

# UN-INVERTING THE PARISI FORMULA

JEAN-CHRISTOPHE MOURRAT

ABSTRACT. The free energy of any system can be written as the supremum of a functional involving an energy term and an entropy term. Surprisingly, the limit free energy of mean-field spin glasses is expressed as an infimum instead, a phenomenon sometimes called an inverted variational principle. Using a stochastic-control representation of the Parisi functional and convex duality arguments, we rewrite this limit free energy as a supremum over martingales of a Wiener space.

## 1. INTRODUCTION AND MAIN RESULT

Let  $h \in \mathbb{R}$  and  $(\beta_p)_{p \geq 2}$  be a sequence of nonnegative real numbers such that the function  $\xi(r) := \sum_{p \geq 2} \beta_p^2 r^p$  is finite for every  $r \in \mathbb{R}$ . For every integer  $N \geq 1$ , let  $(H'_N(\sigma))_{\sigma \in \mathbb{R}^N}$  be the centered Gaussian field with covariance such that, for every  $\sigma, \tau \in \mathbb{R}^N$ ,

$$\mathbb{E}[H'_N(\sigma)H'_N(\tau)] = N\xi\left(\frac{\sigma \cdot \tau}{N}\right),$$

and for every  $\sigma \in \mathbb{R}^N$ , let

$$H_N(\sigma) := H'_N(\sigma) + h \sum_{i=1}^N \sigma_i.$$

The object of study of this paper is the limit free energy

$$(1.1) \quad f := \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \log \left( \frac{1}{2^N} \sum_{\sigma \in \{-1,1\}^N} \exp(H_N(\sigma)) \right).$$

An explicit representation for this limit was conjectured in [32, 33, 34, 35, 24] and then proved rigorously in [16, 39, 30, 31]. In order to state this result, we denote by  $\text{Pr}([0,1])$  the set of probability measures over  $[0,1]$ . For each  $\mu \in \text{Pr}([0,1])$ , we define  $\Phi_\mu = \Phi_\mu(t, x) : [0,1] \times \mathbb{R} \rightarrow \mathbb{R}$  to be the solution to the backwards parabolic equation

$$(1.2) \quad -\partial_t \Phi_\mu(t, x) = \frac{\xi''(t)}{2} \left( \partial_x^2 \Phi_\mu(t, x) + \mu[0,t] (\partial_x \Phi_\mu(t, x))^2 \right),$$

with terminal condition

$$(1.3) \quad \Phi_\mu(1, x) = \log \cosh x.$$

---

*Date:* February 15, 2024.

The Parisi formula states that the limit free energy  $f$  exists and is given by

$$(1.4) \quad f = \inf_{\mu \in \text{Pr}([0,1])} \left\{ \Phi_{\mu}(0, h) - \frac{1}{2} \int_0^1 t \xi''(t) \mu[0, t] dt \right\}.$$

It was shown in [3] that the variational problem in (1.4) admits a unique minimizer. For “most” choices of  $\xi$ , this minimizer encodes the asymptotic law of the overlap between two independent samples from the Gibbs measure associated with the energy function  $H_N$ , see [31, Corollaries 3.1 and 3.4].

Any expression of the form  $\log \mathbb{E}[\exp(f(X))]$  can be rewritten as a supremum over probability measures of an energy term and an entropy term, see for instance [5, Corollaries 4.14 and 4.15]. Such a formulation has a clear interpretation based on physical quantities or large-deviations considerations. It would seem reasonable that aspects of this structure be preserved in the limit of large system size, provided that we understand the asymptotics of the relevant energy and entropy terms. It thus initially came as a surprise that the expression in (1.4) takes the form of an infimum instead, and to this day I am not aware of a compelling physical interpretation of this variational problem. In fact, for more complex models with multiple types, such a variational representation of the limit free energy is no longer valid in general, see [26, Section 6].

The main goal of this paper is to give an alternative representation of the limit free energy  $f$  as a supremum instead. In order to state the result, we define, for every  $x \in \mathbb{R}$ ,

$$(1.5) \quad \phi(x) := \log \cosh(x + h),$$

and let  $\phi^*$  denote the convex dual of  $\phi$ , so that for every  $\lambda \in \mathbb{R}$ ,

$$\phi^*(\lambda) := \sup_{x \in \mathbb{R}} (\lambda x - \phi(x)) \in \mathbb{R} \cup \{+\infty\}.$$

A classical calculation yields that  $\phi^*$  is infinite outside of  $[-1, 1]$ , while for every  $\lambda \in [-1, 1]$ ,

$$\phi^*(\lambda) = \frac{1}{2} [(1 + \lambda) \log(1 + \lambda) + (1 - \lambda) \log(1 - \lambda)] - \lambda h.$$

Let  $\mathcal{P} = (\Omega, (\mathcal{F}_t)_{t \in [0,1]}, \mathbf{P})$  be a filtered probability space. The  $\sigma$ -algebras  $(\mathcal{F}_t)_{t \in [0,1]}$  are assumed to be complete, that is, they contain every subset of any null-measure set. We also assume that  $\mathcal{P}$  is sufficiently rich that one can define a Brownian motion  $(W_t)_{t \in [0,1]}$  over it (in particular, the process  $W$  is adapted and has independent increments with respect to the filtration  $(\mathcal{F}_t)_{t \in [0,1]}$ ). We denote by **Mart** the space of bounded martingales over  $\mathcal{P}$ .

**Theorem 1** (un-inverted Parisi formula). *The limit free energy  $f$  defined in (1.1) is given by*

$$(1.6) \quad f = \sup_{\alpha \in \mathbf{Mart}} \left\{ \mathbf{E} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) \right] - \frac{1}{2} \sup_{t \in [0,1]} \int_t^1 \xi''(s)(s - \mathbf{E}[\alpha_s^2]) ds \right\}.$$

Moreover, there exists a unique  $\alpha \in \mathbf{Mart}$  that realizes the supremum in (1.6).

Theorem 1 is a consequence of a representation of the optimizers of (1.4) and (1.6) as a dual pair in a saddle-point problem. For every  $\mu \in \Pr([0, 1])$  and  $\alpha \in \mathbf{Mart}$ , we define

$$(1.7) \quad \Gamma(\mu, \alpha) := \mathbf{E} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t](t - \alpha_t^2) dt \right].$$

**Theorem 2** (saddle-point problem for optimizers). *Let  $\bar{\mu} \in \Pr([0, 1])$  denote the unique minimizer of (1.4), and let  $\bar{\alpha} \in \mathbf{Mart}$  denote the unique maximizer of (1.6). The limit free energy (1.1) is given by*

$$(1.8) \quad f = \Gamma(\bar{\mu}, \bar{\alpha}) = \inf_{\mu \in \Pr([0,1])} \sup_{\alpha \in \mathbf{Mart}} \Gamma(\mu, \alpha) = \sup_{\alpha \in \mathbf{Mart}} \inf_{\mu \in \Pr([0,1])} \Gamma(\mu, \alpha).$$

Moreover, the following two properties of  $\bar{\mu}$  and  $\bar{\alpha}$  are valid and, taken together, characterize the pair  $(\bar{\mu}, \bar{\alpha})$  uniquely in  $\Pr([0, 1]) \times \mathbf{Mart}$ .

(1) The support of  $\bar{\mu}$  is a subset of the set of maximizers of the mapping

$$(1.9) \quad \begin{cases} [0, 1] & \rightarrow \mathbb{R} \\ t & \mapsto \int_t^1 \xi''(s)(s - \mathbf{E}[\bar{\alpha}_s^2]) ds. \end{cases}$$

(2) Letting  $(X_t)_{t \in [0,1]}$  be the strong solution to

$$(1.10) \quad \begin{cases} X_0 = h, \\ dX_t = \xi''(t) \bar{\mu}[0, t] \partial_x \Phi_{\bar{\mu}}(t, X_t) dt + \sqrt{\xi''(t)} dW_t, \end{cases}$$

we have, for every  $t \in [0, 1]$ ,

$$(1.11) \quad \bar{\alpha}_t = \partial_x \Phi_{\bar{\mu}}(t, X_t) = \partial_x \Phi_{\bar{\mu}}(0, h) + \int_0^t \sqrt{\xi''(s)} \partial_x^2 \Phi_{\bar{\mu}}(s, X_s) dW_s.$$

For each measure  $\mu \in \Pr([0, 1])$ , the Parisi formula (1.4) can be used to obtain an upper bound on the limit free energy  $f$ . In particular, we can write a first replica-symmetric upper bound obtained by imposing the support of  $\mu$  to be a singleton, and then progress along the hierarchy of replica symmetry breaking by allowing the support of  $\mu$  to contain two elements, then three elements, etc. Using the characterization above, we can associate to each  $\mu \in \Pr([0, 1])$  a corresponding lower bound on the free energy, as described

in the next theorem. This result could for instance facilitate the construction of certified numerical approximations of the limit free energy.

**Theorem 3** (RSB lower bound). *For every  $\mu \in \text{Pr}([0, 1])$ , we have*

$$(1.12) \quad 0 \leq \Phi_\mu(0, h) - \frac{1}{2} \int_0^1 t \xi''(t) \mu[0, t] dt - f \leq \frac{1}{2} \int \left( g_\mu(t) - \inf_{[0, 1]} g_\mu \right) d\mu(t),$$

where the function  $g_\mu$  is such that, letting  $(X_t)_{t \in [0, 1]}$  be the strong solution to

$$(1.13) \quad \begin{cases} X_0 = h, \\ dX_t = \xi''(t) \mu[0, t] \partial_x \Phi_\mu(t, X_t) dt + \sqrt{\xi''(t)} dW_t, \end{cases}$$

we have for every  $t \in [0, 1]$  that

$$(1.14) \quad g_\mu(t) := \int_t^1 \xi''(s) \left( \mathbf{E}[(\partial_x \Phi_\mu(s, X_s))^2] - s \right) ds.$$

Moreover, the right-hand side of (1.12) is zero if and only if  $\mu \in \text{Pr}([0, 1])$  is the unique minimizer to (1.4).

I expect that similar results can be proved for models with soft spins, using for instance [27, Proposition 3.2] and [28, Corollary 1.3]. The formulas for the limit free energy given in [29] and [28] in this case are of a form similar to the infimum in (1.4), but with an additional parameter dependence that we then maximize over; this last step is meant to control the norm of the configuration  $\sigma \in \mathbb{R}^N$  and is in line with classical large-deviations theory. A generalization of Theorem 1 to this case would allow us to write the limit free energy as a simple supremum instead.

The present paper is inspired by a number of earlier works which we now briefly review. In [22], a dual problem was identified for the ground state energy of spherical models, and then extended to a similar representation for the free energy at any temperature in [23]. In this context, the Parisi formula for the limit free energy can be rewritten in a form called the Crisanti-Sommers formula. This formula is simpler to manipulate than (1.4), and in particular, it is manifestly convex in the variable  $\mu$ . The dual problem identified in [22, 23] takes the form of an obstacle problem.

Another series of inspiring works relates to the design of efficient algorithms that identify spin configurations whose energy is as high as possible. Important progress on this question was achieved in [38] for spherical models. In this context, the energy value reached by the algorithm can be written explicitly as a simple function of  $\xi$ . The approach was then generalized to the models we consider here in [25, 10, 36], in the form of an approximate message passing algorithm with some tunable parameters. The asymptotic energy value reached by the algorithm can be described by an explicit formula, and optimizing over the free parameters gives us a variational formula for the best possible value ALG reached by this class of algorithms. This optimization problem has a structure similar to that in (1.6). In particular, a key insight of [38] is to consider algorithms that proceed via orthogonal increments, which

in the limit yield martingales which we seek to optimize over as in (1.6). The ground state energy can be represented as a Parisi-type formula analogous to (1.4), see [4], and it was shown in [10] that the original formulation of ALG as a supremum admits a dual representation that is similar to this Parisi formula for the ground state, the only difference being that the minimization is carried over a larger set. Closely related works include [14, 19, 20] where it is shown that a large class of algorithms will fail to find a configuration with energy exceeding ALG; see also [13, 15] for surveys. A method for sampling the Gibbs measure that may also be related to the present work was introduced in [11]. Other related works include [1, 17, 18], where these optimization and sampling problems were investigated in detail for more analytically tractable models whose thermodynamics was derived in [9, 7, 8].

The starting point for the proofs of Theorems 1 and 2 is a representation of the function  $\Phi_\mu$  as the value function of a problem of stochastic control, which was introduced and developed in [6, 3, 21]. One of the main features of this representation is that it makes the convexity of the mapping  $\mu \mapsto \Phi_\mu$  transparent, while I am not aware of any alternative method for verifying this convexity property. We then seek a dual to the minimization problem in (1.4), in the sense of convex analysis, and the functional  $\Gamma$  naturally appears. The interchange of the infimum and supremum in (1.8) requires some care, since for fixed  $\mu$ , the functional  $\Gamma(\mu, \cdot)$  is not concave in general (although it is in the high-temperature regime  $\xi''(1) \leq 1$ ). We proceed by enlarging the probability space so as to “randomize” our choice of martingale, as in the construction of Nash equilibria with mixed strategies. Justifying that the supremum in (1.6) is achieved requires yet another enlargement of the probability space. However, we can then identify this optimizer explicitly according to (1.11), and in particular observe that it is measurable with respect to the filtration of Brownian motion, and so we can ultimately conclude that these enlargements of the probability space were not necessary after all.

## 2. PROOFS

We start by stating the representation of  $\Phi_\mu$  alluded to above and due to [6, 3, 21]. In order to do so, we let

$$(2.1) \quad \mathscr{W} := (C([0, 1]), (\mathcal{F}_t)_{t \in [0, 1]}, \mathbb{P})$$

be the canonical Wiener space, with canonical random variable  $(W_t)_{t \in [0, 1]}$  and with associated expectation  $\mathbb{E}$ . The  $\sigma$ -algebras  $(\mathcal{F}_t)_{t \in [0, 1]}$  are assumed to be complete. We denote by **Mart** the space of bounded martingales over  $\mathscr{W}$ , and by **Prog** the space of bounded progressive processes over  $\mathscr{W}$ . By [6, 3, 21], the quantity  $\Phi_\mu(0, h)$  defined in (1.2)-(1.3) admits the variational

representation

$$(2.2) \quad \Phi_\mu(0, h) = \sup_{\alpha \in \mathbf{Prog}} \mathbf{E} \left[ \phi \left( \int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt \right],$$

and as a consequence, the Parisi formula (1.4) can be rewritten as

$$(2.3) \quad f = \inf_{\mu \in \mathbf{Pr}([0,1])} \sup_{\alpha \in \mathbf{Prog}} \mathbf{E} \left[ \phi \left( \int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] (\alpha_t^2 + t) dt \right].$$

The supremum in (2.2) can be interpreted as a stochastic control problem. Using this interpretation, it was shown in [6, 3, 21] that the supremum in (2.2) is achieved, and an explicit description of the unique maximizer was given. It is relatively classical to verify that these results remain valid even if we enlarge the probability space and possibly allow for the control  $\alpha$  to depend on additional randomness. Notice however that it is less classical to verify that the supremum on the right side of (1.6) does not depend on the choice of probability space, or to verify that the supremum is achieved, so it will be important for our purposes to be careful about such aspects. In the next lemma, we therefore restate the identity (2.2) in the generic probability space  $\mathcal{P}$  (not necessarily the Wiener space  $\mathcal{W}$ ), and describe precisely the identity of the maximizer. We denote by **Prog** the space of bounded progressive processes over the probability space  $\mathcal{P}$ .

**Lemma 2.1** (variational representation of  $\Phi_\mu$  [6, 3, 21]). *For every  $\mu \in \mathbf{Pr}([0, 1])$ , we have*

$$(2.4) \quad \Phi_\mu(0, h) = \sup_{\alpha \in \mathbf{Prog}} \mathbf{E} \left[ \phi \left( \int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt \right].$$

Moreover, let  $(X_t)_{t \in [0,1]}$  be the strong solution to

$$(2.5) \quad \begin{cases} X_0 = h, \\ dX_t = \xi''(t) \mu[0, t] \partial_x \Phi_\mu(t, X_t) dt + \sqrt{\xi''(t)} dW_t. \end{cases}$$

A stochastic process  $\alpha \in \mathbf{Prog}$  achieves the supremum in (2.4) if and only if it satisfies, for almost every  $t \in [0, 1]$  with  $\mu[0, t] > 0$ ,

$$(2.6) \quad \alpha_t = \partial_x \Phi_\mu(t, X_t) = \partial_x \Phi_\mu(0, h) + \int_0^t \sqrt{\xi''(s)} \partial_x^2 \Phi_\mu(s, X_s) dW_s.$$

*Proof.* We learn from [3, Theorem 3] or [21, Lemma 18] that the representation (2.4) is valid if we replace **Prog** by the space of processes that are progressively measurable with respect to the  $\sigma$ -algebra generated by

the Brownian motion  $W$ ; or equivalently, if the probability space  $\mathscr{P}$  is the canonical Wiener space  $\mathscr{W}$ . An examination of the proofs given there in fact also yields the identity (2.4) in the more general context considered here. For the reader's convenience, we briefly justify this here as well. Since the supremum in (2.4) is over a potentially larger set than the one considered in [3, 21], it suffices to show that for every  $\alpha \in \mathbf{Prog}$ , we have

$$(2.7) \quad \Phi_\mu(0, h) \geq \mathbf{E} \left[ \phi \left( \int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt \right].$$

We therefore fix  $\alpha \in \mathbf{Prog}$  and for every  $t \in [0, 1]$ , set

$$Y_t := h + \int_0^t \xi''(s) \mu[0, s] \alpha_s ds + \int_0^t \sqrt{\xi''(s)} dW_s.$$

Using the regularity properties of  $\Phi_\mu$  shown in [21, Theorem 4], we can apply Itô's formula to get that

$$\begin{aligned} \Phi_\mu(1, Y_1) &= \Phi_\mu(0, h) + \int_0^1 \partial_t \Phi_\mu(t, Y_t) dt + \int_0^1 \partial_x \Phi_\mu(t, Y_t) dY_t \\ &\quad + \frac{1}{2} \int_0^1 \partial_x^2 \Phi_\mu(t, Y_t) d\langle Y \rangle_t. \end{aligned}$$

Taking the expectation, we obtain that

$$\begin{aligned} \mathbf{E}[\Phi_\mu(1, Y_1)] &= \Phi_\mu(0, h) + \mathbf{E} \int_0^1 \left( \partial_t \Phi_\mu(t, Y_t) \right. \\ &\quad \left. + \frac{\xi''(t)}{2} (2\mu[0, t] \alpha_t \partial_x \Phi_\mu(t, Y_t) + \partial_x^2 \Phi_\mu(t, Y_t)) \right) dt. \end{aligned}$$

By (1.2) and [21, Theorem 1], for almost every  $t \in [0, 1]$ , we have for every  $x \in \mathbb{R}$  that

$$(2.8) \quad \begin{aligned} -\partial_t \Phi_\mu(t, x) &= \frac{\xi''(t)}{2} \left( \partial_x^2 \Phi_\mu(t, x) + \mu[0, t] (\partial_x \Phi_\mu(t, x))^2 \right) \\ &= \frac{\xi''(t)}{2} \sup_{a \in \mathbb{R}} \left( \partial_x^2 \Phi_\mu(t, x) + 2\mu[0, t] a \partial_x \Phi_\mu(t, x) - \mu[0, t] a^2 \right) \\ &\geq \frac{\xi''(t)}{2} \left( \partial_x^2 \Phi_\mu(t, x) + 2\mu[0, t] \alpha_t \partial_x \Phi_\mu(t, x) - \mu[0, t] \alpha_t^2 \right). \end{aligned}$$

Combining this with the previous display, we infer that

$$\mathbf{E}[\Phi_\mu(1, Y_1)] \leq \Phi_\mu(0, h) + \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \mathbf{E}[\alpha_t^2] dt.$$

In view of (1.3) and (1.5), this is (2.7), so the proof of (2.4) is complete.

The second identity in (2.6) is a consequence of Itô's formula and the regularity properties of  $\Phi_\mu$  given in [21, Theorem 4]. The results of [3, Theorem 3] and [21, Lemma 18] only state that (2.6) is a sufficient condition for optimality, but again the proofs in fact yield that they are also necessary.

This is also apparent from the argument given above, by an examination of the case of equality in (2.8).  $\square$

One key point of the representation in (2.4) is that it is manifestly convex in  $\mu$ . In our next step, we rewrite this convex function as a supremum of affine functions.

**Lemma 2.2** (Supremum of affine functionals). *For every  $\mu \in \text{Pr}([0, 1])$ , we have*

$$(2.9) \quad \Phi_\mu(0, h) = \sup_{\alpha \in \mathbf{Mart}} \mathbf{E} \left[ \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt + \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) \right].$$

Moreover, there exists a unique maximizer to the variational problem in (2.9). Letting  $(X_t)_{t \in [0, 1]}$  be the solution to (2.5), this unique maximizer  $\alpha \in \mathbf{Mart}$  must satisfy (2.6) for every  $t \in [0, 1]$ .

*Proof.* We decompose the proof into two steps.

*Step 1.* Let  $X \in L^2(\mathcal{P})$ . In this step, we show that

$$(2.10) \quad \mathbf{E}[\phi(X)] = \sup_{\lambda \in L^\infty(\mathcal{P})} \mathbf{E}[\lambda X - \phi^*(\lambda)],$$

and that the supremum in (2.10) is achieved at  $\lambda \in L^\infty(\mathcal{P})$  if and only if  $\lambda = \phi'(X)$ .

By the biconjugation theorem, we have that for every  $x \in \mathbb{R}$ ,

$$(2.11) \quad \phi(x) = \sup_{\lambda \in \mathbb{R}} (\lambda x - \phi^*(\lambda)).$$

In particular, we have for every  $x, \lambda \in \mathbb{R}$  that

$$(2.12) \quad \phi(x) + \phi^*(\lambda) \geq \lambda x,$$

and thus

$$\mathbf{E}[\phi(X)] \geq \sup_{\lambda \in L^\infty(\mathcal{P})} \mathbf{E}[\lambda X - \phi^*(\lambda)].$$

As a supremum of convex and lower semi-continuous functions, the function  $\phi^*$  is convex and lower semi-continuous. Since  $\phi^*$  is in fact strictly convex on its effective domain  $[-1, 1]$ , for each  $x \in \mathbb{R}$ , there exists a unique  $\lambda \in \mathbb{R}$  such that

$$(2.13) \quad \phi(x) + \phi^*(\lambda) = \lambda x.$$

Since the derivative of  $\phi^*$  diverges at the endpoints of the interval  $[-1, 1]$ , this unique  $\lambda$  must lie in the open interval  $(-1, 1)$ . Similarly, for each  $\lambda \in (-1, 1)$ , there exists a unique  $x \in \mathbb{R}$  such that the identity (2.13) is realized, and recall that we always have (2.12) in general. In short, the condition (2.13) is equivalent to the statement that  $\phi'(x) = \lambda$ , and in particular, for every  $x \in \mathbb{R}$ ,

$$\phi(x) = \phi'(x)x - \phi^*(\phi'(x)).$$



This yields (2.10) and the fact that the unique maximizer is  $\lambda = \phi'(X)$ .

*Step 2.* We let  $(X_t)$  be the solution to (2.5), and for every  $t \in [0, 1]$ , we let

$$\widehat{\alpha}_t := \partial_x \Phi_\mu(t, X_t) = \partial_x \Phi_\mu(0, h) + \int_0^t \sqrt{\xi''(s)} \partial_x^2 \Phi_\mu(s, X_s) dW_s.$$

Using (1.3) and (1.5), we notice that

$$(2.14) \quad \widehat{\alpha}_1 = \phi'(X_1 - h) = \phi' \left( \int_0^1 \xi''(t) \mu[0, t] \widehat{\alpha}_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right).$$

By the result of the previous step and Lemma 2.1, we can rewrite  $\Phi_\mu(0, h)$  in the form of

$$(2.15) \quad \sup_{\alpha \in \mathbf{Prog}} \sup_{\lambda \in L^\infty(\mathcal{F})} \mathbf{E} \left[ \lambda \left( \int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t \right) - \phi^*(\lambda) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt \right].$$

Moreover, the choice of  $\alpha = \widehat{\alpha}$  and  $\lambda = \widehat{\alpha}_1$  realizes the supremum. Also, by Lemma 2.1, the random variable

$$\int_0^1 \xi''(t) \mu[0, t] \alpha_t dt + \int_0^1 \sqrt{\xi''(t)} dW_t$$

does not depend on the choice of optimizer  $\alpha$  in (2.15). The result of the previous step and (2.14) therefore also yield that the choice of  $\lambda = \widehat{\alpha}_1$  is necessary for any maximizing pair  $(\alpha, \lambda)$  of (2.15). If we also impose the maximizing pair to be such that  $\alpha \in \mathbf{Mart}$  and  $\alpha_1 = \lambda$ , then by the martingale property this yields that  $\alpha = \widehat{\alpha}$ . In the case when these additional constraints are imposed, the functional we optimize over in (2.15) reduces to that in (2.9), so the proof is complete.  $\square$

By (1.4) and Lemma 2.2, this already shows that

$$(2.16) \quad f = \inf_{\mu \in \text{Pr}([0,1])} \sup_{\alpha \in \mathbf{Mart}} \Gamma(\mu, \alpha).$$

Since the minimization problem over  $\mu$  is in fact the same as that in (1.4), it is achieved at the same optimal  $\mu$  and only there, by [3]. Moreover, once  $\mu$  is fixed, the maximization over  $\alpha$  is achieved at the optimal  $\alpha$  described by Lemma 2.2 and only there.

Our next goal is to justify the interversion of the infimum and supremum in (2.16). For every bounded measurable  $f : [0, 1] \rightarrow \mathbb{R}$ , we observe the integration by parts

$$(2.17) \quad \int_0^1 f(s) \mu[0, s] ds = \int_0^1 f(s) \int \mathbf{1}_{\{t \leq s\}} d\mu(t) ds \\ = \int \int_t^1 f(s) ds d\mu(t).$$

It will be convenient to use this observation and rewrite (2.16) as

$$(2.18) \quad f = \inf_{\mu \in \text{Pr}([0,1])} \sup_{\alpha \in \text{Mart}} \mathbf{E} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) + \frac{1}{2} \int \int_t^1 \xi''(s)(\alpha_s^2 - s) ds d\mu(t) \right].$$

**Lemma 2.3** (Interchanging inf and sup). *Let*

$$(2.19) \quad \mathbf{K}_0 := \left\{ \left( \mathbf{E} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) \right], (\mathbf{E}[\alpha_t^2])_{t \in [0,1]} \right) \mid \alpha \in \text{Mart}, \mathbf{E}[\phi^*(\alpha_1)] < +\infty \right\},$$

and let  $\mathbf{K}$  be the closure of the convex hull of  $\mathbf{K}_0$  with respect to the topology of the space  $\mathbb{R} \times L^1([0,1])$ . The limit free energy (1.1) is given by

$$(2.20) \quad f = \sup_{(\chi, \gamma) \in \mathbf{K}} \left\{ \chi + \frac{1}{2} \inf_{t \in [0,1]} \int_t^1 \xi''(s)(\gamma_s - s) ds \right\}.$$

*Proof of Lemma 2.3.* We can rewrite (2.18) in the form of

$$(2.21) \quad f = \inf_{\mu \in \text{Pr}([0,1])} \sup_{(\chi, \gamma) \in \mathbf{K}_0} \left( \chi + \frac{1}{2} \int \int_t^1 \xi''(s)(\gamma_s - s) ds d\mu(t) \right).$$

Let us call  $G(\mu, \chi, \gamma)$  the functional between the large parentheses in (2.21). For each  $\mu$ , the functional  $G(\mu, \cdot, \cdot)$  is affine, so the supremum in (2.21) is not changed if we replace  $\mathbf{K}_0$  by its convex hull. For every  $\gamma, \rho \in L^1([0,1])$ ,

$$(2.22) \quad \sup_{t \in [0,1]} \left| \int_t^1 \xi''(s)(\gamma_s - s) ds - \int_t^1 \xi''(s)(\rho_s - s) ds \right| \leq \xi''(1) \int_0^1 |\gamma_s - \rho_s| ds.$$

The functional  $G(\mu, \cdot, \cdot)$  is therefore continuous for the topology of  $\mathbb{R} \times L^1([0,1])$ . We deduce that the supremum in (2.21) is not changed if we replace  $\mathbf{K}_0$  by  $\mathbf{K}$ , that is,

$$(2.23) \quad f = \inf_{\mu \in \text{Pr}([0,1])} \sup_{(\chi, \gamma) \in \mathbf{K}} \left( \chi + \frac{1}{2} \int \int_t^1 \xi''(s)(\gamma_s - s) ds d\mu(t) \right).$$

We now argue that the set  $\mathbf{K}_0$  is precompact. First, for every  $\alpha \in \text{Mart}$ , the condition  $\mathbf{E}[\phi^*(\alpha_1)] < +\infty$  is equivalent to the condition that  $\alpha_1$  takes values in  $[-1, 1]$  almost surely. As a result, if  $(\chi^{(n)}, \gamma^{(n)})$  denotes a sequence of elements of  $\mathbf{K}_0$ , we have that  $\chi^{(n)}$  and  $\sup_{t \in [0,1]} |\gamma^{(n)}(t)|$  are bounded uniformly over  $n$ . We can therefore find a subsequence along which  $\chi^{(n)}$  and  $\gamma_t^{(n)}$  converge for each  $t \in \mathbb{Q} \cap [0, 1]$ . Using also that for each  $n$ , the mapping  $t \mapsto \gamma_t^{(n)}$  is non-decreasing, we deduce that  $(\chi^{(n)}, \gamma^{(n)})$  converges

in  $\mathbb{R} \times L^1([0, 1])$  along the subsequence. This shows that  $K_0$  is precompact, and therefore that its closed convex hull  $K$  is compact, by [2, Theorem 5.35].

Endowing the space of probability measures  $\Pr([0, 1])$  with the topology of weak convergence turns this space into a compact set. Using again (2.22), we see that  $G$  is jointly continuous. We have already used that for each fixed  $\mu \in \Pr([0, 1])$ , the mapping  $G(\mu, \cdot, \cdot)$  is affine; and for each fixed  $(\chi, \gamma) \in \mathbb{R} \times L^1([0, 1])$ , the mapping  $G(\cdot, \chi, \gamma)$  is affine. By the minimax theorem [12, 37], we can therefore exchange the infimum and the supremum in (2.23) and obtain that

$$f = \sup_{(\chi, \gamma) \in K} \inf_{\mu \in \Pr([0, 1])} \left( \chi + \frac{1}{2} \int \int_t^1 \xi''(s)(\gamma_s - s) ds d\mu(t) \right).$$

The infimum is achieved for measures  $\mu$  that are supported on minimizers of the mapping

$$t \mapsto \int_t^1 \xi''(s)(\gamma_s - s) ds.$$

This yields the announced result.  $\square$

While Lemma 2.3 does indeed perform some interchange of infimum and supremum, the expression of the optimization as a supremum over the set  $K$  is not very satisfactory. The goal of the next lemma is to revert this back to an optimization problem over the space of martingales. The proof of this lemma will be simplified if we assume that the  $\sigma$ -algebra  $\mathcal{F}_0$  of the probability space  $\mathcal{P}$  is sufficiently rich, so that we can perform a “randomization” operation on the space of martingales, as in the construction of Nash equilibria. We could also show the next lemma directly without this assumption, by showing that the “randomized” martingales we use can be approximated arbitrarily closely by martingales that are measurable with respect to the filtration generated by the Brownian motion  $W$ . Since we will ultimately show that the supremum is in fact achieved at a martingale that is measurable with respect to the filtration generated by  $W$ , there is ultimately no loss in making this additional assumption here.

**Lemma 2.4.** *Suppose that the  $\sigma$ -algebra  $\mathcal{F}_0$  of the probability space  $\mathcal{P}$  is sufficiently rich that there exists an  $\mathcal{F}_0$ -measurable random variable that is uniformly distributed over  $[0, 1]$ . Then the identity (1.6) is valid.*

*Proof.* Lemma 2.3 clearly implies that

$$(2.24) \quad f \geq \sup_{\alpha \in \text{Mart}} \left\{ \mathbf{E} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) \right] - \frac{1}{2} \sup_{t \in [0, 1]} \int_t^1 \xi''(s)(s - \mathbf{E}[\alpha_s^2]) ds \right\},$$

so we only need to show the converse bound. Notice carefully that the definition of  $K_0$  in (2.19) depends on the identity of the probability space  $\mathcal{P}$ . For the bound converse to (2.24), we will appeal to Lemma 2.3 applied to the

case when the underlying probability space is the canonical Wiener space  $\mathscr{W}$ . We recall that we denote by  $\mathbf{E}$  the expectation over  $\mathscr{W}$ , and by  $\mathbf{Mart}$  the space of bounded martingales over  $\mathscr{W}$ , while we denote by  $\mathbf{Mart}$  the space of bounded martingales over  $\mathscr{P}$ . There is a canonical injection from  $\mathbf{Mart}$  to  $\mathbf{Mart}$  given by  $\alpha \mapsto \alpha \circ W$ ; but the space  $\mathbf{Mart}$  may be larger than  $\mathbf{Mart}$  in general. We let  $\mathbf{K}_0$  be as in (2.19) but with the underlying probability space being  $\mathscr{W}$ , we let  $\mathbf{K}_1$  be the convex hull of  $\mathbf{K}_0$ , and  $\mathbf{K}$  be the closure of  $\mathbf{K}_1$ . By (2.22), the representation in (2.20) is still valid if we replace  $\mathbf{K}$  by  $\mathbf{K}_1$ . Moreover, every element of  $(\chi, \gamma) \in \mathbf{K}_1$  can be represented in the form

$$(\chi, \gamma) = \sum_{i=1}^n c_i \left( \mathbf{E} \left[ \alpha_1^{(i)} \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1^{(i)}) \right], \left( \mathbf{E}[(\alpha_t^{(i)})^2] \right)_{t \in [0,1]} \right),$$

where  $c_1, \dots, c_n \in [0, 1]$  are such that  $\sum_i c_i = 1$  and  $\alpha^{(1)}, \dots, \alpha^{(n)} \in \mathbf{Mart}$  take values in  $[-1, 1]$ . We identify the martingales  $\alpha^{(1)}, \dots, \alpha^{(n)}$  with elements of  $\mathbf{Mart}$  through the injection  $\alpha \mapsto \alpha \circ W$  mentioned above. Notice that by construction, the martingales  $\alpha^{(1)}, \dots, \alpha^{(n)}$  are independent of  $\mathcal{F}_0$ . Under our assumption on  $\mathscr{P}$ , there exists an  $\mathcal{F}_0$ -measurable random variable  $N$  such that for every  $i \in \{1, \dots, n\}$ ,

$$\mathbf{P}[N = i] = c_i.$$

For every  $t \in [0, 1]$ , we set

$$\beta_t := \alpha_t^{(N)}.$$

Since  $N$  is  $\mathcal{F}_0$ -measurable and the martingales  $\alpha^{(1)}, \dots, \alpha^{(n)}$  are independent of  $\mathcal{F}_0$ , we see that  $\beta \in \mathbf{Mart}$ , and a direct calculation gives that

$$\left( \mathbf{E} \left[ \beta_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\beta_1) \right], \left( \mathbf{E}[(\beta_t)^2] \right)_{t \in [0,1]} \right) = (\chi, \gamma).$$

So we have shown that every pair  $(\chi, \gamma) \in \mathbf{K}_1$  can be represented in the form above. Since we have also observed that

$$f = \sup_{(\chi, \gamma) \in \mathbf{K}_1} \left\{ \chi + \frac{1}{2} \inf_{t \in [0,1]} \int_t^1 \xi''(s) (\gamma_s - s) ds \right\},$$

and recalling also (2.24), we obtain the result.  $\square$

We continue by showing the existence of a martingale that achieves the supremum in (1.6), at the cost of possibly changing the probability space again.

**Lemma 2.5** (Existence of maximizing martingale). *There exists a probability space  $\mathscr{P}$  such that the identity (1.6) is valid and the supremum appearing there is achieved.*

*Proof.* Let  $\mathcal{P}$  be a probability space satisfying the assumption of Lemma 2.4, and let  $(\alpha^{(n)})_{n \in \mathbb{N}}$  be a sequence of elements of **Mart** such that

$$f = \lim_{n \rightarrow \infty} \left\{ \mathbf{E} \left[ \alpha_1^{(n)} \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1^{(n)}) \right] + \frac{1}{2} \inf_{t \in [0,1]} \int_t^1 \xi''(s) (\mathbf{E}[(\alpha_s^{(n)})^2] - s) ds \right\}.$$

For  $n$  sufficiently large, the martingale  $\alpha^{(n)}$  must take values in  $[-1, 1]$ . In particular, for each  $t \in [0, 1]$ , the family of random variables  $(\alpha_t^{(n)})_{n \in \mathbb{N}}$  is tight. By Prokhorov's theorem, we can therefore find a subsequence  $(k_n)_{n \in \mathbb{N}}$ , a probability space  $\overline{\mathcal{P}} = (\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbf{P}})$  and random variables  $(W, (\alpha_t)_{t \in \mathbb{Q} \cap [0,1]})$  over  $\overline{\mathcal{P}}$  such that for every integer  $\ell \geq 1$ ,  $t_1, \dots, t_\ell \in \mathbb{Q} \cap [0, 1]$ , and bounded continuous function  $G : C([0, 1]) \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ ,

$$(2.25) \quad \lim_{n \rightarrow \infty} \mathbf{E} \left[ G(W, \alpha_{t_1}^{(k_n)}, \dots, \alpha_{t_\ell}^{(k_n)}) \right] = \overline{\mathbf{E}} \left[ G(W, \alpha_{t_1}, \dots, \alpha_{t_\ell}) \right].$$

Without loss of generality we assume from now on that the convergence above is valid along the full sequence, that is, we can take  $k_n = n$ . For every  $t \in [0, 1]$ , we let  $\overline{\mathcal{F}}_t$  be the  $\sigma$ -algebra on  $\overline{\mathcal{P}}$  generated by the random variables  $(W_s)_{s \leq t}$  and  $(\alpha_s)_{s \in \mathbb{Q} \cap [0, t]}$ . The process  $W$  is a Brownian motion with respect to this filtration, and  $(\alpha_t)_{t \in \mathbb{Q} \cap [0, 1]}$  is a martingale with respect to  $(\overline{\mathcal{F}}_t)_{t \in \mathbb{Q} \cap [0, 1]}$ . We extend  $\alpha$  by setting, for every  $t \in [0, 1]$ ,

$$\alpha_t := \overline{\mathbf{E}}[\alpha_1 | \overline{\mathcal{F}}_t].$$

This turns  $\alpha$  into a martingale with respect to the filtration  $(\overline{\mathcal{F}}_t)_{t \in [0, 1]}$ . Since  $\alpha_1^{(n)}$  takes values in  $[-1, 1]$  and  $\phi^*$  is continuous over this interval, we have that

$$\lim_{n \rightarrow \infty} \mathbf{E}[\phi^*(\alpha_1^{(n)})] = \overline{\mathbf{E}}[\phi^*(\alpha_1)].$$

In order to control the continuity of the linear mapping  $W \mapsto \int_0^1 \sqrt{\xi''(t)} dW_t$ , we introduce the notation

$$\zeta(t) := \sqrt{\xi''(t)}.$$

We observe that  $\zeta$  is continuously differentiable on  $[0, 1]$  whenever  $\xi''(0) \neq 0$  or  $\xi''(0) = \xi'''(0) = 0$ , while in all cases, we can find a constant  $C < \infty$  such that for every  $t \in (0, 1]$ ,

$$\zeta'(t) \leq Ct^{-1/2}.$$

Integrating by parts (or applying Itô's formula), we have that

$$\zeta(1)W_1 = \int_0^1 \zeta'(t)W_t dt + \int_0^1 \zeta(t) dW_t,$$

and in particular, for a possibly larger value of the constant  $C < \infty$ , we have that for every  $W \in C([0, 1])$ ,

$$\left| \int_0^1 \zeta(t) dW_t \right| \leq C \|W\|_{L^\infty([0, 1])}.$$

From this, (2.25), the fact that  $\alpha_1^{(n)}$  takes values in  $[-1, 1]$ , and an approximation argument, we deduce that

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[ \alpha_1^{(n)} \int_0^1 \sqrt{\xi''(t)} \, dW_t \right] = \overline{\mathbf{E}} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} \, dW_t \right].$$

It also follows from (2.25) that for every  $t \in \mathbb{Q} \cap [0, 1]$ ,

$$\lim_{n \rightarrow \infty} \mathbf{E} [(\alpha_t^{(n)})^2] = \overline{\mathbf{E}}[\alpha_t^2].$$

Since the paths  $t \mapsto \mathbf{E}[(\alpha_t^{(n)})^2]$  and  $t \mapsto \overline{\mathbf{E}}[\alpha_t^2]$  are non-decreasing, one can extend this convergence to all points of continuity of the latter mapping, which form a subset of  $[0, 1]$  of full measure. Using also (2.22), we deduce that

$$\lim_{n \rightarrow \infty} \inf_{t \in [0, 1]} \int_t^1 \xi''(s) (\mathbf{E}[(\alpha_s^{(n)})^2] - s) \, ds = \inf_{t \in [0, 1]} \int_t^1 \xi''(s) (\overline{\mathbf{E}}[\alpha_s^2] - s) \, ds.$$

In conclusion, we have constructed a martingale  $\alpha$  over  $\overline{\mathcal{P}}$  such that

$$f = \overline{\mathbf{E}} \left[ \alpha_1 \int_0^1 \sqrt{\xi''(t)} \, dW_t - \phi^*(\alpha_1) \right] + \frac{1}{2} \inf_{t \in [0, 1]} \int_t^1 \xi''(s) (\overline{\mathbf{E}}[\alpha_s^2] - s) \, ds.$$

Since the converse inequality (2.24) is valid on every probability space, this completes the proof.  $\square$

We are now ready to prove the main results of this paper.

*Proof of Theorems 1 and 2.* Let  $\mathcal{P}$  be a probability space such that the conclusion of Lemma 2.5 is valid. We denote by  $\overline{\alpha} \in \mathbf{Mart}$  a martingale that achieves the supremum in (1.6), and we denote by  $\overline{\mu} \in \text{Pr}([0, 1])$  the probability measure that achieves the infimum in (1.4). We have already observed in (2.16) and in the paragraph below it that by [3], there is exactly one such choice of  $\overline{\mu}$ , and

$$f = \inf_{\mu \in \text{Pr}([0, 1])} \sup_{\alpha \in \mathbf{Mart}} \Gamma(\mu, \alpha) = \sup_{\alpha \in \mathbf{Mart}} \Gamma(\overline{\mu}, \alpha),$$

while Lemma 2.5 and the integration by parts (2.17) state that

$$f = \sup_{\alpha \in \mathbf{Mart}} \inf_{\mu \in \text{Pr}([0, 1])} \Gamma(\mu, \alpha) = \inf_{\mu \in \text{Pr}([0, 1])} \Gamma(\mu, \overline{\alpha}).$$

In particular, for every  $\mu \in \text{Pr}([0, 1])$  and  $\alpha \in \mathbf{Mart}$ , we have

$$\Gamma(\overline{\mu}, \alpha) \leq f \quad \text{and} \quad f \leq \Gamma(\mu, \overline{\alpha}),$$

and thus  $f = \Gamma(\overline{\mu}, \overline{\alpha})$  and

$$(2.26) \quad \Gamma(\overline{\mu}, \alpha) \leq \Gamma(\overline{\mu}, \overline{\alpha}) \leq \Gamma(\mu, \overline{\alpha}).$$

The first optimality condition in (2.26) and Lemma 2.2 imply that  $\overline{\alpha}$  must be as described in the statement of Theorem 2. Notice that this martingale is measurable with respect to the filtration generated by the Brownian motion  $W$ . In other words, there exists  $\widehat{\alpha} \in \mathbf{Mart}$  such that the martingale  $\overline{\alpha} \in \mathbf{Mart}$  is the image of  $\widehat{\alpha}$  under the canonical injection  $\alpha \mapsto \alpha \circ W$  from  $\mathbf{Mart}$

to **Mart** (recall that **Mart** denotes the space of bounded martingales over the Wiener space  $\mathscr{W}$ ). The martingale  $\widehat{\alpha}$  has a canonical image in the space of bounded martingales of every probability space  $\mathscr{P}$  under consideration in Theorem 1. Recalling also from (2.24) that the converse inequality is valid in every probability space, this completes the proof of Theorem 1. The fact that the support of  $\bar{\mu}$  is a subset of the set of maximizers of the mapping in (1.9) is a consequence of the second optimality condition in (2.26) and the integration by parts in (2.17). To show that the properties (1) and (2) in the statement of Theorem 2 characterize the pair  $(\bar{\mu}, \bar{\alpha})$  uniquely, we observe that these conditions imply that (2.26) is valid for every  $\mu \in \text{Pr}([0, 1])$  and  $\alpha \in \mathbf{Mart}$ , using again the integration by parts (2.17) and Lemma 2.2, so we must have

$$f = \inf_{\mu \in \text{Pr}([0,1])} \sup_{\alpha \in \mathbf{Mart}} \Gamma(\mu, \alpha) \geq \inf_{\mu \in \text{Pr}([0,1])} \Gamma(\mu, \bar{\alpha}) = \Gamma(\bar{\mu}, \bar{\alpha})$$

as well as

$$f = \sup_{\alpha \in \mathbf{Mart}} \inf_{\mu \in \text{Pr}([0,1])} \Gamma(\mu, \alpha) \leq \sup_{\alpha \in \mathbf{Mart}} \Gamma(\bar{\mu}, \alpha) = \Gamma(\bar{\mu}, \bar{\alpha}).$$

Hence  $f = \Gamma(\bar{\mu}, \bar{\alpha})$ , the probability measure  $\bar{\mu}$  must be the unique minimizer of (1.4), and the martingale  $\bar{\alpha}$  must be the unique maximizer of (1.6).  $\square$

*Proof of Theorem 3.* Let  $\mu \in \text{Pr}([0, 1])$ , let  $(X_t)_{t \geq 0}$  be the strong solution to (2.5), let  $\alpha \in \mathbf{Mart}$  be such that (2.6) holds for every  $t \in [0, 1]$ , and let  $g_\mu$  be such that, for every  $t \in [0, 1]$ ,

$$g_\mu(t) := \int_t^1 \xi''(s) (\mathbf{E}[\alpha_s^2] - s) ds.$$

By Lemma 2.2, we have that

$$\Phi_\mu(0, h) = \mathbf{E} \left[ \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \alpha_t^2 dt + \alpha_1 \int_0^1 \sqrt{\xi''(t)} dW_t - \phi^*(\alpha_1) \right].$$

Combining this with Theorem 1 yields that

$$f \geq \Phi_\mu(0, h) - \frac{1}{2} \int_0^1 \xi''(t) \mu[0, t] \mathbf{E}[\alpha_t^2] dt + \frac{1}{2} \inf_{[0,1]} g_\mu.$$

The integration by parts (2.17) therefore yields the second inequality in (1.12); we recall that the first one follows from (1.4). By Theorem 2, if we choose  $\mu$  to be the minimizer of (1.4), then the support of  $\mu$  is a subset of the set of minimizers of  $g_\mu$ , so the right-hand side of (1.12) vanishes in this case. Conversely, if  $\mu \in \text{Pr}([0, 1])$  is such that the right-hand side of (1.12) vanishes, then it must clearly be the unique minimizer of (1.4).  $\square$

## REFERENCES

- [1] Louigi Addario-Berry and Pascal Maillard. The algorithmic hardness threshold for continuous random energy models. *Math. Stat. Learn.*, 2(1):77–101, 2019.
- [2] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006.

- [3] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.
- [4] Antonio Auffinger and Wei-Kuo Chen. Parisi formula for the ground state energy in the mixed  $p$ -spin model. *Ann. Probab.*, 45(6B):4617–4631, 2017.
- [5] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.
- [6] Anton Bovier and Anton Klimovsky. The Aizenman-Sims-Starr and Guerra’s schemes for the SK model with multidimensional spins. *Electron. J. Probab.*, 14:no. 8, 161–241, 2009.
- [7] Anton Bovier and Irina Kurkova. Derrida’s generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):439–480, 2004.
- [8] Anton Bovier and Irina Kurkova. Derrida’s generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004.
- [9] D. Capocaccia, Marzio Cassandro, and Pierre Picco. On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.*, 46(3-4):493–505, 1987.
- [10] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Optimization of mean-field spin glasses. *Ann. Probab.*, 49(6):2922–2960, 2021.
- [11] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic stochastic localization. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science*, pages 323–334. IEEE Computer Soc., Los Alamitos, CA, 2022.
- [12] Ky Fan. Minimax theorems. *Proc. Nat. Acad. Sci. U.S.A.*, 39:42–47, 1953.
- [13] David Gamarnik. The overlap gap property: a topological barrier to optimizing over random structures. *Proc. Natl. Acad. Sci. USA*, 118(41):e2108492118, 2021.
- [14] David Gamarnik and Aukosh Jagannath. The overlap gap property and approximate message passing algorithms for  $p$ -spin models. *Ann. Probab.*, 49(1):180–205, 2021.
- [15] David Gamarnik, Christopher Moore, and Lenka Zdeborová. Disordered systems insights on computational hardness. *J. Stat. Mech. Theory Exp.*, (11):Paper No. 114015, 41, 2022.
- [16] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.
- [17] Fu-Hsuan Ho. Sampling from the Gibbs measure of the continuous random energy model and the hardness threshold. Preprint, arXiv:2308.00857.
- [18] Fu-Hsuan Ho and Pascal Maillard. Efficient approximation of branching random walk Gibbs measures. *Electron. J. Probab.*, 27:Paper No. 75, 18, 2022.
- [19] Brice Huang and Mark Sellke. Tight Lipschitz hardness for optimizing mean field spin glasses. Preprint, arXiv:2110.07847.
- [20] Brice Huang and Mark Sellke. Algorithmic threshold for multi-species spherical spin glasses. Preprint, arXiv:2303.12172.
- [21] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.
- [22] Aukosh Jagannath and Ian Tobasco. Low temperature asymptotics of spherical mean field spin glasses. *Comm. Math. Phys.*, 352(3):979–1017, 2017.
- [23] Aukosh Jagannath and Ian Tobasco. Bounds on the complexity of replica symmetry breaking for spherical spin glasses. *Proc. Amer. Math. Soc.*, 146(7):3127–3142, 2018.
- [24] Marc Mézard, Giorgio Parisi, and Miguel Virasoro. *Spin glass theory and beyond: an introduction to the replica method and its applications*, volume 9. World Scientific Publishing Company, 1987.
- [25] Andrea Montanari. Optimization of the Sherrington-Kirkpatrick Hamiltonian. *SIAM Journal on Computing*, (0):FOCS19–1, 2021.



- [26] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.
- [27] Jean-Christophe Mourrat. The Parisi formula is a Hamilton-Jacobi equation in Wasserstein space. *Canad. J. Math.*, 74(3):607–629, 2022.
- [28] Jean-Christophe Mourrat and Dmitry Panchenko. Extending the Parisi formula along a Hamilton-Jacobi equation. *Electron. J. Probab.*, 25:Paper No. 23, 17, 2020.
- [29] Dmitry Panchenko. Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.*, 17(7):793–857, 2005.
- [30] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.
- [31] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.
- [32] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754, 1979.
- [33] Giorgio Parisi. The order parameter for spin glasses: a function on the interval 0-1. *J. Phys. A*, 13(3):1101, 1980.
- [34] Giorgio Parisi. A sequence of approximated solutions to the S-K model for spin glasses. *J. Phys. A*, 13(4):L115–L121, 1980.
- [35] Giorgio Parisi. Order parameter for spin-glasses. *Phys. Rev. Lett.*, 50(24):1946, 1983.
- [36] Mark Sellke. Optimizing mean field spin glasses with external field. Preprint, arXiv:2105.03506.
- [37] Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.
- [38] Eliran Subag. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.*, 74(5):1021–1044, 2021.
- [39] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

(Jean-Christophe Mourrat) DEPARTMENT OF MATHEMATICS, ENS LYON AND CNRS, LYON, FRANCE