{1 \leq i_1, \dots, i_p \leq N}$ is now a tensor of order p made of independent standard Gaussian random variables, independent of the vector \bar{x} , and where for any $x \in \mathbb{R}^N$, we denote by $x^{\otimes p}$ the tensor of order p such that, for every $i_1, \dots, i_p \in \{1, \dots, N\}$,

$$(x^{\otimes p})_{i_1 \dots i_p} = x_{i_1} \cdots x_{i_p}.$$

For two tensors $V = (V_{i_1 \dots i_p})_{1 \leq i_1, \dots, i_p \leq N}$ and $V' = (V'_{i_1 \dots i_p})_{1 \leq i_1, \dots, i_p \leq N}$ of order p , we define

$$V : V' = \sum_{1 \leq i_1, \dots, i_p \leq N} V_{i_1 \dots i_p} V'_{i_1 \dots i_p}.$$

We redefine $H_N(t, h, x)$ to be

$$\begin{aligned} H_N(t, h, x) := & \frac{\sqrt{t}}{N^{\frac{p-1}{2}}} W : x^{\otimes p} + \frac{t}{N^{p-1}} (x \cdot \bar{x})^p - \frac{t}{2N^{p-1}} |x|^{2p} \\ & + \sqrt{h} z \cdot x + h x \cdot \bar{x} - \frac{h}{2} |x|^2, \end{aligned}$$

and set F_N and \bar{F}_N to be as in (2.3) and (2.4). The analogue of Theorem 2.1 in the context of tensors reads as follows.

Theorem 5.1 (Convergence to HJ). *For every $M \geq 1$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\sup_{[0, M]^2} (F_N - f)^2 \right] = 0,$$

where $f(t, h)$ is the viscosity solution of the Hamilton-Jacobi equation

$$(5.1) \quad \begin{cases} \partial_t f - 2^{p-1} (\partial_h f)^p = 0 & \text{in } (0, +\infty)^2, \\ -\partial_h f = 0 & \text{on } (0, +\infty) \times \{0\}, \end{cases}$$

with initial condition $f(0, h) = \psi(h)$.

The next proposition is our replacement for Proposition 2.2.

Proposition 5.2 (Approximate HJ in finite volume). *There exists $C < \infty$ such that for every $N \geq 1$ and uniformly over $[0, \infty)^2$,*

$$|\partial_t \bar{F}_N - 2^{p-1}(\partial_h \bar{F}_N)^p|^2 \leq \frac{C}{N} \partial_h^2 \bar{F}_N + C \mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2 \right] + \frac{C}{Nh},$$

and moreover,

$$(5.2) \quad \partial_h \bar{F}_N \geq 0.$$

Proof of Proposition 5.2. Observe that

$$(5.3) \quad \partial_t F_N(t, h) = \frac{1}{N} \left\langle \frac{1}{2N^{\frac{p-1}{2}} \sqrt{t}} W : x^{\otimes p} + \frac{1}{N^{p-1}} (x \cdot \bar{x})^p - \frac{1}{2N^{p-1}} |x|^{2p} \right\rangle.$$

By Gaussian integration by parts and the Nishimori identity, we deduce that

$$\partial_t \bar{F}_N(t, h) = \frac{1}{2N^p} \mathbb{E} \langle (x \cdot \bar{x})^p \rangle.$$

The expressions (3.7)-(3.8) are still valid (as well as (3.10)). We deduce that

$$(5.4) \quad 2\partial_t \bar{F}_N - 2^p (\partial_h \bar{F}_N)^p = \mathbb{E} \left\langle \left(\frac{x \cdot \bar{x}}{N} \right)^p \right\rangle - \left(\mathbb{E} \left\langle \frac{x \cdot \bar{x}}{N} \right\rangle \right)^p.$$

Using that $a^p - b^p = (a - b)(a^{p-1} + \dots + b^{p-1})$ and the fact that the support of the measure P is bounded, we get that

$$|\partial_t \bar{F}_N - 2^{p-1}(\partial_h \bar{F}_N)^p|^2 \leq \frac{C}{N^2} \mathbb{E} \langle (x \cdot \bar{x} - \mathbb{E} \langle x \cdot \bar{x} \rangle)^2 \rangle.$$

The arguments in the proof of Proposition 2.2 apply without any modification to show that

$$\begin{aligned} \mathbb{E} \langle (x \cdot \bar{x} - \mathbb{E} \langle x \cdot \bar{x} \rangle)^2 \rangle &\leq 4 \mathbb{E} \left\langle \left(H'_N(h, x) - \mathbb{E} \langle H'_N(h, x) \rangle \right)^2 \right\rangle \\ &\leq 4N \partial_h^2 \bar{F}_N(t, h) + 4N^2 \mathbb{E} \left[(\partial_h F_N(t, h) - \partial_h \bar{F}_N(t, h))^2 \right] \\ &\quad + CNh^{-1}. \end{aligned}$$

Combining the two previous displays yields Proposition 5.2. \square

Proof of Theorem 5.1. As in the proof of Theorem 2.1, it suffices to show that if f is such that \bar{F}_N converges locally uniformly to f along a subsequence, then f is a viscosity solution of (5.1). Abusing notation, we do not write explicitly the subsequence along which the convergence holds.

Step 1. We show that f is a viscosity subsolution of (5.1). The proof follows Steps 2-6 of the proof of Theorem 2.1 very closely. The first difference is that we use Proposition 5.2 to replace (4.12) by

$$(5.5) \quad \begin{aligned} &(\partial_t \bar{F}_N - 2^{p-1}(\partial_h \bar{F}_N)^p)(t'_N, h') \\ &\leq \left(\frac{C}{N} \partial_h^2 \bar{F}_N(t'_N, h') + C \mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2(t'_N, h') \right] + \frac{C}{Nh'} \right)^{\frac{1}{2}}. \end{aligned}$$

(Implicit in this expression is the fact that the quantity under the square root on the right side is nonnegative.) The estimates (4.13) and (4.14) still hold and

the proofs given there apply without any modification. We deduce as in (4.15) that

$$\left| (\partial_h G_N)^p(t'_N, h'_N) - \delta_N^{-1} \int_{h'_N}^{h'_N + \delta_N} (\partial_h \bar{F}_N)^p(t'_N, h') dh' \right| \leq C\sqrt{\delta_N}.$$

We average the inequality (5.5) over $h' \in [h'_N, h'_N + \delta_N]$, use Jensen's inequality, the estimate above and (4.14) to obtain that

$$(\partial_t G_N - 2^{p-1}(\partial_h G_N)^p)(t'_N, h'_N) \leq \left(\frac{C}{N} \partial_h^2 G_N(t'_N, h'_N) + C\delta_N \right)^{\frac{1}{2}} + C\sqrt{\delta_N},$$

and then conclude as before that (4.16) holds.

Step 2. We now show that f is a viscosity supersolution of (5.1). Let $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ be such that $f - \phi$ has a strict local minimum at the point (t, h) . We keep the definition of δ_N as in (4.4) for consistency of notation (although here a simpler choice not depending on the rate of convergence of \bar{F}_N to f would also do), and redefine G_N to be

$$G_N(t', h') = \delta_N^{-1} \int_{h'+\delta_N}^{h'+2\delta_N} \bar{F}_N(t', h'') dh''.$$

In this new definition of G_N , we have shifted the interval over which the integral is taken by δ_N to the right in order to avoid the singularity of the error term in Proposition 5.2 near $h = 0$. Since G_N converges to f locally uniformly, there exists a sequence $(t_N, h_N) \in (0, \infty) \times [0, \infty)$ converging to (t, h) as N tends to infinity such that $G_N - \phi$ has a local minimum at (t_N, h_N) . By Proposition 5.2, for every $h' > 0$,

$$(5.6) \quad (\partial_t \bar{F}_N - 2^{p-1}(\partial_h \bar{F}_N)^p)(t_N, h') \geq - \left(\frac{C}{N} \partial_h^2 \bar{F}_N(t_N, h') + C\mathbb{E} \left[(\partial_h F_N - \partial_h \bar{F}_N)^2(t_N, h') \right] + \frac{C}{Nh'} \right)^{\frac{1}{2}}.$$

We also observe that the estimate (4.14) still holds in the present context. Averaging the inequality (5.6) over $h' \in [h_N + \delta_N, h_N + 2\delta_N]$, using Jensen's inequality, and (4.14), we get that

$$(\partial_t G_N - 2^{p-1}(\partial_h G_N)^p)(t_N, h_N) \geq - \left(\frac{C}{N} \partial_h^2 G_N(t_N, h_N) + C\delta_N + \frac{C}{N\delta_N} \right)^{\frac{1}{2}}.$$

Using Jensen's inequality for the left side of (5.6) is justified since $\partial_h \bar{F}_N \geq 0$. Since $\partial_h \bar{F}_N$ is bounded, we have that $|\partial_h^2 G_N(t_N, h_N)| \leq \delta_N^{-1}$. We thus obtain that

$$(5.7) \quad \liminf_{N \rightarrow \infty} (\partial_t G_N - 2^{p-1}(\partial_h G_N)^p)(t_N, h_N) \geq 0.$$

If $h_N > 0$ for infinitely many values of N , then the first derivatives of G_N and ϕ coincide at (t_N, h_N) for these values of N , and we thus deduce from (5.7) that

$$(\partial_t \phi - 2^{p-1}(\partial_h \phi)^p)(t, h) = \lim_{N \rightarrow \infty} (\partial_t \phi - 2^{p-1}(\partial_h \phi)^p)(t_N, h_N) \geq 0.$$

On the other hand, if $h_N = 0$ for infinitely many values of N , then we can reproduce the argument of Step 1 of the proof of Theorem 5.1 to conclude.

Indeed, in this case, we must have $h = 0$,

$$(5.8) \quad \partial_h(G_N - \phi)(t_N, h_N) \geq 0, \quad \text{and} \quad \partial_t(G_N - \phi)(t_N, h_N) = 0.$$

If $-\partial_h\phi(t, h) \geq 0$, then there is nothing to show. Else, using the first statement in (5.8), we have that

$$(\partial_h\phi)^p(t, h) = \lim_{N \rightarrow \infty} (\partial_h\phi)^p(t_N, h_N) \leq \liminf_{N \rightarrow \infty} (\partial_h G_N)^p(t_N, h_N),$$

and thus, using the second statement in (5.8) and then (5.7), we deduce that

$$(\partial_t\phi - 2^{p-1}(\partial_h\phi)^p)(t, h) \geq \limsup_{N \rightarrow \infty} (\partial_t G_N - 2^{p-1}(\partial_h G_N)^p)(t_N, h_N) \geq 0,$$

thereby completing the proof. \square

APPENDIX A. NISHIMORI IDENTITY

We show that under the measure defined in (3.1), the law of the random variable x is the law of \bar{x} conditionally on the relevant observations, and then verify the Nishimori identity stated in (3.2) and (3.3). To encode the observations, we redefine the variable Y to be

$$Y = (Y^{(1)}, Y^{(2)}) := \left(\sqrt{\frac{t}{N}} \bar{x} \bar{x}^t + W, \sqrt{h} \bar{x} + z \right).$$

For every bounded measurable functions f and g , we can write $\mathbb{E}[f(\bar{x})g(Y)]$, up to a normalization constant that depends neither on f nor on g , as

$$\int f(x)g\left(\sqrt{\frac{t}{N}} x x^t + W, \sqrt{h} x + z\right) \exp\left(-\sum_{i,j=1}^N \frac{W_{ij}^2}{2} - \sum_{i=1}^N \frac{z_i^2}{2}\right) dW dz dP_N(x),$$

with the shorthand notation $dW := \prod_{i,j} dW_{ij}$ and $dz := \prod_i dz_i$. A change of variables leads to

$$\int f(x)g(Y^{(1)}, Y^{(2)}) \exp\left(-\frac{1}{2} \sum_{i,j=1}^N \left(Y_{ij}^{(1)} - \sqrt{\frac{t}{N}} x_i x_j\right)^2 - \sum_{i=1}^N \frac{(Y_i^{(2)} - \sqrt{h} x)^2}{2}\right) dY dP_N(x).$$

Denoting the exponential factor above by $\mathcal{E}(x, Y)$, we thus obtain that the law of Y is the law with density given, up to a normalization constant, by

$$\bar{\mathcal{E}}(Y) := \int \mathcal{E}(x, Y) dP_N(x),$$

and that

$$(A.1) \quad \mathbb{E}[f(\bar{x})g(Y)] = \int f(x) \frac{\mathcal{E}(x, Y)}{\bar{\mathcal{E}}(Y)} dP_N(x) g(Y) \bar{\mathcal{E}}(Y) dY.$$

The conditional law of \bar{x} given Y is thus the probability measure given by

$$\frac{\mathcal{E}(x, Y)}{\bar{\mathcal{E}}(Y)} dP_N(x).$$

Moreover, a calculation similar to that in (2.2) yields that this quantity can be rewritten as

$$\frac{e^{H_N(t,h,x)} dP_N(x)}{\int_{\mathbb{R}^N} e^{H_N(t,h,x')} dP_N(x')},$$

which is the Gibbs measure defined in (3.1). We denote by $x^{(1)}, \dots, x^{(k)}$ a sequence of k random variables which, conditionally on \bar{x} , W and z , are independent and distributed according to this measure; we denote their joint (conditional) law by $\langle \cdot \rangle$. We thus have that, for every $k \in \mathbb{N}$ and bounded measurable function $f(x^{(1)}, \dots, x^{(k)}, Y)$,

$$\begin{aligned} \mathbb{E} \langle f(x^{(1)}, \dots, x^{(k)}, Y) \rangle &= \int f(x^{(1)}, \dots, x^{(k)}, Y) \\ &\times \frac{\mathcal{E}(x^{(1)}, Y)}{\bar{\mathcal{E}}(Y)} dP_N(x^{(1)}) \dots \frac{\mathcal{E}(x^{(k)}, Y)}{\bar{\mathcal{E}}(Y)} dP_N(x^{(k)}) \bar{\mathcal{E}}(Y) dY. \end{aligned}$$

In view of the expression for the joint law of \bar{x} and Y obtained in (A.1), we deduce that

$$(A.2) \quad \mathbb{E} \langle f(x^{(1)}, \dots, x^{(k)}, Y) \rangle = \mathbb{E} \langle f(x^{(1)}, \dots, x^{(k-1)}, \bar{x}, Y) \rangle.$$

This implies in particular that (3.2) and (3.3) hold.

APPENDIX B. CLASSICAL RESULTS ON VISCOSITY SOLUTIONS

In this appendix, for the reader's convenience, we prove the comparison principle (Proposition 4.2) and the Hopf-Lax formula (Proposition 4.3) for solutions of the Hamilton-Jacobi equation (2.6). Classical references for such results include [13, 10].

Proof of Proposition 4.2. We argue by contradiction, assuming instead that

$$(B.1) \quad \sup_{[0, \infty)^2} (u - v) > \sup_{\{0\} \times [0, \infty)} (u - v).$$

The argument rests on the idea of “doubling the variables” and considering the maximization of functions of the form

$$(B.2) \quad ((t, h), (t', h')) \mapsto u(t, h) - v(t', h') - \frac{1}{2\alpha} (|t - t'|^2 + |h - h'|^2),$$

where $\alpha > 0$ is a parameter that is ultimately sent to 0. We will decompose this argument into four steps. In Step 1, we modify the functions u and v slightly so that they become strict sub- and supersolutions respectively. In Step 2, we modify the function u further to ensure that the maximum of the function in (B.2) is achieved at a point that remains in a bounded set as $\alpha \rightarrow 0$. In a preliminary Step 0, we build a convenient special function for this purpose. The conclusion is then derived in Step 3.

Step 0. We build a family of functions $(\Phi_\delta)_{\delta \in (0, 1]}$ in $C^\infty([0, \infty)^2)$ such that for every $T, H > 0$ and every $\delta > 0$ sufficiently small (in terms of T and H), the following properties hold.

$$(B.3) \quad \|\Phi_\delta\|_{L^\infty([0, T] \times [0, H])} + \|\nabla \Phi_\delta\|_{L^\infty([0, T] \times [0, H])} \leq \delta,$$

$$(B.4) \quad \forall t \geq 0, \quad \forall h \geq 2\delta^{-2}, \quad \Phi_\delta(t, h) \geq \delta^{-1}h,$$

and

$$(B.5) \quad \partial_t \Phi_\delta \geq \delta^{-\frac{1}{2}} |\partial_h \Phi_\delta| \quad \text{on } [0, T] \times [0, \infty).$$

Let $\chi \in C^\infty(\mathbb{R})$ be a smooth function satisfying $0 \leq \chi \leq 1$ and such that $\chi \equiv 0$ on $(-\infty, 0]$ and $\chi \equiv 1$ on $[1, \infty)$. For every $z \in \mathbb{R}$, we set

$$\sigma(z) := \int_{-\infty}^z e^{z'} (1 - \chi(z')) dz'.$$

Note that for every $z \in (-\infty, 0]$, we have that $\sigma(z) = e^z$ and that σ is constant (and $\sigma \geq 1$) on $[1, \infty)$. Moreover, there exists a constant $C < \infty$ such that

$$(B.6) \quad 0 \leq \sigma' \leq C\sigma.$$

For every $\delta \in (0, 1]$, we consider the function

$$\Phi_\delta(t, h) := \delta^{-1} h \sigma(\delta h - \delta^{-1}) + \delta^{-3} t \sigma(\delta h - \delta^{-1}).$$

Roughly speaking, the function $h \mapsto \sigma(\delta h - \delta^{-1})$ serves as a smoothed indicator function for the set $\{h \geq \delta^{-2}\}$. The properties (B.3) and (B.4) are immediate. For the last property, we observe that

$$\partial_t \Phi_\delta(t, h) = \delta^{-3} \sigma(\delta h - \delta^{-1}),$$

$$\partial_h \Phi_\delta(t, h) = \delta^{-1} \sigma(\delta h - \delta^{-1}) + h \sigma'(\delta h - \delta^{-1}) + \delta^{-2} t \sigma'(\delta h - \delta^{-1}).$$

Since σ' is supported in $(-\infty, 1]$, the second term on the right side vanishes whenever $h \geq \delta^{-2} + \delta^{-1}$, and thus in particular whenever $h \geq 2\delta^{-2}$. We deduce that

$$0 \leq \partial_h \Phi_\delta(t, h) \leq \delta^{-1} \sigma(\delta h - \delta^{-1}) + 2\delta^{-2} \sigma'(\delta h - \delta^{-1}) + \delta^{-2} t \sigma'(\delta h - \delta^{-1}).$$

The inequality (B.5) thus follows using (B.6).

Step 1. We show that without loss of generality, we can assume that there exists $\varepsilon > 0$ such that u is a viscosity solution of

$$(B.7) \quad \begin{cases} \partial_t u - 2(\partial_h u)^2 \leq -\varepsilon & \text{in } (0, +\infty)^2, \\ -\partial_h u \leq -\varepsilon & \text{on } (0, +\infty) \times \{0\}. \end{cases}$$

Indeed, since we assume that $\partial_h u \in L^\infty([0, \infty)^2)$, it suffices to replace u by $u_\varepsilon := u + \varepsilon h - C\varepsilon t$ for a sufficiently large constant C depending on $\|\partial_h u\|_{L^\infty([0, \infty)^2)}$, and then select $\varepsilon > 0$ sufficiently small that the property (B.1) still holds. Similarly, we can assume that the function v is a viscosity solution of

$$(B.8) \quad \begin{cases} \partial_t v - 2(\partial_h v)^2 \geq \varepsilon & \text{in } (0, +\infty)^2, \\ -\partial_h v \geq \varepsilon & \text{on } (0, +\infty) \times \{0\}. \end{cases}$$

Note that these modifications preserve the fact that u and v are uniformly Lipschitz continuous in h .

Step 2. We now “localize” the function u , in the sense that we make sure that the function becomes very negative as we move away from a bounded set. To start with, we can replace u by $u - \frac{\varepsilon}{T-t}$ for some $\varepsilon > 0$ and $T \in (0, \infty)$. This preserves the fact that u solves (B.7) on $(0, T) \times [0, \infty)$, and for $\varepsilon > 0$ sufficiently small and T sufficiently large, it also preserves the property (B.1) in the sense that

$$(B.9) \quad \sup_{[0, T) \times [0, \infty)^2} (u - v) > \sup_{\{0\} \times [0, \infty)} (u - v).$$

This modification of u also ensures that for every $H > 0$,

$$(B.10) \quad \limsup_{\eta \rightarrow 0} \{u(t, h) : (t, h) \in [T - \eta, T) \times [0, H]\} = -\infty.$$

Note that we still have that

$$\sup_{t \in [0, T)} u(t, 0) < +\infty,$$

and that u is uniformly Lipschitz continuous in h , and thus there exists a constant $C < \infty$ such that

$$(B.11) \quad \forall (t, h) \in [0, T) \times [0, \infty), \quad u(t, h) \leq C(1 + h).$$

Our final modification is to replace u by $u_\delta := u - \Phi_\delta$ for the function Φ_δ defined in Step 0. It is clear from (B.3) that for $\delta > 0$ sufficiently small, the properties (B.9) and (B.10) still hold with u replaced by u_δ . It is also clear from (B.4) and (B.11) that for every $\delta > 0$ sufficiently small, we have

$$(B.12) \quad \forall (t, h) \in [0, T) \times [0, +\infty), \quad u_\delta(t, h) \leq C(1 + h) - \delta^{-1}h\mathbb{1}_{\{h \geq 2\delta^{-2}\}}.$$

There remains to verify that u_δ is still a solution of (B.7), possibly after replacing ε by $\varepsilon/2$. Formally, the verification of the boundary condition on $[0, T) \times \{0\}$ is immediate from (B.3), while we have

$$\partial_t u_\delta - 2(\partial_h u_\delta)^2 = \partial_t u - 2(\partial_h u)^2 - \partial_t \Phi_\delta - 4\partial_h u \partial_h \Phi_\delta - 2(\partial_h \Phi_\delta)^2,$$

and since $\partial_h u$ is bounded, it follows from (B.5) that for every $\delta \geq 0$ sufficiently small,

$$\partial_t \Phi_\delta + 4\partial_h u \partial_h \Phi_\delta + 2(\partial_h \Phi_\delta)^2 \geq 0.$$

This formal calculation is easily made rigorous using test functions.

Finally, we observe that as a consequence of (B.12) and the Lipschitz continuity of v in the h variable, there exists a constant $C < \infty$ such that for every $t \in [0, T)$, $h, t', h' \geq 0$, we have

$$u_\delta(t, h) - v(t', h') \leq C(1 + h + h') - \delta^{-1}h\mathbb{1}_{\{h \geq 2\delta^{-2}\}}.$$

We select $\delta > 0$ sufficiently small that for t, h, t', h' as above,

$$(B.13) \quad u_\delta(t, h) - v(t', h') \leq C(1 + h + h' - 2h\mathbb{1}_{\{h \geq 2\delta^{-2}\}}).$$

Step 3. Summarizing the result of the previous steps, we have shown that without loss of generality, we can assume that v solves (B.7) on $(0, \infty) \times [0, \infty)$, that u solves (B.8) on $(0, T) \times [0, \infty)$ and satisfies (B.10) and (B.13) for every $t \in [0, T)$ and $h, t', h' \geq 0$, and that (B.9) holds.

We define, for every $\alpha \in (0, 1]$ and $t \in [0, T)$, $h, t', h' \geq 0$,

$$\Psi_\alpha(t, h, t', h') := u(t, h) - v(t', h') - \frac{1}{2\alpha} (|t - t'|^2 + |h - h'|^2).$$

We claim that the maximum of Ψ_α is achieved at some point $(t_\alpha, h_\alpha, t'_\alpha, h'_\alpha)$, and that this point remains in a bounded set as $\alpha > 0$ is sent to 0. For fixed $\alpha > 0$, consider a sequence of approximate maximizers for Ψ_α denoted by $(t_{n,\alpha}, h_{n,\alpha}, t'_{n,\alpha}, h'_{n,\alpha})$. We deduce from (B.13) that for some $C < \infty$,

$$C(1 + h_{n,\alpha} + h'_{n,\alpha} - 2h_{n,\alpha}\mathbb{1}_{\{h_{n,\alpha} \geq 2\delta^{-2}\}}) - \frac{1}{2\alpha} (|t_{n,\alpha} - t'_{n,\alpha}|^2 + |h_{n,\alpha} - h'_{n,\alpha}|^2) \geq -C,$$

and in particular,

$$\frac{1}{2} (|t_{n,\alpha} - t'_{n,\alpha}|^2 + |h_{n,\alpha} - h'_{n,\alpha}|^2) - C (1 + h_{n,\alpha} + h'_{n,\alpha} - 2h_{n,\alpha} \mathbb{1}_{\{h_{n,\alpha} \geq 2\delta^{-2}\}}) \leq C.$$

If $h_{n,\alpha} < 2\delta^{-2}$, then we can obtain a uniform upper bound on $h'_{n,\alpha}$, and thus also on $|t_{n,\alpha} - t'_{n,\alpha}|$. Otherwise, we can first obtain a uniform upper bound on $|h_{n,\alpha} - h'_{n,\alpha}|$, and then deduce an upper bound on $h_{n,\alpha}$, $h'_{n,\alpha}$ and $|t_{n,\alpha} - t'_{n,\alpha}|$. Finally, we can use (B.10) to conclude that the maximizer of Ψ_α exists and remains in a bounded set as α tends to 0.

Since $u - v$ remains bounded from above over this bounded set, see (B.13), there exists a constant $C < \infty$ such that for every $\alpha > 0$ sufficiently small, we have

$$|t_\alpha - t'_\alpha|^2 + |h_\alpha - h'_\alpha|^2 \leq C\alpha.$$

After extracting a subsequence if necessary, we may assume that $t_\alpha, t'_\alpha \rightarrow t_0$ and $h_\alpha, h'_\alpha \rightarrow h_0$ as $\alpha \rightarrow 0$. Using again (B.10), it is clear that $t_0 < T$. Since $(t_\alpha, h_\alpha, t'_\alpha, h'_\alpha)$ is a maximizer of Ψ_α , we have

$$(B.14) \quad \Psi_\alpha(t_\alpha, h_\alpha, t'_\alpha, h'_\alpha) \geq \sup_{[0, \infty)^2} (u - v) \geq u(t_0, h_0) - v(t_0, h_0).$$

Since we also have

$$\Psi_\alpha(t_\alpha, h_\alpha, t'_\alpha, h'_\alpha) \leq u(t_\alpha, h_\alpha) - v(t'_\alpha, h'_\alpha),$$

and the functions u and v are continuous at (t_0, h_0) , we deduce, using (B.14) twice, that

$$(B.15) \quad \lim_{\alpha \rightarrow 0} \Psi_\alpha(t_\alpha, h_\alpha, t'_\alpha, h'_\alpha) = u(t_0, h_0) - v(t_0, h_0) = \sup_{[0, \infty)^2} (u - v).$$

In view of (B.1) and the second equality in (B.15), we have that $t_0 > 0$, and thus, for every $\alpha > 0$ sufficiently small, we have that $t_\alpha > 0$ and $t'_\alpha > 0$. Note that by definition, the function

$$(t, h) \mapsto u(t, h) - v(t'_\alpha, h'_\alpha) - \frac{1}{2\alpha} (|t - t'_\alpha|^2 + |h - h'_\alpha|^2)$$

has a local maximum at $(t, h) = (t_\alpha, h_\alpha)$. If $h_\alpha = 0$, then the definition of viscosity solutions implies that, for every $\alpha > 0$ sufficiently small,

$$\min \left(-\frac{1}{\alpha} (h_\alpha - h'_\alpha), \frac{1}{\alpha} (t_\alpha - t'_\alpha) - \frac{2}{\alpha^2} (h_\alpha - h'_\alpha)^2 \right) \leq -\varepsilon.$$

Since in this case ($h_\alpha = 0$) the first term in the minimum is nonnegative, we deduce that

$$(B.16) \quad \frac{1}{\alpha} (t_\alpha - t'_\alpha) - \frac{2}{\alpha^2} (h_\alpha - h'_\alpha)^2 \leq -\varepsilon.$$

As can be seen directly, this conclusion also holds when $h_\alpha > 0$. Similarly, we infer from the fact that the function

$$(t', h') \mapsto v(t', h') - u(t_\alpha, h_\alpha) + \frac{1}{2\alpha} (|t_\alpha - t'|^2 + |h_\alpha - h'|^2)$$

has a local minimum at $(t', h') = (t'_\alpha, h'_\alpha)$ that

$$\frac{1}{\alpha} (t_\alpha - t'_\alpha) - \frac{2}{\alpha^2} (h_\alpha - h'_\alpha)^2 \geq \varepsilon.$$

This is in contradiction with (B.16), and thus the proof is complete. \square

We now turn to the proof of Proposition 4.3.

Proof of Proposition 4.3. We decompose the proof into five steps.

Step 1. We show that the function ψ is uniformly Lipschitz. This function is clearly differentiable at every $h > 0$ and

$$\psi'(h) = \mathbb{E} \left\langle \frac{zx}{2\sqrt{h}} + x\bar{x} - \frac{x^2}{2} \right\rangle,$$

where here the notation $\langle \cdot \rangle$ simplifies into

$$\langle f(x) \rangle := \frac{\int_{\mathbb{R}} f(x) \exp(\sqrt{h}zx + hx\bar{x} - \frac{h}{2}x^2) dP(x)}{\int_{\mathbb{R}} \exp(\sqrt{h}zx + hx\bar{x} - \frac{h}{2}x^2) dP(x)}.$$

As in (3.8), we obtain by Gaussian integration by parts that

$$(B.17) \quad \psi'(h) = \frac{1}{2} \mathbb{E} \langle x\bar{x} \rangle.$$

Since we assume that the support of the measure P is bounded, this completes the proof that ψ' is uniformly bounded.

Step 2. For convenience, we extend ψ to be constant equal to $\psi(0)$ on $(-\infty, 0]$, so that for every $t, h \geq 0$,

$$\begin{aligned} f(t, h) &= \sup_{h' \in \mathbb{R}} \left(\psi(h') - \frac{(h - h')^2}{8t} \right) \\ &= \sup_{h' \in \mathbb{R}} \left(\psi(h - h') - \frac{(h')^2}{8t} \right). \end{aligned}$$

For every $h, h_1 \geq 0$, we have

$$\psi(h - h') - \frac{(h')^2}{8t} \leq \psi(h_1 - h') - \frac{(h')^2}{8t} + \|\psi'\|_{L^\infty(\mathbb{R})} |h_1 - h|,$$

and thus, taking the supremum over h' on both sides,

$$f(t, h) \leq f(t, h_1) + \|\psi'\|_{L^\infty(\mathbb{R})} |h_1 - h|.$$

Since the roles of h and h_1 are symmetric, this shows that f is uniformly Lipschitz in the h variable.

Step 3. As preparation for the proof that f is a viscosity solution of (2.6), we prove the dynamic programming principle, namely, that for every $t, s, h \geq 0$,

$$(B.18) \quad f(t + s, h) = \sup_{h' \geq 0} \left(f(t, h') - \frac{(h - h')^2}{8s} \right).$$

By convexity of the square function, we have, for every $t, s, h, h', h'' \geq 0$,

$$(B.19) \quad \left(\frac{h - h'}{t + s} \right)^2 \leq \frac{t}{t + s} \left(\frac{h'' - h'}{t} \right)^2 + \frac{s}{t + s} \left(\frac{h - h''}{s} \right)^2,$$

and thus

$$\begin{aligned} f(t+s, h) &\geq \sup_{h', h'' \geq 0} \left(\psi(h') - \frac{(h'' - h')^2}{8t} - \frac{(h - h'')^2}{8s} \right) \\ &\geq \sup_{h'' \geq 0} \left(f(t, h'') - \frac{(h - h'')^2}{8s} \right). \end{aligned}$$

This proves one inequality in (B.18). The converse inequality is immediate, since the right side of (B.18) is

$$\sup_{h', h'' \geq 0} \left(\psi(h') - \frac{(h'' - h')^2}{8t} - \frac{(h - h'')^2}{8s} \right),$$

and, for each fixed $h, h' \geq 0$, we can achieve the case of equality in (B.19) for some $h'' \geq 0$, so that

$$\inf_{h'' \geq 0} \left(\frac{(h'' - h')^2}{8t} + \frac{(h - h'')^2}{8s} \right) = \frac{(h - h')^2}{8(t+s)}.$$

Step 4. We show that f is a viscosity supersolution of (2.6). Let $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ be such that (t, h) is a local minimum of $f - \phi$. We start by assuming that $h > 0$. By (B.18), we have that for every $p \in \mathbb{R}$ and $s > 0$ sufficiently small,

$$f(t, h) \geq f(t-s, h-sp) - \frac{sp^2}{8}.$$

Since $f - \phi$ has a local minimum at (t, h) , we have that for every $p \in \mathbb{R}$ and $s > 0$ sufficiently small,

$$f(t-s, h-sp) - \phi(t-s, h-sp) \geq f(t, h) - \phi(t, h).$$

Combining these two inequalities and passing to the limit $s \rightarrow 0$, we deduce that

$$\partial_t \phi(t, h) + p \partial_h \phi(t, h) + \frac{p^2}{8} \geq 0,$$

and taking the infimum over $p \in \mathbb{R}$ yields

$$(B.20) \quad (\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \geq 0,$$

as desired. In the case when $h = 0$, the same reasoning applies, except that we need to restrict to values of p such that $p \leq 0$. That is, for every $p \leq 0$,

$$\partial_t \phi(t, h) + p \partial_h \phi(t, h) + \frac{p^2}{8} \geq 0.$$

If $-\partial_h \phi(t, h) \geq 0$, then there is nothing to show. Otherwise, we can choose $p = -4\partial_h \phi(t, h)$ and conclude for the validity of (B.20).

Step 5. We show that f is a viscosity subsolution of (2.6). Let $(t, h) \in (0, \infty) \times [0, \infty)$ and $\phi \in C^\infty((0, \infty) \times [0, \infty))$ be such that (t, h) is a local maximum of $f - \phi$. In view of (B.18) and of the fact that f is uniformly Lipschitz in the h variable, it is clear that for each $s > 0$, there exists $h'_s \in \mathbb{R}$ such that

$$f(t, h) = f(t-s, h-h'_s) - \frac{(h'_s)^2}{8s},$$

and moreover, we have that $h'_s \rightarrow 0$ as $s \rightarrow 0$. Since $f - \phi$ has a local maximum at (t, h) , we have that, for every $s > 0$ sufficiently small,

$$(f - \phi)(t - s, h - h'_s) \leq (f - \phi)(t, h),$$

and thus

$$(B.21) \quad \phi(t, h) - \phi(t - s, h - h'_s) \leq -\frac{(h'_s)^2}{8s}.$$

We start by assuming that $h > 0$, in which case we aim to show that

$$(B.22) \quad (\partial_t \phi - 2(\partial_h \phi)^2)(t, h) \leq 0.$$

We argue by contradiction, assuming the negation of (B.22). By continuity, there exist $\varepsilon, \delta > 0$ such that for every $s, a \in [-\delta, \delta]$, we have

$$(B.23) \quad (\partial_t \phi - 2(\partial_h \phi)^2)(t - s, h - a) \geq \varepsilon,$$

and thus, for every such s and a and every $p \in \mathbb{R}^d$,

$$(\partial_t \phi + p \partial_h \phi)(t - s, h - a) \geq \varepsilon - \frac{p^2}{8}.$$

By integration, we deduce that, for every $s \in (0, \delta]$,

$$\begin{aligned} \phi(t, h) - \phi(t - s, h - a) &= \int_0^1 (s \partial_t \phi + a \partial_h \phi)(t - \alpha s, h - \alpha a) \, d\alpha \\ &= s \int_0^1 \left(\partial_t \phi + \frac{a}{s} \partial_h \phi \right)(t - \alpha s, h - \alpha a) \, d\alpha \\ &\geq s\varepsilon - \frac{a^2}{8s}. \end{aligned}$$

For $s > 0$ sufficiently small, we can choose $a = h'_s$ in the inequality above, and reach a contradiction with (B.21). This shows (B.22) in the case $h > 0$.

When $h = 0$, our starting point has to be modified from (B.23) to the statement that for every $s \in [-\delta, \delta]$ and $a \in [-\delta, 0]$,

$$\min(-\partial_h \phi, \partial_t \phi - 2(\partial_h \phi)^2)(t - s, h - a) \geq \varepsilon.$$

We can then reproduce the argument above and arrive at a contradiction. \square

Acknowledgments. I would like to warmly thank Louigi Addario-Berry and Pascal Maillard for stimulating discussions, the CRM and Jessica Lin for hospitality, and Florent Krzakala and Lenka Zdeborová for organizing a conference in Cargèse which inspired the present paper. I was partially supported by the ANR grants LSD (ANR-15-CE40-0020-03) and Malin (ANR-16-CE93-0003) and by a grant from the NYU–PSL Global Alliance.

REFERENCES

- [1] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008.
- [2] J. Barbier, M. Dia, N. Macris, F. Krzakala, T. Lesieur, and L. Zdeborová. Mutual information for symmetric rank-one matrix estimation: a proof of the replica formula. In *Advances in Neural Information Processing Systems 29*, pages 424–432, 2016.
- [3] J. Barbier and N. Macris. The adaptive interpolation method: a simple scheme to prove replica formulas in Bayesian inference. *Probab. Theory Related Fields*, in press.

- [4] J. Barbier, N. Macris, and L. Miolane. The layered structure of tensor estimation and its mutual information. In *55th Annual Allerton Conference on Communication, Control, and Computing*, pages 1056–1063. IEEE, 2017.
- [5] A. Barra, G. Del Ferraro, and D. Tantari. Mean field spin glasses treated with PDE techniques. *Eur. Phys. J. B*, 86(7):Art. 332, 10, 2013.
- [6] A. Barra, A. Di Biasio, and F. Guerra. Replica symmetry breaking in mean-field spin glasses through the Hamilton-Jacobi technique. *J. Stat. Mech. Theory Exp.*, (9):P09006, 22, 2010.
- [7] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013.
- [8] J. G. Brankov and V. A. Zagrebnov. On the description of the phase transition in the Husimi-Temperley model. *J. Phys. A*, 16(10):2217–2224, 1983.
- [9] S. Chatterjee. *Superconcentration and related topics*. Springer Monographs in Mathematics. Springer, Cham, 2014.
- [10] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [11] Y. Deshpande and A. Montanari. Information-theoretically optimal sparse PCA. In *IEEE International Symposium on Information Theory*, pages 2197–2201, 2014.
- [12] A. El Alaoui and F. Krzakala. Estimation in the spiked Wigner model: a short proof of the replica formula, preprint, arXiv:1801.01593.
- [13] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [14] F. Guerra. Sum rules for the free energy in the mean field spin glass model. *Fields Institute Communications*, 30:161, 2001.
- [15] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.
- [16] M. Lelarge and L. Miolane. Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Related Fields*, in press.
- [17] T. Lesieur, F. Krzakala, and L. Zdeborová. Phase transitions in sparse PCA. In *IEEE International Symposium on Information Theory*, pages 1635–1639, 2015.
- [18] T. Lesieur, L. Miolane, M. Lelarge, F. Krzakala, and L. Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. In *IEEE International Symposium on Information Theory*, pages 511–515, 2017.
- [19] M. Mézard, G. Parisi, and M. Virasoro. *Spin glass theory and beyond: an introduction to the replica method and its applications*, volume 9. World Scientific Publishing Company, 1987.
- [20] C. Newman. Percolation theory: A selective survey of rigorous results. In *Advances in multiphase flow and related problems*. SIAM, 1986.
- [21] D. Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.
- [22] M. Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.
- [23] M. Talagrand. Mean field models for spin glasses: some obnoxious problems. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 63–80. Springer, Berlin, 2007.
- [24] M. Talagrand. *Mean field models for spin glasses. Volume II*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer, Heidelberg, 2011.

(J.-C. Mourrat) DMA, ECOLE NORMALE SUPÉRIEURE, CNRS, PSL UNIVERSITY, PARIS, FRANCE

E-mail address: mourrat@dma.ens.fr