



# On Large Deviations for Gibbs Measures, Mean Energy and Gamma-Convergence

Robert J. Berman<sup>1</sup>

Received: 26 October 2016 / Revised: 24 April 2018 / Accepted: 16 May 2018 /  
Published online: 14 June 2018  
© The Author(s) 2018

**Abstract** We consider the random point processes on a measure space  $(X, \mu_0)$  defined by the Gibbs measures associated with a given sequence of  $N$ -particle Hamiltonians  $H^{(N)}$ . Inspired by the method of Messer–Spohn for proving concentration properties for the laws of the corresponding empirical measures, we propose a number of hypotheses on  $H^{(N)}$  that are quite general but still strong enough to extend the approach of Messer–Spohn. The hypotheses are formulated in terms of the asymptotics of the corresponding mean energy functionals. We show that in many situations, the approach even yields a large deviation principle (LDP) for the corresponding laws. Connections to Gamma-convergence of (free) energy type functionals at different levels are also explored. The focus is on differences between positive and negative temperature situations, motivated by applications to complex geometry. The results yield, in particular, large deviation principles at positive as well as negative temperatures for quite general classes of singular mean field models with pair interactions, generalizing the 2D vortex model and Coulomb gases. In a companion paper, the results will be illustrated in the setting of Coulomb and Riesz type gases on a Riemannian manifold  $X$ , comparing with the complex geometric setting.

**Keywords** Interacting particle systems · Mean-field limit · Large deviations · Coulomb interaction

**Mathematics Subject Classification** 82C22 · 60F10

---

Communicated by Peter J. Forrester.

---

✉ Robert J. Berman  
robertb@chalmers.se

<sup>1</sup> Department of Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, 412 96 Göteborg, Sweden

## 1 Introduction

Let  $X$  be a compact topological space endowed with a probability measure  $\mu_0$ . Given a sequence of symmetric functions  $H^{(N)}$  on the  $N$ -fold products  $X^N$ , which are absolutely integrable with respect to the Borel measure  $\mu_0^{\otimes N}$ , the corresponding *Gibbs measures* at inverse temperature  $\beta_N \in \mathbb{R}$  are defined as the following sequence of symmetric probability measures on  $X^N$  :

$$\mu_{\beta_N}^{(N)} := \frac{1}{Z_{N,\beta_N}} e^{-\beta_N H^{(N)}} \mu_0,$$

assuming that the partition function  $Z_{N,\beta_N}$  is finite:

$$Z_{N,\beta_N} := \int_{X^N} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} < \infty.$$

We also assume that the following limit exists:

$$\beta := \lim_{N \rightarrow \infty} \beta_N \in ] - \infty, \infty ].$$

The ensemble  $(X^N, \mu_{\beta_N}^{(N)})$  (called the *canonical ensemble*) defines a random point process with  $N$  particles on  $X$  which, from the point of view of statistical mechanics, models  $N$  identical particles on  $X$  interacting by the *Hamiltonian* (interaction energy)  $H^{(N)}$  in thermal equilibrium at *inverse temperature*  $\beta_N$ . The corresponding *empirical measure* is the random measure

$$\delta_N : X^N \rightarrow \mathcal{P}(X), \quad (x_1, \dots, x_N) \mapsto \delta_N(x_1, \dots, x_N) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \quad (1.1)$$

taking values in the space  $\mathcal{P}(X)$  of all probability measures on  $X$ . A recurrent theme in statistical mechanics is to study the large  $N$ -limit (i.e., the “macroscopic limit”) of the canonical ensemble through the large  $N$ -limit of the laws of  $\delta_N$ , i.e. through the sequence of probability measures

$$\Gamma_N := (\delta_N)_* \mu_{\beta_N}^{(N)} \quad (1.2)$$

on  $\mathcal{P}(X)$ . In many situations, the laws  $\Gamma_N$  can be shown to concentrate, as  $N \rightarrow \infty$ , at the subset of  $\mathcal{P}(X)$  consisting of the minima of a free energy type functional  $F_\beta$  on  $\mathcal{P}(X)$ ; we will then say that “*the sequence  $\Gamma_N$  has the concentration property.*” For example, if the functional  $F_\beta$  has a unique minimizer  $\mu_\beta$ , then it follows that the random measures  $\delta_N$  converge in law to a unique deterministic measure  $\mu_\beta$ . A stronger exponential notion of concentration—with an explicit speed and rate functional—is offered by the theory of large deviations, by demanding that the laws  $\Gamma_N$  satisfy a *large deviation principle (LDP)* with *speed*  $r_N$  and a *rate functional*  $F$ , symbolically expressed as

$$\Gamma_N(\mu) \sim e^{-r_N F(\mu)}, \quad N \rightarrow \infty.$$

The present paper is inspired by the method introduced by Messer–Spohn [28] to establish the concentration property of the laws of the empirical measures  $\delta_N$  using the Gibbs variational principle, combined with properties of the mean energy of the system. In the original approach in [28],  $H^{(N)}$  was assumed to be the *mean field Hamiltonian* corresponding to a continuous pair interaction potential  $W(x, y)$  :

$$H^{(N)}(x_1, \dots, x_N) := \frac{1}{N} \sum_{1 \leq i, j \leq N} W(x_i, x_j),$$

where  $\beta_N = \beta \in ]0, \infty[$  (this is a mean field interaction in the sense that each particle  $x_i$  is exposed to the average of the pair interactions  $W(x_i, x_j)$  of all  $N$  particles, including the self-interaction). But the approach has also been extended to handle some situations where  $W(x, y)$  is allowed to be singular [15, 22, 24], as in Onsager’s vortex model for 2D turbulence [29], where  $W(x, y) = -\log|x - y|$ . The corresponding mean field Hamiltonian is then “renormalized” by removing the self-interaction terms in order to make sure that  $H^{(N)}$  is generically finite on  $X^N$ . The aim of the present paper is to:

- Propose a number of quite general hypotheses on  $H^{(N)}$ , formulated in terms of the corresponding *mean energy functional*  $E^{(N)}$  on  $\mathcal{P}(X^N)$  (defined by formula 1.3) that are strong enough to extend the approach of Messer–Spohn.
- Show that the approach also yields the stronger exponential concentration property in the sense of an LDP, almost “for free,” in several situations.
- Explore some relations to the notion of Gamma-convergence of functionals: first by reformulating the approach of Messer–Spohn in terms of Gamma-convergence of the induced *mean free energy functionals*  $F_{\beta_N}^{(N)}$  on  $\mathcal{P}(\mathcal{P}(X))$  (formula (2.4)) and then by deducing a Gamma-convergence result for the sequence  $H^{(N)}/N$  on  $X^N$ .

The main motivation comes from the probabilistic approach to the construction of Kähler–Einstein metrics on a complex algebraic manifold  $X$ , introduced in [4, 5]. In that situation, the corresponding Hamiltonians  $H^{(N)}$  are highly nonlinear and singular (and not of the simpler mean field type appearing in formula (1.5), which has previously been used in the statistical mechanical approach to conformal geometry introduced in [23]). But still, as shown in [5], building on [7], the sequence  $H^{(N)}/N$  Gamma-converges towards a certain energy type functional  $E(\mu)$  on  $\mathcal{P}(X)$ . Exploiting superharmonicity properties of  $H^{(N)}/N$ , the corresponding LDP is then established at any positive inverse temperature  $\beta$  in [5], producing Kähler–Einstein metrics with negative Ricci curvature in the large  $N$ -limit. The approach in [5] bypasses the problem of the existence of the macroscopic mean energy (hypothesis H1 below), which is open in the complex geometric setting. On the other hand, as discussed in [5], extending the LDP in [5] to negative  $\beta$ —which is needed to produce Kähler–Einstein metrics with *positive* Ricci curvature—necessitates the existence of the macroscopic mean energy. This is the motivation behind Theorem 1.3 below, which shows that, conversely, hypothesis H1 together with the additional hypotheses H4, implies an

LDP for appropriate negative  $\beta$ . The hypothesis H4 is inspired by the energy-entropy compactness results in [9], which can be viewed as the macroscopic analog of H4 in the complex geometric setting. Incidentally, in the lowest-dimensional case when  $X$  is a Riemann surface, the probabilistic setting in [4, 5] is essentially equivalent to a mean field model with a logarithmic pair interaction, which is thus similar to Onsager's vortex model for 2D turbulence [29]. In the latter situation, the corresponding concentration properties were established in [15, 22], for any  $\beta$  above the critical negative temperature (and the LDP was established using a different method in [12]). This is in line with Onsager's prediction of the existence of macroscopic negative temperature states.

Another motivation for the present paper comes from random matrix theory (or more generally Coulomb gases), which can be viewed as a vortex type model with  $\beta_N \sim N$  (and in particular  $\beta = \infty$ ). The corresponding concentration property was established in [24], using the method of Messer–Spohn as in [15, 22]. Here we observe that, with a simple modification, the concentration property can be upgraded to an LDP (Corollary 1.2). In particular, this allows one to dispense with some technical assumptions (such as regularity properties of the corresponding pair interaction  $W(x, y)$  away from the diagonal) used in the different approaches to LDPs in [1, 2, 16, 34] (see Sect. 1.3).<sup>1</sup>

Yet another motivation comes from approximation and sampling theory and, in particular, the problem of finding nearly minimal configurations for a given energy type interaction on a Riemannian manifold, in the spirit of [20, 32].

Let us also point out that that the restriction that  $X$  be compact can be removed if suitable growth-assumptions of  $H^{(N)}$  at infinity are made, as in the settings in  $\mathbb{R}^n$  considered in [1, 2, 16, 18, 24, 34] (using appropriate tightness estimates). But in order to (hopefully) convey the conceptual simplicity of the arguments we stick with a compact  $X$ .

## 1.1 Hypotheses

We may as well assume that  $X$  coincides with the support of  $\mu_0$ . In the following,  $\mu^{(N)}$  will denote a symmetric probability measure on  $X^N$  and  $Y := \mathcal{P}(X)$ . We recall that the *mean (microscopic) energy* of  $\mu^{(N)}$ , in the usual sense of statistical mechanics, is defined by

$$E^{(N)}(\mu^{(N)}) := \frac{1}{N} \int_{X^N} H^{(N)} \mu^{(N)} \quad (1.3)$$

(assuming that  $H^{(N)} \in L^1(\mu^{(N)})$ ). We introduce the following hypotheses:

- (H1) “Existence of a macroscopic mean energy  $E(\mu)$ ”: There exists a functional  $E(\mu)$  on  $\mathcal{P}(X)$  such that for any  $\mu$  in  $\mathcal{P}(X)$  satisfying  $E(\mu) < \infty$ ,

$$\lim_{N \rightarrow \infty} E^{(N)}(\mu^{\otimes N}) = E(\mu).$$

<sup>1</sup> The Hamiltonians in the random matrix and Coulomb gas literature are usually scaled in a different way so that our zero-temperature ( $\beta = \infty$ ) corresponds to a fixed inverse temperature.

Moreover,  $E(\mu_0) < \infty$ .

- (H2) “Lower bound on the mean energy”: For any sequence of  $\mu^{(N)}$  such that  $\Gamma_N := (\delta_N)_* \mu^{(N)} \rightarrow \Gamma$  weakly in  $\mathcal{P}(Y)$ , we have

$$\liminf_{N \rightarrow \infty} E^{(N)}(\mu^{(N)}) \geq E(\Gamma) := \int_Y E(\mu) \Gamma(\mu).$$

- (H3) “Approximation property”: For any  $\mu$  such that  $E(\mu) < \infty$ , there exists a sequence  $\mu_j$  converging weakly to  $\mu$  such that  $\mu_j$  has finite entropy with respect to  $\mu_0$  and satisfies  $E(\mu_j) \rightarrow E(\mu)$ .
- (H4) “Mean energy/entropy compactness”: If

$$E^{(N)}(\mu^{(N)}) \leq C, \quad D^{(N)}(\mu^{(N)}) \leq C,$$

where  $D^{(N)}(\mu^{(N)})$  is the mean entropy, then the following convergence holds, after perhaps replacing  $\mu^{(N)}$  by a subsequence such that  $\Gamma_N := (\delta_N)_* \mu^{(N)} \rightarrow \Gamma$  weakly in  $\mathcal{P}(Y)$  :

$$\lim_{N \rightarrow \infty} E^{(N)}(\mu^{(N)}) = \int_Y E(\mu) \Gamma(\mu).$$

The first hypothesis will be assumed throughout the paper. The second and third ones will appear naturally in positive and vanishing temperature, respectively, while the fourth one turns out to be useful in some case of negative temperature. However, it may very well be that hypothesis H4 needs to be weakened a bit in order to increase its scope. For example, in the proof of the large  $N$ -concentration properties, one only needs to assume that H4 holds when  $\mu^{(N)}$  is the Gibbs measure corresponding to  $H^{(N)}$  (see Remark 2.11).

Of course, the sign of the temperature may be switched by replacing  $H^{(N)}$  with  $-H^{(N)}$ , but the point is that, in practice, we will consider settings where the sign of  $H^{(N)}$  is fixed by the requirement that  $H^{(N)}$  be bounded from below (which essentially means that the system is assumed to be stable at zero temperature).

### 1.2 Large Deviation Results

We start with the simpler setting of positive temperature:

**Theorem 1.1** *Suppose that hypotheses H1 and H2 hold, and let  $\beta_N$  be a sequence of positive numbers tending to  $\beta \in ]0, \infty[$ . Then the measures  $(\delta_N)_*(e^{-\beta_N H^{(N)}} \mu_0^{\otimes N})$  on  $\mathcal{P}(X)$  satisfy, as  $N \rightarrow \infty$ , a large deviation principle (LDP) with speed  $\beta_N N$  and rate functional*

$$F_\beta(\mu) = E(\mu) + \frac{1}{\beta} D_{\mu_0}(\mu). \tag{1.4}$$

*Equivalently, the LDP holds for the corresponding Gibbs measures with  $F_\beta$  replaced by  $F_\beta - \inf F_\beta$ . Under the additional hypothesis ‘ H3, the result also holds when  $\beta = \infty$ .*

The previous theorem in particular applies to the following “finite order” Hamiltonians of mean field type. Given symmetric functions  $W_m$  on  $X^m$  for  $m \leq M$ , set

$$H^{(N)}(x_1, \dots, x_N) := \sum_{m=1}^M \frac{1}{N^{(m-1)}} \sum_I W_m(x_{i_1}, \dots, x_{i_m}), \quad (1.5)$$

where the inner sum runs over all multi indices  $I = (i_1, \dots, i_m)$  of length  $m$  and with the property that no two indices of  $I$  coincide. Then it is easy to verify H1 and H2 above with

$$E(\mu) := \sum_{m=1}^M \int_{X^m} W_m \mu^{\otimes m}.$$

The main case of interest is when  $M = 2$  and  $H^{(N)}$  is a sum of pair-interactions  $W(x_i, x_j)$ , scaled by  $1/N$ . But since it will require no extra effort in the proofs, we consider the more general “finite order setting.”

**Corollary 1.2** *Let  $W(x_1, \dots, x_m)$  be a symmetric lower semi-continuous function on  $X^m$  taking values in  $]-\infty, \infty]$  and  $\beta_N$  a sequence of positive numbers tending to  $\beta \in ]0, \infty[$ . Assume that the Gibbs measures  $\mu_{\beta_N}^{(N)}$  of the corresponding mean field Hamiltonians are well-defined probability measures. Then the laws  $(\delta_N)_*(\mu_{\beta}^{(N)})$  satisfy an LDP with speed  $\beta_N N$  and rate functional  $F_\beta$ , with*

$$E(\mu) := \int_{X^m} W \mu^{\otimes m}.$$

The corresponding result also holds for  $\beta = \infty$  if hypothesis H3 holds.

In the Euclidean setting and with  $M = 2$ , the previous corollary was established very recently in [18] using somewhat different methods (the results in [18] have also independently been generalized to the setting of the previous theorem and corollary in [19]).

We next turn to the case of negative temperature.

**Theorem 1.3** *Suppose that hypothesis H1 and H4 hold, and fix a negative number  $\beta_0$ . Then the following are equivalent:*

- For any  $\beta > \beta_0$ , we have  $Z_{N,\beta} \leq C_\beta^N$ .
- For any  $\beta > \beta_0$ , the measures  $(\delta_N)_*(e^{-\beta H^{(N_k)}} \mu_0^{\otimes N})$  on  $\mathcal{P}(X)$  satisfy an LDP with speed  $N$  and rate functional

$$\beta F_\beta(\mu) = \beta E(\mu) + D_{\mu_0}(\mu).$$

**Corollary 1.4** *Let  $W(x, y)$  be a symmetric function in  $L^1(X^2, \mu_0^{\otimes 2})$  such that the corresponding partition functions  $Z_{N,\beta}$  satisfy*

$$Z_{N,\beta} \leq C_\beta^N$$

for any  $N \geq 2$  and  $\beta \in ] - \infty, \infty[$ . Then the Gibbs measures  $\mu_\beta^{(N)}$  of the mean field Hamiltonians corresponding to  $W$  satisfy an LDP with speed  $N$  and rate functional  $\beta F_\beta$ , with

$$E(\mu) := \int_{X^2} W \mu^{\otimes 2}.$$

The key observation in the proof of Corollary 1.4 is that the first point in Theorem 1.3 always implies, “for free,” a uniform estimate in the Orlicz (Zygmund) space  $L^1 \text{Log} L^1$ , so that some general Orlicz space duality results [26, 30] can be exploited in order to verify the hypothesis H4.

It seems natural to ask if the previous corollary can be generalized to the case when one only assumes the integrability condition that  $Z_{2,\beta}$  be finite when  $\beta > \beta_0$ , for some (finite) negative number  $\beta_0$ . The following theorem gives an affirmative answer if one strengthens the integrability condition a bit:

**Theorem 1.5** *Let  $X$  be a compact metric space,  $W$  a lower semi-continuous symmetric measurable function on  $X^2$ , and  $\beta_0$  a negative number such that*

$$\sup_{x \in X} \int_X e^{-\beta_0 W(x,y)} \mu_0(y) < \infty.$$

*Then, for any  $\beta > \beta_0$ , the Gibbs measures  $\mu_\beta^{(N)}$  satisfy an LDP as in the previous corollary.*

Specialized to the logarithmic case of the vortex model, i.e. to the case

$$W(x, y) = - \log |x - y|,$$

the previous theorem recovers the LDP in [12] with a new proof. The proof follows closely the corresponding (weaker) concentration result for the vortex model originally established in [15, 22]. The new observation is that with a little twist, the argument in [15, 22] can be supplemented to give the LDP in question.

### 1.3 Relations to Gamma-Convergence at Different Levels

The proofs of the LDPs above are based on the Gamma-convergence of the corresponding free energy functionals  $F_{\beta_N}^{(N)}$  when viewed as functionals on the space  $\mathcal{P}(\mathcal{P}(X))$  (a similar approach is used in the dynamic setting considered in [10] where the assumptions H1 and H2 also appear naturally). Incidentally, as observed in the following corollary, the LDPs then imply the Gamma-convergence of the scaled Hamiltonians  $H^{(N)}/N$  when viewed as functionals on  $\mathcal{P}(X)$ .

**Corollary 1.6** *Suppose that the Hamiltonians  $H^{(N)}$  satisfy H1 and H2. Then  $H^{(N)}/N$  Gamma-converges towards  $E(\mu)$  on  $\mathcal{P}(X)$ . In particular, this applies to the finite order mean field Hamiltonians.*

The previous result generalizes the Gamma-convergence result in [34, Prop 2.8, Remark 2.19] for the mean field Hamiltonians corresponding to pair interactions  $W(x, y)$  (on a domain in  $\mathbb{R}^D$ ), where it was assumed that  $w(x, y)$  essentially only blows up along the diagonal (similar results also appear implicitly in [1, 2, 16]). The proofs in [1, 2, 16, 34] are based on some rather intricate combinatorial constructions, involving small cubes in  $\mathbb{R}^D$ . On the other hand, the latter results yield the stronger result that any measure  $\mu$  with a positive continuous density admits a recovery sequence  $x^{(N)}$  such that

$$\limsup_{N \rightarrow \infty} \sup_{B_{\epsilon_N}} E_N \leq E(\mu),$$

where  $B_{\epsilon_N}$  denotes the  $L^\infty$ -ball with center  $x^{(N)}$  and radius  $\epsilon_N$  of the order  $1/N^{1/D}$ . In turn, as shown in [34], the latter stronger form of Gamma-convergence implies the LDP for the corresponding Gibbs measures when  $\beta_N \gg \log N$  (which is used to make sure that  $\beta_N^{-1} N^{-1} \log \int_{B_{\epsilon_N}} \mu_0^{\otimes N} \rightarrow 0$  when  $\mu_0$  is a volume form). But in fact, as pointed out by the referee, the technical condition that  $\beta_N \gg \log N$  can be replaced by  $\beta_N \gg 1$  by a slight modification of the proof of Prop 2.5 in [16].

Relations between Gamma-convergence and large deviation principles have also been previously studied in [27] but from a somewhat different perspective (see also [13] for some related results).

#### 1.4 Applications to the Coulomb Gas on a Riemannian Manifold

In the companion paper [6], the general large deviation results above are illustrated and further developed for Coulomb and Riesz type gases on a compact  $D$ -dimensional Riemannian manifold  $(X, g)$  and more generally for suitable compact subsets  $K \subset X$  (the case when  $\mu_0$  is a volume form and  $\beta > 0$  has also independently been obtained in [19]).

Here we will only state the corresponding LDP for the Coulomb gas on  $(X, g)$  defined as follows. Let  $W(x, y)$  be  $1/2$  times the integral kernel of the inverse of the positive Laplacian  $-\Delta$  on the space of all functions in  $L^2(X, dV_g)$  with mean zero, where  $dV_g$  denotes the volume form determined by the metric  $g$ . As is classical,  $W$  is symmetric and smooth away from the diagonal, and close to the diagonal it admits the following asymptotics when  $D > 2$ :

$$W(x, y) = \frac{C_D}{d(x, y)^{(D-2)}}(1 + O(1)), \quad D > 2,$$

for a positive constant  $C_D$ . Moreover, when  $D = 2$ ,

$$W(x, y) = -\frac{1}{(4\pi)} \log d^2(x, y) + O(1), \quad D = 2.$$

In particular,  $W$  is lsc and in  $L^1(X \times X)$ . Given a probability measure  $\mu_0$  on  $X$ , the Coulomb gas on  $(X, g, \mu_0)$  at inverse temperature  $\beta_N$  is defined by the Gibbs measures



corresponding to  $(\mu_0, H^{(N)}, \beta_N)$ , where  $H^{(N)}$  is the mean field Hamiltonian corresponding to the pair interaction  $W(x, y)$ . In this setting, Corollary 1.2 and Theorem 1.5 yield, as shown in [6], the following LDP for the laws of the empirical measures of the Coulomb gas, formulated in terms of the potential theoretic properties of the measure  $\mu_0$  :

**Theorem 1.7** *Let  $(X, g)$  be a compact Riemannian manifold, and consider the Coulomb gas at inverse temperature  $\beta_N$  on  $(X, g, \mu_0)$ .*

- *When  $\beta \in ]0, \infty[$ , the LDP holds if the measure  $\mu_0$  is nonpolar.*
- *When  $\beta = \infty$ , the LDP holds if  $\mu_0$  is nonpolar and  $\mu_0$  is determining for its support  $K$ .*
- *When  $D = 2$  and  $\beta < \infty$ , the LDP holds when  $\beta > -4\pi d(\mu_0)$ , where  $d(\mu_0) \in [0, \infty[$  is the sup over all  $t > 0$  such that there exists a positive constant  $C$  (depending on  $t$ ) such that*

$$\mu_0(B_R(x)) \leq CR^t$$

as  $R \rightarrow 0$ , for any Riemannian ball  $B_R(x)$  of radius  $R$  centered at a given point  $x$  in  $X$ .

We briefly recall that a compact subset  $K \subset X$  is *polar* if it is locally contained in the  $-\infty$ -set of a local subharmonic function (or equivalently, if  $K$  has vanishing capacity). Accordingly, a measure  $\mu_0$  is said to be *nonpolar* if it does not charge any polar set. The notion of a *determining* measure  $\mu_0$  appearing in the second point above means that for any continuous function  $u$ ,

$$\|e^{\varphi-u}\|_{L^\infty(K, \mu_0)} = \sup_K e^{\varphi-u},$$

for any quasi-subharmonic function  $\varphi$  on  $X$ ; i.e.,  $\varphi$  is strongly usc and satisfies  $\Delta\varphi \geq -1$ . This notion is closely related to the notion of measures satisfying a *Bernstein–Markov property* in pluripotential theory [8, 11] and measures with *regular asymptotic behavior* in the theory of planar orthogonal polynomials [33]. For example,  $\mu_0$  can be taken to be the  $D$ -dimensional Hausdorff measure on a Lipschitz domain  $K \subset (X, g)$  or the  $(D - 1)$ -dimensional Hausdorff measure on a Lipschitz hypersurface in  $(X, g)$ . The point is that the assumption that  $\mu_0$  is nonpolar and determining implies that the hypothesis H3 is satisfied, as shown in [6] (an alternative proof of the LDP in the case when  $\beta = \infty$  can also be given using the approach in the complex geometric setting in [3]). Finally, we recall that measures satisfying  $d(\mu_0) > 0$ , as in the third point above, are sometimes called *Frostman measures* in the classical litterature (for example, the property in question holds with  $d(\mu_0) = d$  when  $\mu_0$  is the  $d$ -dimensional Hausdorff measure of a compact subset  $K$  of  $X$  of Hausdorff dimension  $d$ ).

More generally, an LDP as in the previous theorem is obtained in [6], when  $W(x, y)$  is taken as the integral kernel of the inverse of  $(-\Delta)^p$  and the (possible fractional) power  $p$  is in  $]0, D/2[$  (or even more generally: when  $(-\Delta)^p$  is replaced by a suitable pseudodifferential operator of order at most  $D$ ). Then the last point in the previous theorem holds in the critical case  $p = D/2$ . However, the LDP for  $\beta = \infty$  appears to

be rather subtle in the general setting and is only shown to hold when  $\mu_0$  is a volume form (or comparable to a volume form), except when  $p \leq 2$ , where it applies to measures  $\mu_0$  that are determining in a suitable sense.

Let us also point out that in the Euclidean setting of the Coulomb and Riesz gases in  $\mathbb{R}^n$ , with  $\mu_0$  given by the Euclidean volume form and  $\beta_N$  of the order  $N$ , a refined “microscopic” large deviation principle “at the level of processes” is obtained in [25]. Such large deviation principles are beyond the scope of the present paper and seem to require different methods—the point here is rather to allow the measure  $\mu_0$  to be very singular (and the inverse temperature to be negative, in some cases).

## 2 Proofs of the Large Deviations Results

### 2.1 General Notation

Given a compact topological space  $X$ , we will denote by  $C^0(X)$  the space of all continuous functions  $u$  on  $X$ , equipped with the sup-norm, and by  $\mathcal{M}(X)$  the space of all signed (Borel) measures on  $X$ . The subset of  $\mathcal{M}(X)$  consisting of all probability measures will be denoted by  $\mathcal{P}(X)$ . We endow  $\mathcal{M}(X)$  with the weak topology; i.e.,  $\mu_j$  is said to converge to  $\mu$  weakly in  $\mathcal{M}(X)$  if

$$\langle \mu_j, u \rangle \rightarrow \langle \mu, u \rangle := \int_X u \mu$$

for any continuous function  $u$  on  $X$ , i.e., for any  $u \in C^0(X)$  (in other words, the weak topology is the weak-star topology when  $\mathcal{M}(X)$  is identified with the topological dual of  $C^0(X)$ ). Since  $X$  is compact, so is  $\mathcal{P}(X)$ . Given a lower semi-continuous function  $F$  on

$$Y := \mathcal{P}(X),$$

we will, abusing notation slightly, also write  $F$  for the induced linear lower semi-continuous functional on  $\mathcal{P}(Y)$ :

$$F(\Gamma) := \int_{\mathcal{P}(Y)} F(\mu) \Gamma(\mu).$$

Equivalently, under the natural embedding  $\mu \mapsto \delta_\mu$  of  $Y$  into  $\mathcal{P}(Y)$ , the function  $F(\Gamma)$  is the unique lower semi-continuous affine extension of  $F$  to  $\mathcal{P}(Y)$ .

We will denote by  $S_N$  the permutation group acting on  $X^N$  and by  $\mathcal{P}(X^N)^{S_N}$  the space of symmetric measures  $\mu_N$  (i.e.,  $S_N$ -invariant) on  $X^N$ . Also note that, following standard practice, we will denote by  $C$  a generic constant whose value may change from line to line.

### 2.1.1 Entropy

We will write  $D(\nu_1, \nu_2)$  for the *relative entropy* (also called the *Kullback-Leibler divergence* in information theory) of two measures  $\nu_1$  and  $\nu_2$  on a topological space  $Z$  : if  $\nu_1$  is absolutely continuous with respect to  $\nu_2$ , i.e.,  $\nu_1 = f\nu_2$ , one defines

$$D(\nu_1, \nu_2) := \int_Y \log(\nu_1/\nu_2)\nu_1,$$

and otherwise one declares that  $D(\nu_1, \nu_2) := \infty$ . Note the sign convention used:  $D$  is minus the *physical* entropy. In our setting, the space  $Z$  will always be of the form  $X^N$ , and we will then take the reference measure  $\nu_2 = \mu_0^{\otimes N}$  and write  $D(\cdot) := D(\cdot, \mu_0^{\otimes N})$ . It will also be convenient to define the *mean entropy* of a probability measure  $\mu_N$  on  $X^N$  (i.e.,  $\mu_N \in \mathcal{P}(X^N)$ ) as

$$D^{(N)}(\mu_N) := \frac{1}{N}D(\mu_N, \mu_0^{\otimes N}).$$

Then it follows directly that

$$D^{(N)}(\mu^{\otimes N}) = D(\mu). \tag{2.1}$$

Moreover, denoting by  $(\mu_N)_j$  the  $j$ th marginal  $\mu_N$  (which defines a probability measure on  $X^j$ ),

$$D^{(N)}(\mu_N) \geq D^{(j)}((\mu_N)_j), \tag{2.2}$$

as follows from the concavity of the function  $t \mapsto \log t$  on  $\mathbb{R}_+$  (see, for example, [22]).

## 2.2 Preliminaries

### 2.2.1 Large Deviation Principles

Let us start by recalling the general definition of a large deviation principle (LDP) for a sequence of measures.

**Definition 2.1** Let  $Y$  be a Polish space, i.e., a complete separable metric space.

- (i) A function  $I : \mathcal{Y} \rightarrow ]-\infty, \infty]$  is a *rate function* if it is lower semi-continuous. It is a *good rate function* if it is also proper; i.e.,  $I^{-1}] - \infty, a]$  is compact for any given  $a \in \mathbb{R}$ .
- (ii) A sequence  $\Gamma_N$  of measures on  $Y$  satisfies a *large deviation principle* with *speed*  $r_N$  and *rate function*  $I$  if

$$\limsup_{N \rightarrow \infty} \frac{1}{r_N} \log \Gamma_N(\mathcal{F}) \leq - \inf_{\mu \in \mathcal{F}} I(\mu)$$

for any closed subset  $\mathcal{F}$  of  $Y$  and

$$\liminf_{N \rightarrow \infty} \frac{1}{r_N} \log \Gamma_N(\mathcal{G}) \geq - \inf_{\mu \in \mathcal{G}} I(\mu)$$

for any open subset  $\mathcal{G}$  of  $Y$ .

*Remark 2.2* The LDP is said to be *weak* if the upper bound is only assumed to hold when  $\mathcal{F}$  is compact. Anyway, we will only consider the case when  $Y$  is compact, and hence the notion of a weak LDP and an LDP then coincide (and moreover, any rate functional is automatically good).

**Lemma 2.3** (Bryc). *Let  $Y$  be a compact Polish space. Suppose that there exists a function  $f$  on  $C^0(Y)$  such that for any  $\Phi \in C^0(Y)$ ,*

$$f_N(\Phi) := \frac{1}{r_N} \log \int e^{r_N \Phi} \Gamma_N \rightarrow f(\Phi).$$

*Then  $\Gamma_N$  satisfies an LDP with speed  $r_N$  and rate functional*

$$I(\mu) = \sup_{\Phi \in C^0(Y)} (\Phi(\mu) - f(\Phi))$$

*(by Varadhan’s lemma the converse also holds).*

We also have the following simple lemma:

**Lemma 2.4** *The measures  $\tilde{\Gamma}_N := (\delta_N)_*(e^{-\beta H^{(N)}} \mu_0^{\otimes N})$  are finite and satisfy the asymptotics in Bryc’s lemma with rate functional  $\tilde{I}(\mu)$  and speed  $N$  if and only if the corresponding probability measures  $\Gamma_N := (\delta_N)_*(\mu_\beta^{(N)})$  on  $\mathcal{P}(X)$  satisfy an LDP at speed  $N$  with rate functional  $I := \tilde{I} - C_\beta$ , where  $C_\beta := \inf_{\mu \in \mathcal{P}(X)} \tilde{I}(\mu)$  and the sequence  $-N^{-1} \log Z_{N,\beta}$  is convergent in  $\mathbb{R}$  (and then the limit is equal to  $C_\beta$ ).*

*Proof* By definition,

$$\frac{1}{N} \log \int e^{N\Phi} \Gamma_N = \frac{1}{N} \log \int e^{N\Phi} \tilde{\Gamma}_N - \frac{1}{N} \log Z_{N,\beta}, \quad Z_{N,\beta} = \int e^{N\tilde{I}} \tilde{\Gamma}_N.$$

Hence, if the measures  $\tilde{\Gamma}_N$  satisfy the asymptotics in Bryc’s lemma with rate functional  $\tilde{I}(\mu)$ , then the measures  $\Gamma_N$  satisfy the asymptotics in Bryc’s lemma with rate functional  $\tilde{I}(\mu) - C_\beta$ , where  $C_\beta$  is the limit of  $-\frac{1}{N} \log Z_{N,\beta}$  as  $N \rightarrow \infty$ . Since  $\mu_\beta^{(N)}$  is a probability measure, the LDP for  $\Gamma_N$  implies that the inf of  $\tilde{I}(\mu) - C_\beta$  vanishes; i.e.,  $C_\beta$  is the inf of  $\tilde{I}(\mu)$ . The converse is proved in a similar way. □

### 2.2.2 Gamma-Convergence

We recall that a sequence of functions  $f_j$  on a topological space  $\mathcal{X}$  is said to *Gamma-converge* to a function  $f$  on  $\mathcal{X}$  if

$$\begin{aligned} x_j \rightarrow x \text{ in } \mathcal{X} &\implies \liminf_{j \rightarrow \infty} f_j(x_j) \geq f(x) \\ \forall x &\exists x_j \rightarrow x \text{ in } \mathcal{X} : \lim_{j \rightarrow \infty} f_j(x_j) = f(x) \end{aligned} \tag{2.3}$$

(such a sequence  $x_j$  is called a *recovery sequence*); see [14]. More generally, given a subset  $\mathcal{S} \subseteq \mathcal{X}$ , we will say that  $f_j$  *Gamma-converges to  $f$  relative to  $\mathcal{S}$*  if the existence of a recovery sequence in  $\mathcal{X}$  is only demanded when  $x \in \mathcal{S}$ .

**Lemma 2.5** *Assume that  $f_j$  Gamma-converges to  $f$  relative to  $\mathcal{S} \subset \mathcal{X}$ . Then  $f|_{\mathcal{S}}$  is lower semi-continuous.*

*Proof* Consider a sequence  $s_i \rightarrow s$  in  $\mathcal{S}$ . For each  $s_i$  we take a recovery sequence  $y_i^{(j)}$  in  $X$  converging to  $s_i$ . Setting  $y_i := y_i^{(n_i)}$ , for a suitable increasing function  $i \mapsto n_i$ , yields a sequence  $y_i$  in  $\mathcal{X}$  converging to  $s$  such that  $f(s_i) \geq f_{n_i}(y_i) - 1/i$ . Setting  $x_{n_i} := y_i$  (and  $x_j := s$  when  $j$  is not of the form  $j = n_i$  for any  $i$ ) and using the first implication in the definition of the relative Gamma-convergence of the sequence  $f_j$ , we thus deduce that  $\liminf_{i \rightarrow \infty} f(s_i) \geq f(s)$ , as desired.  $\square$

**Lemma 2.6** *Let  $\mathcal{X}$  be a compact topological space and assume that  $f_j$  Gamma-converges to  $f$  relative to a set  $\mathcal{S}$  containing all minima of  $f$ . Then*

$$\liminf_{j \rightarrow \infty} \inf_{\mathcal{X}} f_j = \inf_{\mathcal{X}} f.$$

*Proof* Given  $s \in \mathcal{S}$  we take a recovery sequence  $x_n$  and observe that

$$f(s) \geq f_n(x_n) + o(1) \geq \inf f_n + o(1) = f_n(y_n) + o(1) \geq f(y) + o(1)$$

for some  $y_n, y \in \mathcal{X}$ , by the compactness and the assumption of Gamma-convergence. In particular, when  $s$  realizes the minimum of  $f$  so does  $y$ , and hence equalities must hold above, which concludes the proof.  $\square$

### 2.2.3 Legendre–Fenchel Transforms

Let  $f$  be a function on a topological vector space  $V$ . Then its Legendre–Fenchel transform is defined as the following convex lower semi-continuous function  $f^*$  on the topological dual  $V^*$ :

$$f^*(w) := \sup_{v \in V} \langle v, w \rangle - f(v)$$

in terms of the canonical pairing between  $V$  and  $V^*$ . In the present setting, we will take  $V = C^0(X)$  and  $V^* = \mathcal{M}(X)$ , the space of all signed Borel measures on  $X$ . Then  $f^{**} = f$  for any lower semi-continuous convex function (by standard duality in locally convex topological vector spaces [17]).

### 2.3 Proof of Theorem 1.1

Set

$$E_N(x_1, \dots, x_N) := H^{(N)}(x_1, \dots, x_N)/N$$

so that the *mean energy* (1.3) can be written as

$$E^{(N)}(\mu_N) := \int_{X^N} E_N \mu_N.$$

We denote by  $F_{\beta_N}^{(N)}$  the corresponding *mean free energy functional* on  $\mathcal{P}(X^N)^{S_N}$ , at inverse temperature  $\beta_N$  :

$$F_{\beta_N}^{(N)}(\mu_N) := E^{(N)}(\mu_N) + \frac{1}{\beta_N} D^{(N)}(\mu_N), \quad (2.4)$$

assuming that  $H^{(N)} \in L^1(\mu^{(N)})$ , which ensures that  $E^{(N)}(\mu_N)$  is well defined and finite. If  $H^{(N)}$  is not in  $L^1(\mu^{(N)})$ , we set  $F_{\beta_N}^{(N)}(\mu_N) := \infty$  if  $\beta_N \in ]-\infty, \infty]$  and  $F_{\beta_N}^{(N)}(\mu_N) := -\infty$  if  $\beta_N < 0$ . Thus  $\beta_N F_{\beta_N}^{(N)}$  takes values in  $] -\infty, \infty]$ .

Now set  $Y := \mathcal{P}(X)$ . By embedding  $\mathcal{P}(X^N/S_N)$  into  $\mathcal{P}(Y)$ , using the push-ward map  $(\delta_N)_*$ , we can identify the mean free energies  $F_{\beta_N}^{(N)}$  with functionals on  $\mathcal{P}(Y)$ , extended by  $\infty$  to all of  $\mathcal{P}(Y)$ . We will identify  $Y$  with its image in  $\mathcal{P}(Y)$  under the embedding  $\mu \mapsto \delta_\mu$ .

The starting point of the proof of the LDP is the following reformulation of Bryc's lemma in terms of the Legendre–Fenchel transform, using the Gibbs variational principle:

**Lemma 2.7** (*Bryc+Gibbs*): *Suppose that the Legendre–Fenchel transforms  $f_N$  of the free energy functionals  $F_{\beta_N}^{(N)}$  converge point-wise to a function  $f$  on  $C^0(Y)$  :*

$$\lim_{N \rightarrow \infty} f_N(\Phi) = f(\Phi);$$

*i.e.*,

$$\lim_{N \rightarrow \infty} \inf_{\Gamma \in \mathcal{P}(Y)} \left( F_{\beta_N}^{(N)}(\Gamma) + \langle \Phi, \Gamma \rangle \right) = -f(-\Phi).$$

*Then the LDP holds with speed  $N\beta_N$  and rate functional*

$$I(\mu) := f^*(\delta_\mu),$$

*where  $f^*$  is the Legendre–Fenchel transform of  $f$ .*

*Proof* The Gibbs variational principle says that if  $\mu_{\beta_N}^{(N)}$  is a well-defined probability measure, then

$$\inf_{\mathcal{P}(X^N)^{S_N}} F_{\beta_N}^{(N)} = F_{\beta_N}^{(N)}(\mu_{\beta_N}^{(N)}) = -\frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N N E^{(N)}} \mu_0^{\otimes N}. \tag{2.5}$$

Indeed, rewriting  $F_{\beta_N}^{(N)}(\mu_N) = \frac{1}{\beta_N} D(\mu_N, \mu_{\beta_N}^{(N)}) - \frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N N E^{(N)}} \mu_0^{\otimes N}$ , this follows immediately from the fact that  $D \geq 0$  (which in turn follows from Jensen’s inequality). Hence, replacing  $H^{(N)}$  with the new Hamiltonian  $H^{(N)} + N\delta_N^*(\Phi)$  and applying Bryc’s lemma concludes the proof.  $\square$

*Remark 2.8* Varadhan’s lemma implies that the converse of the previous lemma also holds.

In order to verify the criterion in the previous lemma, we will use the following lemma:

**Lemma 2.9** *Under the hypotheses H1 and H2 and  $\beta \in ]0, \infty[$ , the mean free energies  $F_{\beta_N}^{(N)}$  Gamma-converge to the lower semi-continuous linear functional  $F_\beta(\Gamma)$  on  $\mathcal{P}(Y)$ , relative to  $Y$ , where  $F_\beta(\mu)$  is the macroscopic free energy on  $Y$  :*

$$F_\beta := E + D/\beta.$$

*If moreover H3 holds, then the corresponding result also holds when  $\beta = \infty$ .*

*Proof* First assume that  $\beta < \infty$ . The lower bound follows directly from hypotheses H1 and H2 together with the fact that the mean entropy functionals satisfy the lower bound in the Gamma-convergence (by subadditivity [31]; see also Theorem 5.5 in [21] for generalizations). To prove the existence of recovery sequences, we fix an element  $\Gamma$  of the form  $\delta_\mu$  and take the recovery sequence to be of the form  $(\delta_N)_* \mu^{\otimes N}$ . Then the required convergence follows from H1 together with the product property (2.1.) Finally, when  $\beta = \infty$ , the previous argument for the existence of a recovery sequence still applies as long as  $\mu$  satisfies  $D(\mu) < \infty$ . The general case then follows by a simple diagonal approximation argument using H3.  $\square$

Now, since the limiting functional  $F_\beta(\Gamma)$  is affine and lower semi-continuous (by Lemma 2.5) and the set  $Y$  is extremal in  $\mathcal{P}(Y)$ , the infimum of  $F$  on  $\mathcal{P}(Y)$  is attained in  $Y$  (for example, by Choquet’s theorem). Fixing a continuous function  $\Phi$  on  $C^0(Y)$  and replacing  $H^{(N)}$  with the new Hamiltonian  $H^{(N)} + N\delta_N^*(\Phi)$ , Lemma 2.6 thus shows that the criterion in Lemma 2.7 is satisfied. Hence the LDP holds with lower semi-continuous rate functional  $I(\mu) = f^*(\delta_\mu)$ . Finally, extending  $I$  to  $\mathcal{P}(Y)$  by linearity, this means that  $I(\Gamma)$  is the Legendre–Fenchel transform of  $f$ , i.e.,  $I = f^*$ . But in our case  $f$  is itself defined as  $f := F^*$ , and hence,  $I = F^{**} = F$  since  $F$  is convex (and even affine) and lower semi-continuous.

*Remark 2.10* An inspection of the proof of Theorem 1.1 above reveals that, in the case  $\beta = \infty$ , the hypothesis H3 may be replaced by the following weaker one:

- (H3)’ The functional  $F_\beta$  Gamma-converges towards  $E$ , as  $\beta \rightarrow \infty$ .

## 2.4 Proof of Corollary 1.2

First note that H1 is trivially satisfied. To verify the second hypothesis H2, we may as well, by linearity, assume that  $M = m$  and that there is just one term with  $W := W_m(x_1, \dots, x_m)$ . Since  $W$  is lower semi-continuous, there exists a sequence of continuous, functions  $W_R$  increasing to  $W$  as  $R \rightarrow \infty$ , and we denote by  $E_{W_R}$  the corresponding functionals on  $\mathcal{P}(X)$ . It follows readily from the definitions that for any fixed  $R > 0$ ,

$$E_{W_R}(\delta_N(x_1, \dots, x_N)) + O\left(\frac{1}{N}\right) = E_{W_R}^{(N)}(x_1, x_2, \dots, x_N),$$

and, in particular,

$$E_W^{(N)}(\mu_N) \geq \int E_{W_R}(\delta_N(x_1, \dots, x_N)) \mu_N + C_R/N.$$

But since  $E_{W_R}$  is continuous,

$$\int E_{W_R}(\delta_N(x_1, \dots, x_N)) \mu_N = \int_{\mathcal{P}(X)} E_{W_R}(\mu)(\delta_N)_* \mu_N \rightarrow \int_{\mathcal{P}(X)} E_{W_R}(\mu) \Gamma.$$

Hence,

$$\liminf_{N \rightarrow \infty} \int_{X^N} E^{(N)} \mu_N \geq \int E_{W_R}(\mu) \Gamma$$

for any  $R > 0$ . Finally, letting  $R \rightarrow \infty$  and using the monotone convergence theorem of integration theory concludes the proof.

## 2.5 Proof of Theorem 1.3

First observe that if the LDP in the second point of the theorem holds, then integrating over all of  $\mathcal{P}(X)$  reveals that the first point holds. To prove the converse, we fix  $\beta > \beta_0$  and note that the Gibbs variational principle applied at the inverse temperature  $\beta - \epsilon$  gives

$$(\beta - \epsilon) F_{\beta - \epsilon}^{(N)} \geq -C_\epsilon,$$

which we rewrite as

$$\beta F_\beta^{(N)} \geq \epsilon E^{(N)} - C_\epsilon.$$

Thus, by the Gibbs variational principle,

$$\beta F_\beta^{(N)}(\nu^{\otimes N}) \geq \beta F_\beta^{(N)}(\mu_\beta^{(N)}) \geq \epsilon E^{(N)}(\mu_\beta^{(N)}) - C_\epsilon$$



for any fixed  $\nu \in \mathcal{P}(Y)$ . In particular, taking  $\nu = \mu_0$  and using the hypothesis H1 (which implies that  $F_\beta^{(N)}(\mu_0^{\otimes N}) \rightarrow E(\mu_0)$  as  $N \rightarrow \infty$ ) gives that

$$E^{(N)}(\mu_\beta^{(N)}) \leq C'.$$

But then the previous inequalities force

$$D^{(N)}(\mu_\beta^{(N)}) \leq C''.$$

Hence, by the hypothesis H4,

$$\lim_{N \rightarrow \infty} E^{(N)}(\mu_\beta^{(N)}) = \int_{\mathcal{P}(X)} E(\mu)\Gamma(\mu)$$

for any limit point  $\Gamma$  of the laws of  $\delta_N$ . As a consequence, we deduce precisely as before that the desired asymptotics for the  $\beta_N F_{\beta_N}^{(N)}$  hold. Finally, repeating the argument with  $H^{(N)}$  replaced by the new Hamiltonian  $H^{(N)} + N\delta_N^*(\Phi)$  (which satisfies the same hypothesis) concludes the proof, just as before.

*Remark 2.11* If one only wants to prove that the laws of  $\delta_N$  concentrate on the minima of  $F_\beta$  (rather than proving an LDP), it is enough to show that the convergence of the free energies hold for  $\Phi = 0$  (as in the original approach in [28]). As revealed by the previous proof, this only requires that the hypothesis H4 holds for the particular sequence  $\mu_\beta^{(N)}$ .

### 2.6 Proof of Corollary 1.4

First observe that the assumption on  $Z_{N,\beta}$  in the case  $N = 2$  gives the following exponential integrability property of  $W$ :

$$\int_{X^2} e^{-\beta W} \mu_0^{\otimes 2} < \infty \tag{2.6}$$

for any  $\beta$ . Hence, by Theorem 1.3, it will be enough to show that the previous integrability property implies that H4 is satisfied (for any sequence  $\mu_N$ ). To this end, we will apply a duality argument. First recall that given a measure space  $(\mathcal{X}, \mu)$  and a finite Young function  $\theta$  on  $\mathbb{R}$  (i.e., a non-negative even lower semi-continuous convex function such that  $\theta(0) = 0$ ), the corresponding *large Orlicz space* is defined by

$$L_\theta(\mathcal{X}, \mu) := \left\{ f : \exists \alpha > 0 : \int \theta(\alpha f) \mu < \infty \right\}$$

and the corresponding *small Orlicz space* is defined by

$$M_\theta(\mathcal{X}, \mu) := \left\{ f : \forall \alpha > 0 : \int \theta(\alpha f) \mu < \infty \right\}$$

(where all functions  $f$  are assumed measurable). The space  $L_\theta(\mathcal{X}, \mu)$  may be equipped with a norm  $\|\cdot\|_\theta$ , called the Luxemburg norm, which turns  $L_\theta(\mathcal{X}, \mu)$  and its subspace  $M_\theta(\mathcal{X}, \mu)$  into Banach spaces:

$$\|f\|_\theta := \inf \left\{ b > 0 : \int \theta(b^{-1}f)\mu \leq 1 \right\},$$

i.e., the gauge of the set (unit-ball)

$$\left\{ f : \int \theta(f)\mu \leq 1 \right\}.$$

By the Hölder–Young inequality,

$$\left| \int fg\mu \right| \leq 2 \|f\|_\theta \|g\|_{\theta^*},$$

where  $\theta^*$  is the Young function defined as the Legendre–Fenchel transform of  $\theta$ . In particular, for any  $g \in L_{\theta^*}$   $f \mapsto \int fg\mu$  defines a continuous function on  $L_\theta$  with bounded operator norm; i.e.,  $L_{\theta^*} \subset L_\theta^*$ , where  $L^*$  denotes the Banach space dual of a Banach space  $L$ , endowed with the operator norm. To apply this in the present context, we note that

$$E^{(N)}(\mu_N) = \int_{X^2} W\rho_N\mu_0^{\otimes 2},$$

where  $\rho_N$  is the density of the second marginal of  $\mu_N$ . The assumption that  $D^{(N)}(\mu^{(N)}) \leq C$  implies that

$$\int_{X^2} (\rho_N \log \rho_N)\mu_0^{\otimes 2} \leq C2,$$

according to (2.2).

Now set  $\theta(s) := e^s - s - 1$ , when  $s \geq 0$ . Then  $\theta^*(t) = (t + 1) \log(1 + t) - t$  when  $t \geq 0$ . By the previous entropy inequality for  $\rho_N$ , the sequence  $\{\rho_N\}$  stays in a fixed ball in  $L_{\theta^*}$ , and hence, by the Hölder-Young inequality,  $\{\rho_N\}$  stays in a fixed ball in the dual Banach space  $L_\theta^*$ . By weak compactness, it then follows that there exists  $\Lambda \in L_\theta^*$  such that for any  $g \in L_\theta$ ,

$$\int_{X^2} \rho_N g \mu_0^{\otimes 2} \rightarrow \langle \Lambda, g \rangle$$

(after perhaps passing to a subsequence). Now, since  $\rho_N\mu_0^{\otimes 2}$  is a probability measure, we may also assume that there exists  $\rho \in L_{\theta^*}$  such that

$$\int_{X^2} \rho_N u \mu_0^{\otimes 2} \rightarrow \int_{X^2} \rho u \mu_0^{\otimes 2}$$

for any continuous function  $u$ . In our case  $g = W$ , and we just need to check that  $\langle \Lambda, g \rangle = \langle \rho \mu_0^{\otimes 2}, g \rangle$ . But, by assumption,  $W \in M_\theta$ , and by the general duality theorems in [26,30] the topological dual of  $M_\theta$  identifies with  $L_{\theta^*}$ , i.e. any continuous functional  $\Lambda$  on  $M_\theta$  is obtained by integrating against a (unique)  $\rho \in L_{\theta^*}$ , which concludes the proof.

**2.7 Proof of Theorem 1.5**

Given a compact metric space  $X$ , we endow  $Y (:= \mathcal{P}(X))$  with the Wasserstein  $L^1$ -metric  $d$ , which is compatible with the weak topology:

$$d(\mu, \nu) = \sup_{f: L(f) \leq 1} \int f(\mu - \nu),$$

where  $f$  is Lipschitz continuous on  $X$  with Lipschitz constant  $L(f) \leq 1$ . Since  $\int(\mu - \nu) = 0$ , we may as well assume that  $f(x_0) = 0$  for a fixed point  $x_0$  and hence that  $|f(x)| \leq C_X$ , where  $C_X$  is independent of  $f$  (since  $X$  is compact and, in particular, has bounded diameter).

Let us first show that, when  $\beta > \beta_0$ ,

$$Z_{N,\beta} \leq A_\beta^{(N-1)}, \quad A_\beta = \sup_{x \in X} \int_X e^{-\beta W(x,y)} \mu_0(y). \tag{2.7}$$

To see this, rewrite  $-\beta H^{(N)} = \frac{1}{N} \sum_{i=1}^N f_i$ , where  $f_i$  is the sum of  $\beta W(x_i, x_j)$  over all  $j$  such that  $j \neq i$ . The arithmetic-geometric means inequality gives

$$\int_{X^N} e^{-\beta H^{(N)}} \mu_0^{\otimes N} \leq \sum_{i=1}^N \frac{1}{N} \int_{X^N} e^{f_i} \mu_0^{\otimes N}.$$

Now, for a given  $i$ , we have (by first integrating over the  $N - 1$  variables different from  $x_i$  and then taking the sup over  $x_i$ )

$$\int_{X^N} e^{f_i} \mu_0^{\otimes N} = \int_X \left( \int_X e^{-W(x,y)} \mu_0(y) \right)^{N-1} \mu_0(x) \leq A_\beta^{N-1},$$

where, by assumption,  $A_\beta < \infty$  (since  $A_{\beta_0} < \infty$  and  $W$  is lsc on the compact space  $X^2$  and thus bounded from below).

Next we fix a continuous function  $\Phi$  on  $Y := \mathcal{P}(X)$ . Without loss of generality, we may as well assume that  $W, \Phi \geq 0$ .

First observe that when  $\beta > \beta_0$ , we have

$$Z_{N,\beta}[\Phi] := \int_{X^N} e^{-\beta(H^{(N)} + N\delta_N^*(\Phi))} \mu_0^{\otimes N} \leq C_\beta^N, \tag{2.8}$$

as follows directly from the bound (2.7) (using that  $\Phi$  is bounded).

Using the convergence of the mean energies and Gibbs variational principle, as before, we thus have

$$-\log C_\beta \leq \limsup_{N \rightarrow \infty} -\frac{1}{N} \log Z_{N,\beta}[\Phi] \leq \inf_{\mu \in \mathcal{P}(X)} (\beta E(\mu) + \Phi(\mu) + D(\mu)). \tag{2.9}$$

Now, to prove the LDP we need, in view of Lemma 2.7, to complement the upper bound on  $-\frac{1}{N} \log Z_{N,\beta}[\Phi]$  in formula (2.9) with a corresponding lower bound. To this end it would, by Theorem 1.3, be enough to establish the hypothesis H4 in the present context (for example by trying to extend the Orlitz space duality argument in the proof of Corollary 1.4). Such an approach remains to be developed (however, a macroscopic version of H4 does hold; see formula (2.12)). Here we will instead take another road, exploiting the stronger  $L^p$ -bounds provided by the following lemma (inspired by [15,22]):

**Lemma 2.12** *Let  $\Phi$  be a given Lipschitz continuous function on  $Y := \mathcal{P}(X)$  and fix  $\beta > \beta_0$ . Then the following estimate holds for the densities  $\rho_j^{(N)}$  of the  $j$ th marginal of the Gibbs measures corresponding to the Hamiltonian  $H^{(N)} + N\delta_N^*(\Phi)$  :*

$$\rho_j^{(N)}(x_1, \dots, x_j) \leq C_j e^{-\frac{1}{N} \sum_{k \neq l \leq j} W(x_k, x_l)}$$

as  $N \rightarrow \infty$ . In particular, for any  $p > 1$ ,  $\rho_j^{(N)}(x_1, \dots, x_j)$  is uniformly bounded in  $L^p$  as  $N \rightarrow \infty$ .

*Proof* To fix ideas, we start with the case  $\Phi = 0$ , following closely the proof of Theorem 3.1 in [15]. Set

$$W(X, Y) := \sum_{x \in X, y \in Y} W(x_i, y_j), \quad d\mu(Y) := \mu_0^{\otimes N-j},$$

where  $X := \{x_1, \dots, x_j\}$  and  $Y := \{x_{j+1}, \dots, x_N\}$ . Then we can decompose  $E^{(N)}(x_1, \dots, x_N) = \frac{1}{N} W(X, X) + \frac{1}{N} W(X, Y) + \frac{1}{N} W(Y, Y)$ . Accordingly,

$$\rho_j^{(N)}(X) = e^{-\frac{\beta}{N} W(X,X)} \frac{1}{Z_N} \int e^{-\frac{\beta}{N} W(X,Y)} e^{-\frac{\beta}{N} W(Y,Y)} d\mu(Y).$$

Applying Hölder’s inequality with  $p = N/j$  (and thus  $q = 1 + j/(N - j)$ ), the integral in the right-hand side is bounded from above by

$$= \left( \int e^{-\beta W(X,Y)} d\mu(Y) \right)^{1/N} \left( \int e^{-\frac{\beta}{N} q W(Y,Y)} d\mu(Y) \right)^{1/q}.$$

By assumption, the integral appearing in the first factor above is bounded from above by a  $A^N$  for some positive constant  $A$ . It will thus be enough to show that the second integral is controlled by  $Z_N$  in the sense that it is bounded from above by a uniform constant times  $Z_N$ . To this end, we will apply Hölder’s inequality again, now with

conjugate exponents  $u$  and  $w$  with  $u$  sufficiently close to 1 (to be quantified below). We thus rewrite

$$q = \frac{1}{u} + \left( q - \frac{1}{u} \right)$$

and apply Hölder’s inequality. Since  $w(q - 1/u) = 1 + w(q - 1) = 1 + \frac{wj}{(N-j)}$ , this gives

$$\int e^{-\frac{\beta}{N}qW(Y,Y)}d\mu(Y) \leq \left( \int e^{-\frac{\beta}{N}W(Y,Y)}d\mu(Y) \right)^{1/u} \left( \int e^{-\frac{\beta}{N}(1+\frac{wj}{(N-j)})W(Y,Y)}d\mu(Y) \right)^{1/w}. \tag{2.10}$$

The first factor is controlled by  $Z_N$  (since  $W \geq 0$ ). Moreover, taking  $w = \epsilon(N - j)/j$  for a sufficiently small positive number  $\epsilon$  ensures that the integral in the second factor is controlled by  $Z_{N,(1+\epsilon)\beta} \leq B^N$  (by (2.8)). Since  $w$  is of the order  $N$ , this concludes the proof when  $\Phi = 0$ . To treat the general case, we will use the following:

Claim:  $\left| \Phi \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) - \Phi \left( \frac{1}{N-j} \sum_{i=j+1}^N \delta_{x_i} \right) \right| \leq C \frac{1}{N}.$

Accepting the claim for the moment and introducing the notation  $\Phi(\frac{1}{N-j} \sum_{i=j+1}^N \delta_{x_i}) = \phi(Y)$ , we have

$$\rho_j^{(N)}(X) \leq \frac{e^C}{Z_N} e^{-\frac{\beta}{N}W(X,X)} \int e^{-\frac{\beta}{N}W(X,Y)} e^{-\frac{\beta}{N}(W(Y,Y)+N^2\phi(Y))} d\mu(Y).$$

We then use first Hölder’s inequality with  $p$  and  $q$  and then with  $u$  and  $v$  exactly as above to get the same factors as above apart from the last factor in formula (2.10), which now becomes

$$\int e^{-\frac{\beta}{N}(1+\gamma)(W(Y,Y)+N^2\phi(Y))} d\mu(Y),$$

which is bounded from above by  $C'^N$ , according to the estimate (2.8) (when  $\gamma$  is sufficiently small). This proves the lemma once we have verified the claim above. To this end, we assume to simplify the notation that  $j = 1$  (the general case is similar) and observe that setting  $\mu := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$  and  $\nu := \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}$  gives

$$N(\mu - \nu) = \delta_{x_1} - \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}.$$

Hence, for any  $f$  such that  $|f| \leq C_X$ , we have

$$\left| \int f(\mu - \nu) \right| \leq \frac{1}{N} \left( C_X + \frac{1}{N-1} \sum_{i=2}^N C_X \right) \leq \frac{1}{N} 2C_X,$$

and hence  $d(\mu, \nu) \leq 2C_X/N$ . But then the claim follows directly from the Lipschitz continuity of  $\Phi$ . □

Now, to verify the missing lower bound on  $-\frac{1}{N} \log Z_{N,\beta}[\Phi]$ , we first claim that it will be enough to verify the case when  $\Phi$  is Lipschitz continuous. Indeed, any continuous function  $\Phi$  on a compact metric space  $Y$  can be written as a uniform limit  $\Phi^{(R)}$  of Lipschitz continuous function (for example,  $\Phi^{(R)}(x) := \inf_Y (\Phi(y) + R d(x, y))$ ) increases to  $\Phi$ , as  $R \rightarrow \infty$ , and has Lipschitz constant  $R$ . Moreover we may, after relabeling the sequence, assume that  $|\Phi_\epsilon - \Phi| \leq \epsilon$ . But since  $\Phi \mapsto -\frac{1}{N\beta} \log Z_{N,\beta}[\Phi]$  is increasing and  $-\frac{1}{N\beta} \log Z_{N,\beta}[\Phi + c] = -\frac{1}{N\beta} \log Z_{N,\beta}[\Phi] + c$  for any  $c \in \mathbb{R}$ , we get

$$\left| \frac{1}{N\beta} \log Z_{N,\beta}[\Phi] - \frac{1}{N\beta} \log Z_{N,\beta}[\Phi_\epsilon] \right| \leq \epsilon,$$

which proves the claim. Next, we recall that, by the Gibbs variational principle,

$$-\frac{1}{N} \log Z_{N,\beta}[\Phi] = F_\beta^{(N)}(\mu_\beta^{(N)}),$$

where  $\mu_\beta^{(N)}$  denotes the sequence of Gibbs measures, at inverse temperature  $\beta$ , corresponding to the Hamiltonian  $H^{(N)} + N\delta_N^*(\Phi)$ , for  $\Phi$  Lipschitz continuous and decompose

$$E^{(N)}(\mu_\beta^{(N)}) = \int_{X^2} W\rho_2^{(N)} \mu_0^{\otimes 2} + \langle \Phi, \Gamma_N \rangle.$$

By continuity the second term above converges towards  $\langle \Phi, \Gamma \rangle$ . To prove the desired lower bound on  $F_\beta^{(N)}(\mu_\beta^{(N)})$ , it will thus, just as in the proof of Corollary 1.4, be enough to show that

$$\lim_{N \rightarrow \infty} \int_{X^2} W\rho_2^{(N)} \mu_0^{\otimes 2} = \int W\rho_2 \mu_0^{\otimes 2} \tag{2.11}$$

for any weak limit point  $\rho_2 \mu_0^{\otimes 2}$  of  $\rho_2^{(N)} \mu_0^{\otimes 2}$ . To this end, we recall that, by the previous lemma,  $\rho_2^{(N)}$  is uniformly bounded in  $L^p$ , as  $N \rightarrow \infty$ , for any fixed  $p > 1$ . Hence, by standard  $L^p$ -duality, the limit (2.11) follows from the fact that  $W \in L^q$  for some (any)  $q > 1$ , since by assumption,  $e^{\epsilon W} \in L^1$  for any sufficiently small positive number  $\epsilon$ . This concludes the proof of Theorem 1.5.

2.7.1 An Alternative Direct Proof of the Lower Semi-continuity of  $\beta F_\beta$

A consequence of the LDP established above is that the corresponding (scaled) free energy functional  $\beta F_\beta$  is lsc on  $\mathcal{P}(X)$ . As we show next, this could also be shown directly by establishing a macroscopic version of the hypothesis H4 (using Orlicz space duality). This indicates that there could be a more direct proof of the LDP that avoids Lemma 2.12, as discussed above.

**Lemma 2.13** *Under the assumptions of Theorem 1.5, the following holds: for any  $\beta > \beta_0$  the (scaled) free energy functional  $\beta F_\beta$  on  $\mathcal{P}(X)$  is lower semi-continuous.*

*Proof* First observe that by the inequality (2.9) (applied to  $\Phi = 0$ ) :

$$\beta F_\beta \geq -C_\beta.$$

Now, applying the previous bound to  $\beta - \epsilon > \beta_0$  reveals that

$$\beta F_\beta \geq \epsilon E - C'.$$

Given  $\mu$  in  $\mathcal{P}(X)$  with  $E(\mu) < \infty$ , we set

$$u_\mu(x) := \int W(x, y)\mu(y).$$

Then

$$E(\mu) = \int_X u_\mu \mu.$$

Using the previous estimate it will, to prove the lemma, be enough to verify the following “macroscopic” version of H4 for any sequence  $\mu_j$  converging weakly towards  $\mu$  :

$$\text{“MacroH4”} : D(\mu_j) \leq C \implies E(\mu_j) \rightarrow E(\mu) \tag{2.12}$$

(compare [9, Theorem 2.17]). To this end, set

$$u_j := u_{\mu_j} := \int W(x, y)\mu_j(y).$$

First observe that it will be enough to prove that

$$\|u_j - u\|_\theta \rightarrow 0, \tag{2.13}$$

where  $\theta$  is of exponential type, as in the proof of Corollary 1.4). Indeed,  $E(\mu_j) - E(\mu) = \langle u_{\mu_j}, \mu_j \rangle - \langle u_\mu, \mu \rangle = \langle u_{\mu_j} - u_\mu, \mu_j \rangle - \langle u_\mu, \mu - \mu_j \rangle = \langle u_{\mu_j} - u_\mu, \mu_j \rangle - \langle \mu, u_\mu - u_{\mu_j} \rangle$ , using that  $W(x, y)$  is symmetric in the last step. Hence, the Hölder–Young inequality gives

$$|E(\mu_j) - E(\mu)| \leq 2 \|u_j - u\|_\theta (\|\rho_j\|_{\theta^*} + \|\rho\|_{\theta^*}).$$

Since the second factor in the right-hand side above is uniformly bounded (since  $D(\mu_j) \leq C$ ), this shows that it is indeed enough to prove (2.13). To prove the latter convergence, observe that the Hölder–Young inequality implies that

$$|u_j| \leq C, \quad |u| \leq C$$

(using the integrability assumption on  $W$ ). Indeed, if  $\mu = \rho\mu_0$ , then

$$|u_\mu(x)| = \left| \int W(x, y)\rho(y)\mu_0(y) \right| \leq 2 \|W(x, \cdot)\|_\theta \|\rho\|_{\theta^*},$$

where the first term is uniformly bounded in  $x$ , by assumption, and the second term is controlled by  $D(\mu)$ . Next, note that the convergence (2.13) that we intend to prove is equivalent to proving

$$\int_X \theta(a(u_j - u))\mu_0 \rightarrow 0$$

for any fixed positive number  $a$ . Now, since  $\theta(t) \leq |t|e^{|t|}$  and the sup norms of  $|u_j|$  and  $|u|$  are bounded from above by  $C$ , we have

$$\int_X \theta(a(u_j - u))\mu_0 \leq a \int_X |u_j - u|e^{a|u_j - u|}\mu_0 \leq ae^{2aC} \int_X |u_j - u|\mu_0.$$

All that remains is thus to show that

$$\|u_j - u\|_{L^1(X, \mu_0)} \rightarrow 0. \tag{2.14}$$

By the lower semi-continuity of  $W$ ,

$$\liminf_{j \rightarrow \infty} \int u_j \mu_0 \geq \int u \mu_0,$$

the desired convergence (2.14) will follow from general measure theory if  $\int u_j \mu_0 \rightarrow \int u \mu_0$  (using that  $u_\mu \geq 0$  if  $W$  is normalized so that  $W \geq 0$ ). But

$$\int_X u_j \mu_0 = \int_X v(y)\mu_j(y), \quad v := \int_X W(x, y)\mu_0(y),$$

where  $v$  is bounded, by the previous argument (since  $\mu_0$  trivially has finite entropy). In particular,  $v$  is in the little Orlicz space  $M_\theta(X, \mu_0)$ , and since  $D(\mu_j) \leq C$ , the desired convergence (2.14) then follows from the duality argument towards the end of the proof of Corollary 1.4. □



### 2.8 Relations to Gamma-Convergence of $E^{(N)}$ on $\mathcal{P}(X)$ : Proof of Corollary 1.6

First observe that the required lower bound on  $E^{(N)}(x^{(N)})$  is obtained by taking  $\mu_N \in \mathcal{P}(X^N)^{S_N}$  to be the normalized  $S_N$ -orbit in  $X^N$  of the Dirac measure supported at a given  $x^{(N)} \in X^N$ :

$$\mu_N = \frac{1}{N!} \sum_{\sigma \in S_N} \delta_{\sigma(x^{(N)})}.$$

Then  $E^{(N)}(x^{(N)}) = E^{(N)}(\mu_N)$ . Now, if  $\delta_N(x^{(N)})$  converges towards  $\mu$  in  $\mathcal{P}(X)$ , then  $(\delta_N)_* \mu_N$  converges towards  $\Gamma := \delta_\mu$ , and hence the desired lower bound on  $E^{(N)}(x^{(N)})$  follows from the assumed lower bound on  $E^{(N)}(\mu_N)$ .

Next, to construct recovery sequences, we fix some  $\beta > 0$ , say  $\beta = 1$ , and a probability measure  $\mu_0$ . By Theorem 1.1 an LDP holds with rate functional  $E(\mu) + D_{\mu_0}(\mu)$ . In particular, the lower bound in the LDP gives that, for any given  $\mu$  and  $\mu_0$  in  $\mathcal{P}(X)$ ,

$$\begin{aligned} -E(\mu) - D_{\mu_0}(\mu) &\leq \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{B_\epsilon(\mu)} e^{-H^{(N)}} \mu_0^{\otimes N} \leq \\ &\leq \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( - \inf_{B_\epsilon(\mu)} E^{(N)} + 0 \right). \end{aligned}$$

Hence, taking a sequence  $x_\epsilon^{(N)} \in B_\epsilon(\mu)$  such that  $E^{(N)}(x_\epsilon^{(N)}) \leq \inf_{B_\epsilon(\mu)} E^{(N)} + \epsilon$  gives

$$-E(\mu) - D_{\mu_0}(\mu) \leq \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \left( -E^{(N)}(x_\epsilon^{(N)}) \right).$$

In particular, taking  $\mu_0 = \mu$  (which gives  $D_{\mu_0}(\mu) = 0$ ), we deduce, by a diagonal argument, the existence of a recovery sequence for any given  $\mu \in \mathcal{P}(X)$ .

## 3 Concluding Remarks

### 3.1 A Weaker form of the Hypothesis H2

Let us come back to the setting of Theorem 1.1 and observe that the hypothesis H2 may be replaced by the following one, which is a priori weaker (see the beginning of Sect. 2.8):

- (H2') For any sequence of  $x^{(N)} \in X^N$  such that  $\delta_N(x^{(N)}) \rightarrow \mu$  weakly in  $\mathcal{P}(X)$ , we have

$$\liminf_{N \rightarrow \infty} \frac{1}{N} H^{(N)}(x^{(N)}) \geq E(\mu).$$

In other words, (H2') says that  $E^{(N)} := H^{(N)}/N$ , when viewed as a functional on  $\mathcal{P}(X)$ , satisfies the lower bound property that is one of the two requirements for the Gamma-convergence of  $E^{(N)}$  towards  $E(\mu)$ , where  $E(\mu)$  denotes, as before, the macroscopic mean energy whose existence is postulated in hypotheses H1.

**Theorem 3.1** *The conclusion of Theorem 1.1 remains valid if H2 is replaced by H2'.*

*Proof* Just as in the proof of Theorem 1.1, in order to verify the convergence in Bryc's lemma, we may without loss of generality assume that  $\Phi = 0$ . Moreover, exactly as before, H1 combined with the Gibbs variational principle yields the upper bound on  $-\log Z_{N,\beta_N}$ . Thus we just have to prove the following bound:

$$\limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} \leq - \inf_{\mathcal{P}(X)} F_\beta. \tag{3.1}$$

To this end, for any fixed  $\delta > 0$ , we cover the compact space  $\mathcal{P}(X)$  by a finite number  $M_\delta$  of balls of radius  $\delta$ . Then

$$\limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} \leq 0 + \limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{B_\delta(\mu_{\delta,N})} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N}, \tag{3.2}$$

where  $\mu_{\delta,N}$  is the center of the ball with the largest integral (using that  $(N\beta_N)^{-1} \log M_\delta \rightarrow 0$ ). Denote by  $\mu_\delta$  a weak limit point in  $\mathcal{P}(X)$  of the family  $\mu_{\delta,N}$  as  $N \rightarrow \infty$ . Then  $B_\delta(\mu_{\delta,N}) \subset B_{2\delta}(\mu_\delta)$  as  $N \rightarrow \infty$  (along the subsequence of  $\{N\}$ ). Next, denote by  $\mu$  a weak limit point of  $\mu_\delta$  as  $\delta \rightarrow 0$ , i.e.,  $\mu$  is the limit of  $\mu_{\delta_j}$  as  $\delta_j \rightarrow 0$ , and fix  $\epsilon > 0$ . For any sufficiently small  $\delta_j$ , we have  $B_{2\delta_j}(\mu_{\delta_j}) \subset B_\epsilon(\mu)$ . Hence, for any such  $\delta_j$ , we have

$$B_{\delta_j}(\mu_{\delta_j,N}) \subset B_\epsilon(\mu)$$

as  $N \rightarrow \infty$  (along the subsequence of  $\{N\}$ ). Thus

$$\limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} \leq \limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{B_\epsilon(\mu)} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N}. \tag{3.3}$$

Moreover, by Sanov's theorem,

$$\limsup_{\epsilon \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{B_\epsilon(\mu)} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} \leq - \liminf_{\epsilon \rightarrow \infty} \liminf_{N \rightarrow \infty} \inf_{B_\epsilon(\mu)} E^{(N)} - \frac{1}{\beta} D(\mu), \tag{3.4}$$

where, by hypothesis H2', the right-hand side is bounded from above by  $-E(\mu) - \frac{1}{\beta} D(\mu) := -F_\beta(\mu)$ . Hence,

$$\limsup_{N \rightarrow \infty} \frac{1}{N\beta_N} \log \int_{X^N} e^{-\beta_N H^{(N)}} \mu_0^{\otimes N} \leq -F_\beta(\mu) \leq - \inf_{\mathcal{P}(X)} F_\beta,$$

which concludes the proof of the bound (3.1) (strictly speaking we have proved the bound for some subsequence of the  $\{N\}$ , but this is enough since we could have started

by replacing the sequence  $\{N\}$  with a subsequence  $N_j$  with the property that the limsup in formula (3.1) is a lim along the subsequence  $N_j$ .  $\square$

A result essentially equivalent to the previous theorem appears in [13].

Combining Theorem 3.1 and Corollary 1.6 thus reveals that the hypotheses H1 and H2' actually imply the Gamma-convergence of  $\frac{1}{N}H^{(N)}$  towards  $E$  on  $\mathcal{P}(X)$ . But it seems unlikely that, in general, the assumption that  $\frac{1}{N}H^{(N)}$  Gamma-converges towards a functional  $E$  on  $\mathcal{P}(X)$  is sufficient to deduce an LDP (even if one also assumes H3). On the other hand, as shown in [5], one does get an LDP for any  $\beta \in ]0, \infty]$  under an assumption of quasi-superharmonicity:

**Theorem 3.2** [5] *Let  $H^{(N)}$  be a sequence of lower semi-continuous symmetric functions on  $X^N$ , where  $X$  is a compact Riemannian manifold. Assume that :*

- The sequence  $\frac{1}{N}H^{(N)}$  on  $X^N$  (identified with a sequence of functions on  $\mathcal{P}(X)$ ) Gamma-converges towards a functional  $E$  on  $\mathcal{P}(X)$ .
- $H^{(N)}$  is uniformly quasi-superharmonic, i.e.,  $\Delta_{x_1}H^{(N)}(x_1, x_2, \dots, x_N) \leq C$  on  $X^N$ .

Then, for any sequence of positive numbers  $\beta_N \rightarrow \beta \in ]0, \infty]$ , the measures  $\Gamma_N := (\delta_N)_* e^{-\beta_N H^{(N)}}$  on  $\mathcal{M}_1(X)$  satisfy, as  $N \rightarrow \infty$ , an LDP with speed  $\beta_N N$  and good rate functional

$$F_\beta(\mu) = E(\mu) + \frac{1}{\beta} D_{dV}(\mu). \tag{3.5}$$

This is not hard to see when  $\beta = \infty$ , but for  $\beta < \infty$ , the proof hinges on a submean inequality for quasi-subharmonic functions with a distortion factor that is subexponential in the dimension, proved in [5].

**Acknowledgements** This work was supported by grants from the ERC and the KAW foundation. It is a pleasure to thank Sebastien Boucksom, Vincent Guedj, Philippe Eyssidieu, and Ahmed Zeriahi for the stimulating collaboration [9] and the referee for careful reading of the manuscript and many helpful remarks. Also thanks to the special issue editors Peter Forrester, Doug Hardin, and Sylvia Serfaty for the invitation to contribute to the special issue of *Constructive Approximation*.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Ben Arous, G., Guionnet, A.: Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Rel. Fields* **108**(4), 517–542 (1997)
2. Ben Arous, G., Zeitouni, O.: Large deviations from the circular law. *ESAIM Probab. Stat.* **2**, 123–134 (1998)
3. Berman, R.J.: Determinantal point processes and fermions on complex manifolds: large deviations and Bosonization. *Comm. Math. Phys.* **327**(1), 1–47 (2014)
4. Berman, R.J.: Kähler–Einstein metrics, canonical random point processes and birational geometry. In: *AMS Proceedings of the 2015 Summer Research Institute on Algebraic Geometry (to appear)*. [arXiv:1307.3634](https://arxiv.org/abs/1307.3634)

5. Berman, R.J.: Large deviations for Gibbs measures with singular Hamiltonians and emergence of Kähler–Einstein metrics. *Comm. Math. Phys.* **354**(3), 1133–1172 (2017)
6. Berman, R.J.: Large Deviations for Gibbs Measures and Global Potential Theory: Riemannian Versus Kähler Manifolds (**in preparation**)
7. Berman, R.J., Boucksom, S.: Growth of balls of holomorphic sections and energy at equilibrium. *Invent. Math.* **181**(2), 337 (2010)
8. Berman, R.J., Boucksom, S., Witt Nyström, D.: Fekete points and convergence towards equilibrium measures on complex manifolds. *Acta Math.* **207**(1), 1–27 (2011)
9. Berman, R.J., Boucksom, S., Eyssidieu, P., Guedj, V., Zeriahi, A.: Kähler–Einstein metrics and the Kähler–Ricci flow on log Fano varieties. *Crelle’s J.* (**to appear**). [arXiv:1111.7158](https://arxiv.org/abs/1111.7158)
10. Berman, R.J., Onnheim, M.: Propagation of Chaos for a Class of First Order Models with Singular Mean Field Interactions. [arXiv:1610.04327](https://arxiv.org/abs/1610.04327)
11. Bloom, T., Levenberg, N., Piazzon, P., Wielonsky, F.: Bernstein–Markov: A Survey. *Dolomites Res. Notes Approx.* Vol. (Special Issue) 75–91 (2015). [arXiv:1512.00739](https://arxiv.org/abs/1512.00739)
12. Bodineau, T., Guionnet, A.: About the stationary states of vortex systems. *Ann. Inst. Henri Poincaré Probab. Stat.* **35**, 205–237 (1999)
13. Boucksom, S.: Limite thermodynamique et théorie du potentiel. *SMF Gazette* Octobre. No. 146 (2015)
14. Braides, A.: *Gamma-Convergence for Beginners*. Oxford University Press, Oxford (2002)
15. Caglioti, E., Lions, P.-L., Marchioro, C., Pulvirenti, M.: A special class of stationary flows for two-dimensional Euler equations: a statistical mechanics description. *Commun. Math. Phys.* **143**(3), 501–525 (1992)
16. Chafaï, D., Gozlan, N., Zitt, P.-A.: First-order global asymptotics for confined particles with singular pair repulsion. *Ann. Appl. Probab.* **24**(6), 2371–2413 (2014)
17. Dembo, A., Zeitouni, O.: Large deviation techniques and applications. Corrected reprint of the second (1998) edition. In: *Stochastic Modelling and Applied Probability*, 38. Springer, Berlin (2010)
18. Dupuis, P., Laschos, V., Ramanan, K.: Large Deviations for Empirical Measures Generated by Gibbs Measures with Singular Energy Functionals. [arXiv:1511.06928](https://arxiv.org/abs/1511.06928)
19. García Zelada, D.: A Large Deviation Principle for Empirical Measures on Polish Spaces: Application to Singular Gibbs Measures on Manifolds. [arXiv:1703.02680](https://arxiv.org/abs/1703.02680)
20. Hardin, D.P., Saff, E.B.: Discretizing manifolds via minimum energy points. *Not. Am. Math. Soc.* **51**(10), 1186–1194 (2004)
21. Hauray, M., Mischler, S.: On Kac’s chaos and related problems. *J. Funct. Anal.* (2014). [arXiv:1205.4518](https://arxiv.org/abs/1205.4518)
22. Kiessling, M.K.H.: Statistical mechanics of classical particles with logarithmic interactions. *Comm. Pure Appl. Math.* **46**, 27–56 (1993)
23. Kiessling, M.K.H.: Statistical mechanics approach to some problems in conformal geometry. *Phys. A: Stat. Mech. Appl.* **279**(1–4), 353–368 (2000)
24. Kiessling, Michael K.-H., Spohn, H.: A note on the eigenvalue density of random matrices. *Comm. Math. Phys.* **199**(3), 683–695 (1999)
25. Leblé, T., Serfaty, S.: Large Deviation Principle for Empirical Fields of Log and Riesz Gases. [arXiv:1502.02970](https://arxiv.org/abs/1502.02970)
26. Leonard, C.: Orlicz Spaces. <http://leonard.perso.math.cnrs.fr/papers/Leonard-Orlicz%20spaces.pdf>
27. Mariani, M.: A Gamma-Convergence Approach to Large Deviations. [arXiv:1204.0640](https://arxiv.org/abs/1204.0640)
28. Messer, J., Spohn, H.: Statistical mechanics of the isothermal Lane–Emden equation. *J. Stat. Phys.* **29**(3), 561–578 (1982)
29. Onsager: Statistical hydrodynamics. *Supplemento al Nuovo Cimento* 6:279–287 (1949)
30. Rao, M.M., Ren, Z.D.: *Theory of Orlicz Spaces*, Volume 146 of Pure and Applied Mathematics. Marcel Dekker, New York (1991)
31. Robinson, D.W., Ruelle, D.: Mean entropy of states in classical statistical mechanics. *Comm. Math. Phys.* **5**, 288–300 (1967)
32. Saff, E.B., Kuijlaars, A.B.J.: Distributing many points on a sphere. *Math. Intell.* **19**(1), 5–11 (1997)
33. Saff, E., Totik, V.: *Logarithmic Potentials with Exterior Fields*. Springer, Berlin (1997) (with an appendix by Bloom, T)
34. Serfaty, S.: Coulomb gases and Ginzburg–Landau vortices. *Zurich Lectures in Advanced Mathematics*. European Mathematical Society (EMS), Zürich (2015)