A Two–Scale Proof of a Logarithmic Sobolev Inequality

Natalie Grunewald[*], Felix Otto[†], Maria G. Reznikoff[‡], and Cédric Villani[§]

July 25, 2005

Abstract

We consider an N–site lattice system with continuous spin variables governed by a Ginzburg–Landau–type potential. Because we are interested in the Kawasaki dynamics, we work with the canonical ensemble in which the mean m is given. We prove a logarithmic Sobolev inequality (LSI) which is uniform in m and has the optimal scaling in the system size N. The method involves a two–scale "block–spin" decomposition. Choosing sufficiently large blocks leads to convexification of the coarse–grained Hamiltonian; consequently, the Bakry–Emery principle implies a macroscopic LSI. On the other hand, the Holley–Stroock lemma implies a microscopic LSI as long as the block–spin size is bounded. We show that the macro– and microscopic LSI can be combined to yield a global LSI. The main ingredient in this final step is the Talagrand inequality.

Keywords. Logarithmic Sobolev inequality, spin system, Kawasaki dynamics, canonical ensemble.

AMS Subject Classification Numbers. Primary: 60K35; Secondary: 60J25, 82B21.

[*]Inst. für Angewandte Mathematik, Universität Bonn, Germany, grunewald@iam.uni-bonn.de.
[†]Inst. für Angewandte Mathematik, Universität Bonn, Germany, otto@iam.uni-bonn.de.
[‡]Inst. für Angewandte Mathematik, Universität Bonn, Germany, reznikoff@iam.uni-bonn.de.
[§]UMPA, École Normale Supérieure de Lyon, France, cvillani@umpa.ens-lyon.fr.

Contents

1 Introduction

Main result. In this paper we consider an N–site lattice system with continuous spin variables governed by a Ginzburg–Landau–type potential. We prove a uniform logarithmic Sobolev inequality (LSI) for the Gibbs measure of the *canonical ensemble* with a perturbed Gaussian potential. Analysis of the canonical ensemble is motivated by Kawasaki dynamics, in which the mean of the spins is preserved.

To be precise, suppose that ψ is a bounded perturbation of a Gaussian in the sense of conditions (5) and (6) below. (Think, for instance, of a double–well potential with quadratic growth at infinity.) The grand canonical measure $\mu_N \in \mathcal{P}(\mathbb{R}^N)$ has density

$$\frac{d\mu_N}{d\mathcal{L}^N}(x) \;=\; Z^{-1} \exp\big(-\sum_{i=1}^{N} \psi(x_i)\big),$$

where Z denotes the (generic) normalization constant for the probability measure. The canonical measure $\nu_{N,m}$ comes from restricting μ_N to $X_{N,m}$, the $(N-1)$–

dimensional hyperplane with mean m:

$$X_{N,m} \;=\; \left\{ (x_1,\ldots,x_N) \in \mathbb{R}^N \;\Big|\; \frac{1}{N}\sum_{i=1}^N x_i \;=\; m \right\}.$$

We offer a new proof of the optimal scaling of the LSI constant for Kawasaki dynamics on the periodic lattice. In $d = 1$, for instance, this means that for every smooth f on $X_{N,m}$, the measure $\nu_{N,m}$ satisfies

$$Ent(f^2) \;\le\; C\,N^2 \int_{X_{N,m}} \sum_{i=1}^N \Big(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_{i+1}}\Big)^2 d\nu_{N,m}, \tag{1}$$

where the constant C is uniform in N and m, and the relative entropy is defined as

$$Ent(f) := \int_{X_{N,m}} f \log f d\nu_{N,m} - \int_{X_{N,m}} f d\nu_{N,m} \int_{X_{N,m}} \log f d\nu_{N,m}.$$

We establish (1) by proving a uniform LSI for Glauber dynamics; see Theorem 1 for a precise statement. Roughly, we show

$$Ent(f^2) \;\le\; C \int_{X_{N,m}} \sum_{i=1}^N \Big(\frac{\partial f}{\partial x_i}\Big)^2 d\nu_{N,m}, \tag{2}$$

where we think of f as being extended to be constant normal to $X_{N,m}$. The Kawasaki LSI (1) follows from (2) by a discrete Poincaré inequality. Namely, for any function f with $\sum_{i=1}^N \frac{\partial f}{\partial x_i} = 0$ (i.e. f is constant normal to $X_{N,m}$),

$$\sum_{i=1}^N \Big(\frac{\partial f}{\partial x_i}\Big)^2 \le C\,N^2 \sum_{i=1}^N \Big(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_{i+1}}\Big)^2. \tag{3}$$

The analogous statement holds for the periodic lattice $\{1,\ldots,L\}^d \subset \mathbb{Z}^d$, with N^2 replaced by L^2; thus (2) implies the optimal L^2–scaling of the Kawasaki LSI constant in arbitrary dimension.

Method and motivation. We emphasize that rather than contributing a new result, we contribute a new technique. Indeed, (1) and (2) are proved in [LPY] and [Cha] by the Lu–Yau martingale method, summarized below. Since our result is not new, it seems appropriate to begin by saying a few words about our method and motivation.

The method which we introduce has as its essence a two–scale decomposition. In the language of physics, it involves a decomposition of a system of N spins into "block–spins" of size K. In the language of measure–theory, it involves the disintegration of the Gibbs measure $\nu_{N,m}$ into macroscopic and microscopic components.

The idea is to derive logarithmic Sobolev inequalities on the macro– and microscopic scales which combine to give (2). On the one hand, Proposition 1 says that choosing K sufficiently large convexifies the free energy of the subsystem, i.e. the "coarse–grained Hamiltonian." This induces a macroscopic LSI (Lemma 7) via the Bakry–Emery principle. On the other hand, for any finite K, the Gibbs measure of the subsystem is a bounded perturbation of a Gaussian. This induces a microscopic LSI (Lemma 8) via the Holley–Stroock lemma.

Combining the macro– and microscopic LSI requires an extra step. Although the relative entropy disintegrates in a straightforward way, the gradient terms do not. In order to bridge the gap, we introduce Proposition 2, which allows one to bound the macroscopic gradients by the original gradients using the fact that LSI implies the Talagrand inequality [OV]. In a different spirit but with similar ingredients, recent work of Blower and Bolley (cf. [BB], Theorem 1.3) also reconstructs a global LSI for a system based on LSI for its components.

Background. The logarithmic Sobolev inequality is a powerful tool for studying spin systems; in particular, it captures the convergence to equilibrium. (For a nicely written introduction to LSI, see [L], [R], or [GZ].) The spectral gap inequality, which has also received a lot of attention in the literature, follows from LSI via linearization (see for instance [Cha], p. 342).

There are (at least) two salient lines of research in the area of logarithmic Sobolev inequalities. One involves including interaction terms in the Hamiltonian, which considerably complicates the dynamics and raises the possibility of phase transition. For more about spin systems with interaction terms and, in particular, uniform LSI in the perturbative regime of sufficiently weak interactions, see [BH, H, SZ, Yo, Yo2, Z, GZ]. The second line of research and the one addressed in this paper is the study of the canonical ensemble, which as mentioned above is motivated by an interest in Kawasaki dynamics. In the case of a strictly convex potential, it is not hard to demonstrate a uniform LSI for the canonical ensemble using the Bakry–Emery principle [BE]. The nonconvex case, however, requires more.

The Lu–Yau martingale method is a well–known technique which has been used to prove LSI for the canonical ensemble with a nonconvex potential [LY, LPY, Cha]. LSI for a bounded perturbation of a Gaussian potential and Kawasaki dynamics is proved in [LPY] via the martingale method. An adaptation of the method in [Cha] extends the result to the (stronger) bound for Glauber dynamics.

For completeness and to contrast with the method presented in this paper, we briefly summarize the martingale method. The first step is to establish LSI for the one–site marginals. Subsequently, one seeks a recursive relationship for the N–site LSI constant in terms of the $(N-1)$–site LSI constant. Turning to the conditional expectations

$$f_k := \mathbb{E}_{\nu_{N,m}}(f|x_1,\ldots,x_k),$$

one appeals to a Markovian decomposition of the relative entropy into a sum of terms of the form

$$a_k := \mathbb{E}_{\nu_{N,m}}\big(f_k \log f_k - f_{k-1}\log f_{k-1}\big),$$

each of which depends only on a single spin. After applying the single–site LSI to each term, one wants to conclude by bounding the derivatives of a_k in terms of the derivatives of f. The central ingredient involves estimating the covariance terms from the Markovian decomposition by a variance term and a gradient term. Clever but elementary estimates produce the desired recursive relation and complete the argument.

Our method is more simple–minded. The martingale method, with the one–site distributions, the control of covariances on large enough blocks, and the recursive relationship between the LSI constant on $(N-1)$–blocks and N–blocks, operates on several scales. Ours operates on just two: the coarse measure on the blocks, and the fine measure on the microscale. Moreover, we require just one thing from equilibrium statistical mechanics: the strict convexity of the limiting free energy. This is a natural object on which to rely; it is precisely the strict convexity of the limit that rules out phase transition.

A gradient–flow on the space of measures. We prefer the following equivalent reformulation of (2): the measure ν satisfies the LSI with constant C if for every $\mu \in \mathcal{P}(X)$

$$\int_X \log\frac{d\mu}{d\nu}\,d\mu \;\le\; C\int_X |\nabla \log\frac{d\mu}{d\nu}|^2\,d\mu, \tag{4}$$

with the gradient corresponding to the choice of dynamics; see below. Our "measure–focused approach" is motivated by the following point of view, explained in [OV]: The Fokker–Planck or forward Kolmogorov equation of the stochastic process can be expressed formally as

$$\frac{\partial \mu_t}{\partial t} = -\mathrm{grad}\, Ent(\mu_t),$$

a gradient flow of the relative entropy on $\mathcal{P}(X)$, the space of probability measures on the state space X. One gives a sense to "grad" by endowing $\mathcal{P}(X)$ with the appropriate Riemannian structure, introduced in [O]. Using this formalism, the LSI can be written:

$$Ent(\mu_t) \;\le\; C\,|\mathrm{grad}\, Ent(\mu_t)|^2,$$

which makes it clear that LSI implies exponential convergence to the equilibrium measure:

$$\frac{d}{dt}(Ent(\mu_t)) \;=\; \langle \mathrm{grad}\, Ent(\mu_t), \frac{d\mu_t}{dt}\rangle_{\mu_t} \;=\; -|\mathrm{grad}\, Ent(\mu_t)|^2 \;\le -\frac{1}{C}Ent(\mu_t).$$

Moreover, the Bakry–Emery principle and Talagrand inequality both fit naturally into this framework (cf. [OV], Section 3).

From the gradient–flow perspective, there are two ingredients for the spin system: the Hamiltonian, which determines the Gibbs measure and thus the notion of relative entropy, and the gradient on the state space X which appears in (4). The gradient determines the type of dynamics and appears in the generator of the underlying stochastic process. In this paper we consider the usual gradient on $X_{N,m}$ coming from the Euclidean structure of $\mathbb{R}^N$, relevant for Glauber dynamics.

2 Set–up and statement of the result

2.1 General notation

All configuration spaces X will be affine subspaces of some $\mathbb{R}^N$ of some dimension K. The space X inherits the geometry of the ambient $\mathbb{R}^N$. To be more precise, the tangent space TX to X is a linear subspace of $\mathbb{R}^N$ of dimension K which inherits the scalar product $v \cdot w$ and the related norm $|v|$ from the ambient space. The ambient structure also allows one to give an unambiguous meaning to the gradient $\nabla\zeta \in TX$ and the Hessian $\mathrm{Hess}\zeta \in \mathrm{End}(TX)$ of a function ζ on X. This is also how the K–dimensional Hausdorff measure $\mathcal{H}^K$ on X is to be understood.

We denote by $\mathcal{P}(X)$ the space of probability measures on X. If $\mu \in \mathcal{P}(X)$ is absolutely continuous with respect to $\nu \in \mathcal{P}(X)$, we denote the Radon–Nykodim density by $\frac{d\mu}{d\nu}$. In particular, $\frac{d\mu}{d\mathcal{H}^K}$ denotes the Lebesgue density.

2.2 Gibbs measures

Let the potential $\psi \colon \mathbb{R} \to \mathbb{R}$ be given. For our main result, we shall assume that ψ is a bounded perturbation of a quadratic potential and twice differentiable:

$$\sup_{\mathbb{R}} |\delta\psi(x)| \;<\; \infty \quad \text{where } \delta\psi(x) = \psi(x) - \frac{1}{2}x^2, \tag{5}$$

$$\sup_{\mathbb{R}} |\frac{d^2\psi}{dx^2}(x)| \;<\; \infty. \tag{6}$$

We consider the related Gibbs measure $\mu \in \mathcal{P}(\mathbb{R})$ given by its Lebesgue density

$$\frac{d\mu}{d\mathcal{L}^1}(x) \;=\; Z^{-1}\,\exp(-\psi(x)). \tag{7}$$

Here and in the sequel, Z denotes the (generic) normalization factor (also called the partition function) which ensures that the measure is a probability measure.

Next, for $N \in \mathbb{N}$ we introduce the product measure $\mu_N \in \mathcal{P}(\mathbb{R}^N)$:

$$\mu_N \;=\; \underbrace{\mu \otimes \cdots \otimes \mu}_{N \text{ times}}.$$

In terms of Lebesgue–densities,

$$\frac{d\mu_N}{d\mathcal{L}^N}(x) \;=\; \Pi_{i=1}^N \frac{d\mu}{d\mathcal{L}^1}(x_i) \;=\; Z^{-1}\exp\big(-\sum_{i=1}^N \psi(x_i)\big). \tag{8}$$

By analogy with statistical mechanics, μ_N is called the grand canonical ensemble. In the language of probability theory, it describes the joint distribution of the N independent random variables $x_1, \ldots, x_N$ which are identically distributed according to the law μ.

Finally, for given $m \in \mathbb{R}$, we consider the affine subspace $X_{N,m}$ and the probability measure $\nu_{N,m} \in \mathcal{P}(X_{N,m})$ defined by

$$X_{N,m} \;=\; \left\{ (x_1, \ldots, x_N) \in \mathbb{R}^N \,\middle|\, \frac{1}{N}\sum_{i=1}^N x_i \;=\; m \right\}$$

and

$$\frac{d\nu_{N,m}}{d\mathcal{H}^{N-1}}(x) \;=\; Z^{-1}\frac{d\mu_N}{d\mathcal{L}^N}(x). \tag{9}$$

In distinction to μ_N, $\nu_{N,m}$ is called the canonical ensemble. It is the above–mentioned distribution of the random variables $x_1, \ldots, x_N$ conditioned on the event that their mean value is given by m, i.e. $N^{-1}\sum_{i=1}^N x_i = m$. It describes the distribution of the fluctuations around the macroscopic observables.

2.3 Logarithmic Sobolev inequality

Consider two probability measures $\mu, \nu \in \mathcal{P}(X)$ where μ is absolutely continuous with respect to ν, with Radon–Nykodim density $h = \frac{d\mu}{d\nu}$. Then the expression

$$\int_X \log\frac{d\mu}{d\nu}\,d\mu \;=\; \int_X h\,\log h\,d\nu$$

is non–negative and zero if and only if $\mu = \nu$. It is called the relative entropy of μ with respect to ν. The expression

$$\int_X |\nabla \log\frac{d\mu}{d\nu}|^2\,d\mu \;=\; \int_X \frac{1}{h}\,|\nabla h|^2\,d\nu \;=\; 4\int_X |\nabla\sqrt{h}|^2\,d\nu$$

is called the Fisher information.

Definition 1. *We say that $\nu \in \mathcal{P}(X)$ satisfies the logarithmic Sobolev inequality (LSI) with constant C provided*

$$\int_X \log\frac{d\mu}{d\nu}\,d\mu \;\le\; C\int_X |\nabla\log\frac{d\mu}{d\nu}|^2\,d\mu \tag{10}$$

holds for all $\mu \in \mathcal{P}(X)$.

The main result of this paper is:

Theorem 1. *Let ψ satisfy (5) & (6). Then $\nu_{N,m}$ defined as in (9) satisfies the LSI with a constant dependent only on ψ (that is, not on $N \in \mathbb{N}$ and $m \in \mathbb{R}$).*

Our proof will use a few basic facts about the LSI. We state them here because they are also helpful in discussing Theorem 1.

Lemma 1 (Factorization). *Let $\mu_i \in \mathcal{P}(X_i)$, $i = 1, 2$. Assume that μ_1 and μ_2 satisfy the LSI with constants C_1 and C_2, respectively. Consider the product measure $\mu_1 \otimes \mu_2 \in \mathcal{P}(X_1 \times X_2)$. It satisfies the LSI with constant $\max\{C_1, C_2\}$.*

Lemma 2 (Holley–Stroock). *Let $\mu \in \mathcal{P}(X)$ satisfy the LSI with constant C. For a bounded function $\delta\Psi\colon X \to \mathbb{R}$ consider the measure $\tilde{\mu} \in \mathcal{P}(X)$ defined through*

$$\frac{d\tilde{\mu}}{d\mu}(x) \;=\; Z^{-1} \exp(-\delta\Psi(x)).$$

Then $\tilde{\mu}$ satisfies the LSI with constant $\exp(\operatorname{osc}_X \delta\Psi)\, C$, where $\operatorname{osc}_X \delta\Psi = \sup_X \delta\Psi - \inf_X \delta\Psi$ denotes the oscillation of $\delta\Psi$.

Lemma 3 (Bakry–Emery). *Let Ψ be a twice differentiable function on X, a K–dimensional affine subspace of $\mathbb{R}^N$. Assume there exists a constant $c > 0$ such that*

$$\forall x \in X \quad \forall v \in TX \quad v \cdot \mathrm{Hess}\Psi(x)\, v \;\geq\; c\, |v|^2.$$

Consider $\mu \in \mathcal{P}(X)$ defined by

$$\frac{d\mu}{d\mathcal{H}^K} \;=\; Z^{-1} \exp(-\Psi(x)).$$

Then μ satisfies LSI with constant $(2c)^{-1}$.

For the proof of Lemma 1 see, for instance, [GZ]. See [HS] for Lemma 2 and [BE] for Lemma 3 (or also [OV]).

2.4 Grand canonical ensemble

The statement of Theorem 1 is also valid if we replace the canonical ensemble $\nu_{N,m}$ by the grand canonical ensemble μ_N. This is an easy consequence of the previous lemmas. Indeed, according to Lemma 3, the LSI holds for the Gaussian measure g on $\mathbb{R}$, that is

$$\frac{dg}{d\mathcal{L}^1}(x) \;=\; \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$$

with constant $\frac{1}{2}$. Next, in view of (5) and according to Lemma 2, the LSI holds for μ defined in (7) with the constant

$$C \;=\; \frac{1}{2} \exp(\operatorname{osc}_{\mathbb{R}} \delta\psi).$$

Finally, we learn from Lemma 1 that the LSI holds for μ_N with the same constant C.

Hence it is only with the conditioning of the grand canonical ensemble on the mean that Theorem 1 is interesting.

2.5 Strict convexity

Let us for a moment assume that ψ is uniformly strictly convex, i.e. that there is a constant $c > 0$ such that

$$\forall x \in \mathbb{R} \quad \frac{d^2\psi}{dx^2}(x) \ \geq \ c.$$

Now consider the potential Ψ on $\mathbb{R}^N$ defined by

$$\Psi(x) \ = \ \sum_{i=1}^{N} \psi(x_i).$$

The potential is strictly uniformly convex with the same constant:

$$\forall x \in \mathbb{R}^N \quad \forall v \in \mathbb{R}^N \quad v \cdot \mathrm{Hess}\Psi(x)\, v \ = \ \sum_{i=1}^{N} v_i \, \frac{d^2\psi}{dx^2}(x_i)\, v_i \ \geq \ c\, |v|^2.$$

Notice that in view of (7), (8) and (9), $\nu_{N,m}$ is of the form

$$\frac{d\nu_{N,m}}{d\mathcal{H}^{N-1}}(x) \ = \ Z^{-1} \exp(-\Psi(x)).$$

Thus an application of Lemma 3 yields that $\nu_{N,m}$ satisfies the LSI with constant $(2c)^{-1}$.

Hence it is only with the conditioning on the mean in conjunction with the non–convexity of ψ that Theorem 1 is interesting.

2.6 Equivalence of ensembles

It is an almost folkloristic statement that the canonical ensemble and a *suitably modified* grand canonical ensemble are equivalent. Let us explain what is meant by this.

It can be seen from (7), (8) and (9) that the canonical ensemble $\nu_{N,m}$ is not affected by changing ψ by a *linear* function. More precisely, we consider the measure μ_λ on $\mathbb{R}$ given by

$$\frac{d\mu_\lambda}{d\mathcal{L}^1}(x) \ = \ Z^{-1} \ \exp(\lambda\, x - \psi(x)). \tag{11}$$

The idea is to choose $\lambda = \lambda(m)$ such that $\mu_\lambda = \mu_m$ has first moment m:

$$\int_{\mathbb{R}} x \mu_\lambda(dx) \ = \ m. \tag{12}$$

Notice that equation (12) can be rewritten as

$$\frac{d\bar{\psi}^*}{d\lambda}(\lambda) \ = \ m, \tag{13}$$

where $\bar{\psi}^*$ is defined via

$$\bar{\psi}^*(\lambda) \;=\; \log \int_{\mathbb{R}} \exp(\lambda\, x - \psi(x))\, \mathcal{L}^1(dx). \tag{14}$$

It is well–known and easy to show (cf. Lemma 13 in Subsection 5.1) that $\bar{\psi}^*$ is uniformly strictly convex, that is, there exists a $c > 0$ such that

$$\forall \lambda \in \mathbb{R}, \quad \frac{d^2\bar{\psi}^*}{d\lambda^2}(\lambda) \;\geq\; c.$$

This means that (13) has a unique solution which can be expressed in terms of the Legendre transform $\bar{\psi}$ of $\bar{\psi}^*$. More precisely, we have

$$\lambda \;=\; \frac{d\bar{\psi}}{dm}(m),$$

where $\bar{\psi}$ is defined via

$$\bar{\psi}(m) \;=\; \sup_{\lambda \in \mathbb{R}} \left(\lambda\, m - \bar{\psi}^*(\lambda)\right).$$

We recall another elementary property of the Legendre transform: Provided m and λ are related by (13), we have

$$\bar{\psi}(m) + \bar{\psi}^*(\lambda) \;=\; m\, \lambda. \tag{15}$$

We thus have the representation

$$\frac{d\mu_m}{d\mathcal{L}^1}(x) \overset{(11),(14)}{=} \exp(-\bar{\psi}^*(\lambda) + \lambda\, x - \psi(x)) \tag{16}$$

$$\overset{(15)}{=} \exp(\bar{\psi}(m) + \lambda\, (x - m) - \psi(x)). \tag{17}$$

As before, we denote by $\mu_{N,m}$ the product measure on $\mathbb{R}^N$ arising from μ_m:

$$\mu_{N,m} \;=\; \underbrace{\mu_m \otimes \cdots \otimes \mu_m}_{N \text{ times}}. \tag{18}$$

This is our modified grand canonical ensemble. It is constructed such that the expected value of the mean coincides with m

$$\int_{\mathbb{R}^N} \left(N^{-1} \sum_{i=1}^{N} x_i \right) \mu_{N,m}(dx) \;=\; m. \tag{19}$$

As it did for μ_N, the conditioning of $\mu_{N,m}$ on $N^{-1}\sum_{i=1}^N x_i = m$ gives rise to $\nu_{N,m}$. In view of (19), this conditioning is expected to be less dramatic. One expects the measures $\mu_{N,m}$ and $\nu_{N,m}$ to be "close" for $N \gg 1$.

2.7 Free energy

Statistical mechanics suggests to measure the closeness of canonical ensemble $\nu_{N,m}$ and grand canonical ensemble $\mu_{N,m}$ in terms of the free energy. The free energy of the canonical ensemble is by definition

$$-\frac{1}{N}\log\left(\int_{X_{N,m}}\exp(\sum_{i=1}^{N}\lambda x_i-\psi(x_i))\,\mathcal{H}^{N-1}(dx)\right) \;=\; -\lambda m+\psi_N(m),$$

where

$$\psi_N(m) \;:=\; -\frac{1}{N}\log\left(\int_{X_{N,m}}\exp(-\sum_{i=1}^{N}\psi(x_i))\,\mathcal{H}^{N-1}(dx)\right). \qquad (20)$$

Notice that in view of (7), (8), and (9) we may rewrite $\nu_{N,m}$ as follows:

$$\frac{d\nu_{N,m}}{d\mathcal{H}^{N-1}}(x) \;=\; \exp\big(N\psi_N(m)-\sum_{i=1}^{N}\psi(x_i)\big), \qquad (21)$$

The free energy of the grand canonical ensemble $\mu_{N,m}$, on the other hand, may be expressed using (14) and (18) as:

$$-\bar{\psi}^*(\lambda) \;\overset{(15)}{=}\; -\lambda m+\bar{\psi}(m),$$

so that the difference of free energies is given by $\bar{\psi}(m)-\psi_N(m)$.

It is a well–known and crucial observation that the difference of the free energies can be characterized in terms of the central limit theorem. More precisely, let $g_{N,m}(ds)$ denote the distribution of $s = N^{-1/2}\sum_{i=1}^{N}(x_i-m)$ with respect to the modified grand canonical ensemble:

$$\int_{\mathbb{R}}\zeta(s)\,g_{N,m}(ds) \;=\; \int_{\mathbb{R}^N}\zeta\big(N^{-1/2}\sum_{i=1}^{N}(x_i-m)\big)\,\mu_{N,m}(dx). \qquad (22)$$

Then we have:

Lemma 4 (Cramer).

$$\exp\big(N\,(\bar{\psi}(m)-\psi_N(m))\big) \;=\; \frac{dg_{N,m}}{d\mathcal{L}^1}(0).$$

From this representation and a finer version of the central limit theorem, we will infer Proposition 1, the closeness of the free energies in the C^2–topology. While we believe the result is probably well–known, for completeness we include an elementary proof. Some of the elements may also be found, for instance, in [F] Chapter XVI, [KL] Appendix 2, [GPV] Section 3, and [LPY], p.752 and Section 5.

Proposition 1.

$$\lim_{N\uparrow\infty}\frac{d^2\psi_N}{dm^2}(m) \;=\; \frac{d^2\bar{\psi}}{dm^2}(m), \quad \textit{uniformly in } m\in\mathbb{R}.$$

This piece of information about the free energy is all we need to know about the closeness of ensembles. In particular, we will not invoke closeness of low–dimensional marginals of $\nu_{N,m}$ and $\mu_{N,m}$ as done in [GPV, LPY, Cap]. In fact, we need Proposition 1 only to conclude

Corollary 1. *There exist a $c > 0$ and an $N_0 < \infty$ dependent only on ψ such that for $N \geq N_0$ we have*

$$\forall m \in \mathbb{R} \quad \frac{d^2\psi_N}{dm^2}(m) \;\geq\; c.$$

The corollary follows immediately from the strict convexity of $\bar{\psi}$ and Proposition 1.

3 The two–scale argument

3.1 Projection and push–forward

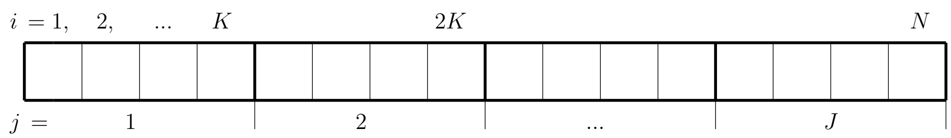

Figure 1: Microscopic and macroscopic observables.

Without loss of generality we will assume that $N \in \mathbb{N}$ can be written as

$$N \;=\; J\,K \quad \text{with} \quad J, K \;\in\; \mathbb{N}.$$

Central to our argument is the following linear map

$$P_{N,K}\colon X_{N,m} \to X_{J,m},\ x = (x_1, \cdots, x_N) \mapsto y = (y_1, \cdots, y_J),$$
$$\text{where} \quad y_j \;=\; K^{-1} \textstyle\sum_{i=(j-1)K+1}^{jK} x_i.$$

We view $P_{N,K}$ as the operator which relates the microscopic state x to the macroscopic observables (see Figure 1). The macroscopic observables are the local mean values $y = P_{N,K}x$.

The map $P_{N,K}$ induces a mapping between probability measures

$$P^{\#}_{N,K}\colon \mathcal{P}(X_{N,m}) \to \mathcal{P}(X_{J,m}), \quad \mu \mapsto P^{\#}_{N,K}\mu \;\stackrel{\text{short}}{=}\colon\; \bar{\mu}$$

defined by

$$\forall \zeta \in C^0_0(X_{J,m}) \quad \int_{X_{J,m}} \zeta(y)\,\bar{\mu}(dy) \;=\; \int_{X_{N,m}} \zeta(P_{N,K}x)\,\mu(dx). \tag{23}$$

Let us denote the fibers of the projection as follows

$$\forall y \in X_{J,m} \quad X_{N,K,y} \;:=\; \{x \in X_{N,m} | P_{N,K}x = y\}.$$

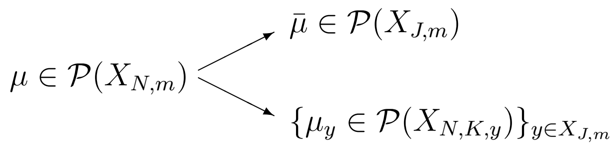

Figure 2: Disintegration of a measure.

Because of the coarea formula, that is,

$$\begin{aligned}\int_{X_{J,m}} \int_{X_{N,K,y}} \zeta(x)\, \mathcal{H}^{N-J}(dx)\, \mathcal{H}^{J-1}(dy) \;&=\; \int_{X_{N,m}} \zeta(x) \sqrt{\det(P_{N,K}\, P_{N,K}^t)}\, \mathcal{H}^{N-1}(dx) \\ &=\; K^{-(J-1)/2} \int_{X_{N,m}} \zeta(x)\, \mathcal{H}^{N-1}(dx),\end{aligned}$$

(23) assumes the following form on the level of densities:

$$\frac{d\bar{\mu}}{d\mathcal{H}^{J-1}}(y) \;=\; K^{(J-1)/2} \int_{X_{N,K,y}} \frac{d\mu}{d\mathcal{H}^{N-1}}(x)\, \mathcal{H}^{N-J}(dx). \tag{24}$$

In words, $\bar{\mu}$ is the push–forward of μ under $P_{N,K}$. In the language of probability theory: If μ denotes the law of the random variable x, then $\bar{\mu}$ denotes the law of $P_{N,K}x$.

3.2 Fibers and disintegration

The fact that $\{X_{N,K,y}\}_{y \in X_{J,m}}$ is a decomposition of $X_{N,m}$ can be used to disintegrate a probability measure $\mu \in \mathcal{P}(X_{N,m})$; see Figure 2. We refer for instance to [AGS] [Chapter 5.3]. There exists a unique (up to $\bar{\mu}$–negligible sets) family of measures $\{\mu_y \in \mathcal{P}(X_{N,K,y})\}_{y \in X_{J,m}}$ such that

$$\mu \;=\; \int_{X_{J,m}} \mu_y\; \bar{\mu}(dy),$$

which is short for

$$\forall \zeta \in C_0^0(X_{N,m}) \quad \int_{X_{N,m}} \zeta(x)\, \mu(dx) \;=\; \int_{X_{J,m}} \int_{X_{N,K,y}} \zeta(x)\, \mu_y(dx)\, \bar{\mu}(dy). \tag{25}$$

On the level of densities this means

$$K^{(J-1)/2} \frac{d\mu}{d\mathcal{H}^{N-1}}(x) \;=\; \frac{d\bar{\mu}}{d\mathcal{H}^{J-1}}(y)\, \frac{d\mu_y}{d\mathcal{H}^{N-J}}(x), \quad \text{where} \quad y \;=\; P_{N,K}\, x. \tag{26}$$

In the language of probability theory: If μ is the law describing the random variable x, then μ_y is the distribution of x conditioned on the event that $P_{N,K}x = y$.

3.3 Disintegration of the Gibbs state

We apply the disintegration of the previous subsection to our Gibbs measure

$$\nu \overset{\text{short}}{=} \nu_{N,m} \in \mathcal{P}(X_{N,m}).$$

This gives rise to

$$\bar{\nu} \overset{\text{short}}{=} P_{N,K}^{\#}\nu \in \mathcal{P}(X_{J,m}) \quad \text{and} \quad \{\nu_y \in \mathcal{P}(X_{N,K,y})\}_{y\in X_{J,m}}.$$

The following characterization is an easy consequence of the fact that the fibers factorize,

$$X_{N,K,y} \;=\; \otimes_{j=1}^{J} X_{K,y_j}, \tag{27}$$

and of the formulas (24), (26) ,(21) and (20) :

Lemma 5 (Disintegration of the Gibbs state).

i)

$$\frac{d\bar{\nu}}{d\mathcal{H}^{J-1}}(y) \;=\; K^{(J-1)/2}\,\exp\big(N\psi_N(m) - K\sum_{j=1}^{J}\psi_K(y_j)\big).$$

ii)

$$\frac{d\nu_y}{d\mathcal{H}^{N-J}}(x) \;=\; \exp\big(K\sum_{j=1}^{J}\psi_K(y_j) - \sum_{i=1}^{N}\psi(x_i)\big).$$

3.4 Local Gibbs states

The disintegration of a probability measure $\mu \in \mathcal{P}(X_{N,m})$ into $\bar{\mu} \in \mathcal{P}(X_{J,m})$ and $\{\mu_y \in \mathcal{P}(X_{N,K,y})\}_{y\in X_{J,m}}$ leads to a disintegration of the relative entropy; see point i) under Lemma 6 below. Only the "second half" of the analogous disintegration holds for the Fisher information; see point ii) under Lemma 6. The "first half" is captured in point iii) in terms of the local Gibbs states:

Definition 2 (Local Gibbs state). *Let $\mu \in \mathcal{P}(X_{N,m})$ with its disintegration $\bar{\mu} \in \mathcal{P}(X_{J,m})$ and $\{\mu_y \in \mathcal{P}(X_{N,K,y})\}_{y\in X_{J,m}}$ be given. Let $\bar{\nu} \in \mathcal{P}(X_{J,m})$ and $\{\nu_y \in \mathcal{P}(X_{N,K,y})\}_{y\in X_{J,m}}$ denote the disintegration of the Gibbs measure $\nu \in \mathcal{P}(X_{N,m})$. Then the local Gibbs state $\tilde{\mu} \in \mathcal{P}(X_{N,m})$ is the measure which comes from $\bar{\mu} \in \mathcal{P}(X_{J,m})$ and $\{\nu_y \in \mathcal{P}(X_{N,K,y})\}_{y\in X_{J,m}}$ (Figure 3). In terms of densities:*

$$K^{(J-1)/2}\frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}}(x) \;=\; \frac{d\bar{\mu}}{d\mathcal{H}^{J-1}}(y)\,\frac{d\nu_y}{d\mathcal{H}^{N-J}}(x) \quad \textit{where } y = P_{N,K}x. \tag{28}$$

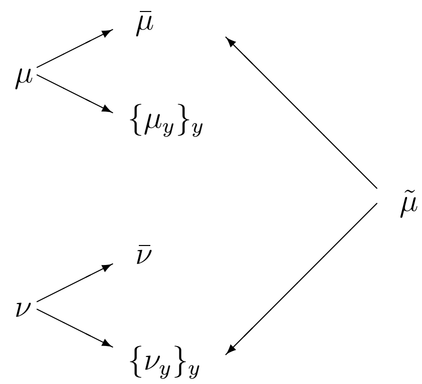

Figure 3: Local Gibbs state.

Lemma 6 (Entropy and Fisher information under disintegration).

i)

$$\int_{X_{N,m}} \log \frac{d\mu}{d\nu}\, d\mu = \int_{X_{J,m}} \log \frac{d\bar\mu}{d\bar\nu}\, d\bar\mu + \int_{X_{J,m}} \left(\int_{X_{N,K,y}} \log \frac{d\mu_y}{d\nu_y}\, d\mu_y \right) \bar\mu(dy).$$

ii)

$$\int_{X_{N,m}} |\nabla^{\|} \log \frac{d\mu}{d\nu}|^2\, d\mu \;=\; \int_{X_{J,m}} \left(\int_{X_{N,K,y}} |\nabla \log \frac{d\mu_y}{d\nu_y}|^2\, d\mu_y \right) \bar\mu(dy),$$

where $\nabla^{\|}$ denotes the gradient parallel to the fibers.

iii)

$$\int_{X_{N,m}} |\nabla \log \frac{d\tilde\mu}{d\nu}|^2\, d\tilde\mu \;=\; K^{-1} \int_{X_{J,m}} |\nabla \log \frac{d\bar\mu}{d\bar\nu}|^2\, d\bar\mu.$$

3.5 Two–scale LSI and proof of Theorem 1

We prove Theorem 1 by combining LSI for the macroscopic measure $\bar\nu$ and the microscopic measures ν_y (see Figure 4 for a sketch). We begin by stating the three ingredients. The first lemma establishes the LSI for the coarse grained Gibbs measure $\bar\nu \in \mathcal{P}(X_{J,m})$. It is an easy consequence of the representation of $\bar\nu$ in Lemma 5, of Corollary 1 (uniform strict convexity of ψ_K for sufficiently large K), and of the Bakry–Emery principle (cf. Lemma 3).

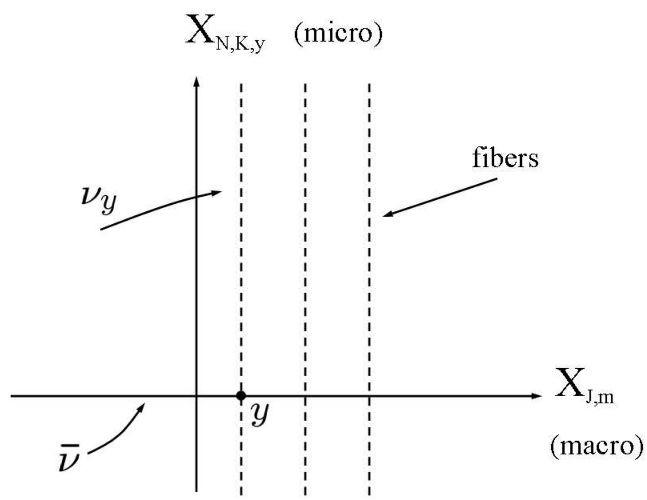

Figure 4: We combine a macroscopic LSI for $\bar\nu$ and a microscopic LSI for ν_y.

Lemma 7 (Macro–LSI). *For all $m \in \mathbb{R}$ and $\bar\mu \in \mathcal{P}(X_{J,m})$ we have*

$$\int_{X_{J,m}} \log \frac{d\bar\mu}{d\bar\nu}\, d\bar\mu \;\le\; \left(2\,K\, \inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2}\right)^{-1} \int_{X_{J,m}} |\nabla \log \frac{d\bar\mu}{d\bar\nu}|^2\, d\bar\mu.$$

The second lemma establishes the LSI for the conditional Gibbs measures $\nu_y \in \mathcal{P}(X_{N,K,y})$ uniformly in $y \in X_{J,m}$. It is an easy consequence of the representation of ν_y in Lemma 5, of Lemma 1 applied to the factorization (27), and of the Holley–Stroock argument (cf. Lemma 2).

Lemma 8 (Micro–LSI). *For all $y \in X_{J,m}$ and $\mu_y \in \mathcal{P}(X_{N,K,y})$ we have*

$$\int_{X_{N,K,y}} \log \frac{d\mu_y}{d\nu_y}\, d\mu_y \;\le\; \frac{1}{2} \exp\left(K \operatorname{osc}_{\mathbb{R}} \delta\psi\right) \int_{X_{N,K,y}} |\nabla \log \frac{d\mu_y}{d\nu_y}|^2\, d\mu_y.$$

Lemmas 7 and 8 show the trade–off in the choice of the "block–spin size" K. On the one hand, K has to be sufficiently large so that the fluctuations convexify the free energy ψ_K of the subsystem. On the other hand, K has to be so small that the Gibbs measure ν_y of the subsystem can be treated as a finite perturbation of a Gaussian.

The third and most delicate ingredient relates the Fisher information of $\tilde\mu$ to that of μ. It comes in the form of a perturbative result: The simple inequality which holds in the Gaussian case is modified.

Proposition 2.

$$\left(\int_{X_{N,m}} |\nabla \log \frac{d\tilde{\mu}}{d\nu}|^2 d\tilde{\mu}\right)^{1/2}$$
$$\leq \left(\int_{X_{N,m}} |\nabla^{\perp} \log \frac{d\mu}{d\nu}|^2 \, d\mu\right)^{1/2}$$
$$+ \sup_{\mathbb{R}} |\frac{d^2 \delta\psi}{dx^2}| \, \exp(K \operatorname{osc}_{\mathbb{R}} \delta\psi) \left(\int_{X_{N,m}} |\nabla^{\parallel} \log \frac{d\mu}{d\nu}|^2 \, d\mu\right)^{1/2} .$$

Here $\nabla^{\perp}$ and $\nabla^{\parallel}$ denote the projection of the gradient on $X_{N,m}$ perpendicular and tangential to the fibers, respectively.

We are ready to prove Theorem 1.
Proof of Theorem 1.
We have for any $\mu \in \mathcal{P}(X_{N,m})$:

$$\int_{X_{N,m}} \log \frac{d\mu}{d\nu} \, d\mu$$
$$= \int_{X_{J,m}} \log \frac{d\bar{\mu}}{d\bar{\nu}} \, d\bar{\mu} + \int_{X_{J,m}} \left(\int_{X_{N,K,y}} \log \frac{d\mu_y}{d\nu_y} \, d\mu_y\right) \bar{\mu}(dy),$$

according to Lemma 6 i)

$$\leq (2\, K \inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2})^{-1} \int_{X_{J,m}} |\nabla \log \frac{d\bar{\mu}}{d\bar{\nu}}|^2 \, d\bar{\mu}$$
$$+ \frac{1}{2} \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big) \int_{X_{J,m}} \left(\int_{X_{N,K,y}} |\nabla \log \frac{d\mu_y}{d\nu_y}|^2 d\mu_y\right) \bar{\mu}(dy)$$

according to Lemmas 7 and 8

$$= (2 \inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2})^{-1} \int_{X_{N,m}} |\nabla \log \frac{d\tilde{\mu}}{d\nu}|^2 \, d\tilde{\mu}$$
$$+ \frac{1}{2} \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big) \int_{X_{N,m}} |\nabla^{\parallel} \log \frac{d\mu}{d\nu}|^2 \, d\mu$$

according to Lemma 6 ii) and iii)

$$\leq (\inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2})^{-1} \int_{X_{N,m}} |\nabla^{\perp} \log \frac{d\mu}{d\nu}|^2 \, d\mu$$
$$+ \left((\inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2})^{-1} \sup_{\mathbb{R}} |\frac{d^2\delta\psi}{dx^2}|^2 \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big) + \frac{1}{2}\right)$$
$$\cdot \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big) \int_{X_{N,m}} |\nabla^{\parallel} \log \frac{d\mu}{d\nu}|^2 \, d\mu$$

according to Proposition 2 and Young's inequality.

Hence ν satisfies the LSI with constant

$$\max\Big\{ \big(\inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2}\big)^{-1},$$
$$\Big(\big(\inf_{\mathbb{R}} \frac{d^2\psi_K}{dm^2}\big)^{-1} \sup_{\mathbb{R}} |\frac{d^2 \delta\psi}{dx^2}|^2 \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big) + \frac{1}{2}\Big) \exp\big(K \operatorname{osc}_{\mathbb{R}} \delta\psi\big)\Big\}.$$

□

3.6 Proofs of the lemmas

Proof of Lemma 4
Notice that for all $\zeta \in C_0^0(\mathbb{R})$ we have

$$\begin{aligned}
\int_{\mathbb{R}} \zeta(s)\, \frac{dg_{N,m}}{d\mathcal{L}^1}(s)\mathcal{L}^1(ds) &= \int_{\mathbb{R}} \zeta(s)\, g_{N,m}(ds) \\
&\overset{(22)}{=} \int_{\mathbb{R}^N} \zeta(N^{-\frac{1}{2}} \sum_{i=1}^N (x_i - m))\, \frac{d\mu_{N,m}}{d\mathcal{L}^N}(x)\, \mathcal{L}^N(dx) \\
&= \int_{\mathbb{R}} \zeta(s) \int_{\{N^{-\frac{1}{2}} \sum_{i=1}^N (x_i - m) = s\}} \frac{d\mu_{N,m}}{d\mathcal{L}^N}(x)\mathcal{H}^{N-1}(dx)\, \mathcal{L}^1(ds),
\end{aligned}$$

where we have used the coarea formula. Hence we obtain

$$\frac{dg_{N,m}}{d\mathcal{L}^1}(0) = \int_{\{N^{-\frac{1}{2}} \sum_{i=1}^N (x_i - m) = 0\}} \frac{d\mu_{N,m}}{d\mathcal{L}^N}(x)\mathcal{H}^{N-1}(dx). \tag{29}$$

Now we remark that

$$\frac{d\mu_{N,m}}{d\mathcal{L}^N}(x) = \Pi_{i=1}^N \frac{d\mu_m}{d\mathcal{L}^1}(x_i) \overset{(17)}{=} \exp(N\bar{\psi}(m))\, \exp\Big(\lambda \sum_{i=1}^N (x_i - m) - \sum_{i=1}^N \psi(x_i)\Big),$$

so that (29) turns into

$$\begin{aligned}
\frac{dg_{N,m}}{d\mathcal{L}^1}(0) &= \exp(N\bar{\psi}(m)) \int_{\{N^{-\frac{1}{2}} \sum_{i=1}^N (x_i - m) = 0\}} \exp(-\sum_{i=1}^N \psi(x_i))\mathcal{H}^{N-1}(dx) \\
&\overset{(20)}{=} \exp(N\bar{\psi}(m))\, \exp(-N\psi_N(m)).
\end{aligned}$$

□

Proof of Lemma 5.
We start with (21), which we rewrite as

$$\frac{d\nu}{d\mathcal{H}^{N-1}}(x) = \exp(N\psi_N(m))\, \Pi_{j=1}^J \exp(-\sum_{i=(j-1)K+1}^{jK} \psi(x_i)).$$

Together with the factorization (27) and (24) we obtain

$$\begin{aligned}
\frac{d\bar\nu}{d\mathcal{H}^{J-1}}(y) &= K^{(J-1)/2}\,\exp(N\psi_N(m))\,\Pi_{j=1}^J\int_{X_{K,y_j}}\exp(-\sum_{i=1}^K\psi(x_i))\,\mathcal{H}^{K-1}(dx)\\
&\overset{(20)}{=} K^{(J-1)/2}\,\exp(N\psi_N(m))\,\Pi_{j=1}^J\exp(-K\psi_K(y_j))\\
&= K^{(J-1)/2}\,\exp\big(N\psi_N(m)-K\sum_{j=1}^J\psi_K(y_j)\big).
\end{aligned}$$

This proves part i) of Lemma 5.

Part ii) follows from part i) together with (26) and (21).

□

PROOF OF LEMMA 6.
We deduce from (26) for μ and for μ replaced by ν that

$$\frac{d\mu}{d\nu}(x) \;=\; \frac{d\bar\mu}{d\bar\nu}(y)\,\frac{d\mu_y}{d\nu_y}(x) \quad\text{where}\quad y \;=\; P_{N,K}x, \tag{30}$$

hence

$$\log\frac{d\mu}{d\nu}(x) \;=\; \log\frac{d\bar\mu}{d\bar\nu}(y)+\log\frac{d\mu_y}{d\nu_y}(x). \tag{31}$$

Therefore,

$$\begin{aligned}
&\int_{X_{N,m}}\log\frac{d\mu}{d\nu}\,d\mu\\
&= \int_{X_{N,m}}(\log\frac{d\bar\mu}{d\bar\nu})(P_{N,K}x)\,\mu(dx)+\int_{X_{N,m}}(\log\frac{d\mu_{P_{N,K}x}}{d\nu_{P_{N,K}x}})(x)\,\mu(dx)\\
&\overset{(23),(25)}{=} \int_{X_{J,m}}(\log\frac{d\bar\mu}{d\bar\nu})(y)\,\bar\mu(dy)+\int_{X_{J,m}}\int_{X_{N,K,y}}(\log\frac{d\mu_y}{d\nu_y})(x)\,\mu_y(dx)\,\bar\mu(dy).
\end{aligned}$$

This establishes part i) of the Lemma.

We infer from (31) that

$$\nabla^{\|}(\log\frac{d\mu}{d\nu})(x) \;=\; \nabla(\log\frac{d\mu_y}{d\nu_y})(x),$$

where the second gradient is the gradient on the fibers. In particular,

$$|\nabla^{\|}(\log\frac{d\mu}{d\nu})(x)|^2 \;=\; |\nabla(\log\frac{d\mu_y}{d\nu_y})(x)|^2.$$

Hence we obtain part ii) of the lemma from integrating this identity against μ and using (25).

For part iii) of the lemma we observe that by definition (28) and by (30) for $\tilde{\mu}$,

$$\frac{d\tilde{\mu}}{d\nu}(x) \;=\; \frac{d\bar{\mu}}{d\bar{\nu}}(y) \quad \text{where} \quad y \;=\; P_{N,K}x,$$

so that

$$\nabla(\log\frac{d\tilde{\mu}}{d\nu})(x) \;=\; P_{N,K}^t\nabla(\log\frac{d\bar{\mu}}{d\bar{\nu}})(y).$$

In particular, we have

$$\begin{aligned}|\nabla(\log\frac{d\tilde{\mu}}{d\nu})(x)|^2 \;&=\; \nabla(\log\frac{d\bar{\mu}}{d\bar{\nu}})(y)\cdot P_{N,K}\,P_{N,K}^t\,\nabla(\log\frac{d\bar{\mu}}{d\bar{\nu}})(y)\\ &=\; K^{-1}\,|\nabla(\log\frac{d\bar{\mu}}{d\bar{\nu}})(y)|^2.\end{aligned}$$

Now part iii) follows from integrating this identity against $\tilde{\mu}$ and using (23).

□

Proof of Lemma 7.

According to Lemma 5, $\bar{\nu}$ is of the form

$$\frac{d\bar{\nu}}{d\mathcal{H}^{J-1}}(y) \;=\; Z^{-1}\exp\big(-K\sum_{j=1}^{J}\psi_K(y_j)\big).$$

Since the potential

$$\Psi_K(y) \;=\; K\sum_{j=1}^{J}\psi_K(y_j)$$

satisfies

$$v\cdot\mathrm{Hess}\Psi_K(y)\,v \;=\; K\sum_{j=1}^{J}v_j\,\frac{d^2\psi_K}{dx^2}(y_j)\,v_j \;\geq\; K\,\big(\inf_{\mathbb{R}}\frac{d^2\psi_K}{dx^2}\big)\,|v|^2,$$

we may apply Lemma (3).

□

Proof of Lemma 8.

We recall the factorization (27) of $X_{N,K,y}$:

$$x\in X_{N,K,y} \quad\Longleftrightarrow\quad \big\{\forall j\in\{1,\cdots,J\} \quad (x_{(j-1)K+1},\cdots,x_{jK})\in X_{K,y_j}\big\}.$$

According to Lemma 5, ν_y factorizes as

$$\begin{aligned}\frac{d\nu_y}{d\mathcal{H}^{N-J}}(x) \quad&=\quad \Pi_{j=1}^J\exp\big(K\psi_K(y_j)-\sum_{i=(j-1)K+1}^{jK}\psi(x_i)\big)\\ &\overset{(21)}{=}\quad \Pi_{j=1}^J\frac{d\nu_{K,y_j}}{d\mathcal{H}^{K-1}}(x_{(j-1)K+1},\cdots,x_{jK}).\end{aligned}$$

Thus in view of Lemma 1, it suffices to show that $\nu_{K,m} \in \mathcal{P}(X_{K,m})$ satisfies the LSI with constant $\frac{1}{2}\exp(K \operatorname{osc}_{\mathbb{R}} \delta\psi)$.

For this we argue as in Subsection 2.4: According to Lemma 3, the Gaussian measure $g \in \mathcal{P}(X_{K,m})$ given by

$$\frac{dg}{d\mathcal{H}^{K-1}}(x) \;=\; Z^{-1}\exp(-\sum_{i=1}^{K}\frac{1}{2}x_i^2)$$

satisfies the LSI with constant $\frac{1}{2}$. Writing

$$\frac{d\nu_{K,m}}{dg}(x) \;=\; Z^{-1}\,\exp(-\sum_{i=1}^{K}\delta\psi(x_i)),$$

we learn from Lemma 2 that $\nu_{K,m}$ satisfies the LSI with constant

$$\frac{1}{2}\exp\left(\operatorname{osc}_{x\in X_{K,m}}\sum_{i=1}^{K}\delta\psi(x_i)\right) \;\le\; \frac{1}{2}\exp\left(K\operatorname{osc}_{\mathbb{R}}\delta\psi\right).$$

□

4 Proof of Proposition 2

In order to establish Proposition 2, it is convenient to introduce the Wasserstein distance, defined below. The squared Wasserstein distance $d_X(\mu,\nu)^2$ to the Gibbs measure ν, the relative entropy $\int_X \log\frac{d\mu}{d\nu}\,d\mu$ with respect to the Gibbs measure, and the Fisher information $\int_X |\nabla\log\frac{d\mu}{d\nu}|^2\,d\mu$ form a natural triplet, see [O, OV]. The LSI relates relative entropy and Fisher information whereas the Talagrand inequality relates Wasserstein distance and relative entropy. In the proof of Lemma 10 below, we shall make use of the result that the LSI implies the Talagrand inequality, cf. [OV], Theorem 1.

Definition 3 (Wasserstein distance). *Let X be an affine subspace of $\mathbb{R}^N$. The squared Wasserstein distance $\mathrm{d}_X(\mu,\nu)^2$ between $\mu,\nu\in\mathcal{P}(X)$ is defined as the infimum of the "transportation cost"*

$$\int_{X\times X}|x-x'|^2\,\pi(dxdx')$$

over all "transference plans" $\pi\in\mathcal{P}(X\times X)$ satisfying

$$\forall\zeta\in C_0^0(X)\quad\left\{\begin{array}{rcl}\int_{X\times X}\zeta(x)\,\pi(dxdx') &=& \int_X\zeta(x)\,\mu(dx)\\ \int_{X\times X}\zeta(x')\,\pi(dxdx') &=& \int_X\zeta(x)\,\nu(dx)\end{array}\right\}.$$

Proposition 2 is a direct consequence of the following two lemmas combined with Lemma 6 ii).

Lemma 9.

$$\left(\int_{X_{N,m}} |\nabla \log \frac{d\tilde{\mu}}{d\nu}|^2 d\tilde{\mu}\right)^{1/2} \leq \left(\int_{X_{N,m}} |\nabla^\perp \log \frac{d\mu}{d\nu}|^2 \, d\mu\right)^{1/2}$$
$$+\sup_{\mathbb{R}} |\frac{d^2\delta\psi}{dx^2}| \left(\int_{X_{J,m}} \mathrm{d}_{X_{N,K,y}}(\mu_y,\nu_y)^2 \bar{\mu}(dy)\right)^{1/2}.$$

Lemma 10. *For all $y \in X_{J,m}$ and $\mu_y \in \mathcal{P}(X_{N,K,y})$ we have*

$$\mathrm{d}_{X_{N,K,y}}(\mu_y,\nu_y)^2 \leq (\exp(K \operatorname{osc}_{\mathbb{R}} \delta\psi))^2 \int_{X_{N,K,y}} |\nabla \log \frac{d\mu_y}{d\nu_y}|^2 \, d\mu_y.$$

Proof of Lemma 9. For abbreviation we introduce the potential

$$\Phi = \log \frac{d\tilde{\mu}}{d\nu}$$

and notice that by definition of the local Gibbs state $\tilde{\mu}$, Φ is a function of the macroscopic variables only:

$$\Phi(x) = \hat{\Phi}(P_{N,K}x).$$

The same holds for the gradient $\nabla\Phi$ and the Laplacian $\triangle\Phi = \nabla\cdot\nabla\Phi$:

$$\nabla\Phi(x) = P_{N,K}^t \nabla\hat{\Phi}(P_{N,K}x) \quad \text{and} \quad \triangle\Phi(x) = \operatorname{tr}\big(P_{N,K}^t \mathrm{Hess}\hat{\Phi}(P_{N,K}x)P_{N,K}\big). \tag{32}$$

On the other hand we have

$$\Phi(x) = (\log \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}})(x) - (\log \frac{d\nu}{d\mathcal{H}^{N-1}})(x) \stackrel{(21)}{=} \log \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} - N\psi_N(m) + \Psi(x),$$

where

$$\Psi(x) = \sum_{i=1}^N \psi(x_i).$$

This implies

$$\nabla\Phi = \frac{d\mathcal{H}^{N-1}}{d\tilde{\mu}} \nabla \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} + \nabla\Psi, \tag{33}$$

so that

$$\int_{X_{N,m}} |\nabla\Phi|^2 \, d\tilde{\mu} \stackrel{(33)}{=} \int_{X_{N,m}} \frac{d\mathcal{H}^{N-1}}{d\tilde{\mu}} \nabla \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} \cdot \nabla\Phi \, d\tilde{\mu} + \int_{X_{N,m}} \nabla\Psi\cdot\nabla\Phi \, d\tilde{\mu}. \tag{34}$$

For the first integral on the right–hand side we observe

$$\begin{aligned} \int_{X_{N,m}} \frac{d\mathcal{H}^{N-1}}{d\tilde{\mu}} \nabla \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} \cdot \nabla\Phi \, d\tilde{\mu} &= \int_{X_{N,m}} \nabla \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} \cdot \nabla\Phi \, d\mathcal{H}^{N-1} \\ &= -\int_{X_{N,m}} \frac{d\tilde{\mu}}{d\mathcal{H}^{N-1}} \triangle\Phi \, d\mathcal{H}^{N-1} \\ &= -\int_{X_{N,m}} \triangle\Phi \, d\tilde{\mu}. \end{aligned} \tag{35}$$

Because of (32), we have by definition of $\tilde{\mu}$

$$-\int_{X_{N,m}} \triangle\Phi \, d\tilde{\mu} \;=\; -\int_{X_{N,m}} \triangle\Phi \, d\mu. \tag{36}$$

We now obtain, repeating the arguments under (34) and (35) for μ instead of $\tilde{\mu}$:

$$\begin{aligned}
&\int_{X_{N,m}} |\nabla\Phi|^2 \, d\tilde{\mu} \\
&\overset{(34),(35),(36)}{=} \; -\int_{X_{N,m}} \triangle\Phi \, d\mu + \int_{X_{N,m}} \nabla\Psi\cdot\nabla\Phi \, d\tilde{\mu} \\
&= \int_{X_{N,m}} \nabla \log \frac{d\mu}{d\nu}\cdot\nabla\Phi \, d\mu + \int_{X_{N,m}} \nabla\Psi\cdot\nabla\Phi \, d(\tilde{\mu}-\mu).
\end{aligned} \tag{37}$$

We now turn to the second term in (37). By definition of the local Gibbs measure $\tilde{\mu}$ we have

$$\begin{aligned}
&\int_{X_{N,m}} \nabla\Psi\cdot\nabla\Phi \, d(\mu-\tilde{\mu}) \\
&= \int_{X_{J,m}} \int_{X_{N,K,y}} \nabla\Psi\cdot\nabla\Phi \, d(\mu_y-\nu_y)\, \bar{\mu}(dy) \\
&\overset{(32)}{=} \int_{X_{J,m}} \nabla\hat{\Phi}(y)\cdot\left(\int_{X_{N,K,y}} P_{N,K}\nabla\Psi \, d(\mu_y-\nu_y)\right)\bar{\mu}(dy).
\end{aligned} \tag{38}$$

We notice that the gradient on $X_{N,m}$ is given by

$$\nabla\Psi(x) \;=\; (x-m\mathbf{1}_N)+\nabla\delta\Psi(x),$$

where $\mathbf{1}_N\in\mathbb{R}^N$ is the vector with entries 1 (so that $\nabla\Psi\in TX_{N,m}$) and

$$\delta\Psi(x) \;=\; \sum_{i=1}^{N}\delta\psi(x_i).$$

Thus we have for $x\in X_{N,K,y}$

$$P_{N,K}\nabla\Psi(x) \;=\; y-m\mathbf{1}_J+P_{N,K}\nabla\delta\Psi(x),$$

and therefore

$$\int_{X_{N,K,y}} P_{N,K}\nabla\Psi \, d(\mu_y-\nu_y) \;=\; \int_{X_{N,K,y}} P_{N,K}\nabla\delta\Psi \, d(\mu_y-\nu_y).$$

Inserting this into (38) one obtains

$$\begin{aligned}
&\int_{X_{N,m}} \nabla\Psi\cdot\nabla\Phi \, d(\mu-\tilde{\mu}) \\
&= \int_{X_{J,m}} P_{N,K}^t\nabla\hat{\Phi}(y)\cdot\left(\int_{X_{N,K,y}} \nabla\delta\Psi \, d(\mu_y-\nu_y)\right)\bar{\mu}(dy).
\end{aligned} \tag{39}$$

The combination of (37) and (39) yields the following identity:

$$\begin{aligned}\int_{X_{N,m}} |\nabla\Phi|^2\, d\tilde\mu &= \int_{X_{N,m}} P^t_{N,K}\nabla\hat\Phi(P_{N,K}x)\cdot\nabla\log\frac{d\mu}{d\nu}(x)\,\mu(dx)\\ &\quad + \int_{X_{J,m}} P^t_{N,K}\nabla\hat\Phi(y)\cdot\left(\int_{X_{N,K,y}}\nabla\delta\Psi\, d(\mu_y-\nu_y)\right)\bar\mu(dy).\end{aligned}\tag{40}$$

We now turn to the estimates. By Cauchy–Schwarz we deduce from (40):

$$\begin{aligned}&\int_{X_{N,m}} |\nabla\Phi|^2\, d\tilde\mu\\ &\le \left(\int_{X_{N,m}} |P^t_{N,K}\nabla\hat\Phi(P_{N,K}x)|^2\,\mu(dx)\int_{X_{N,m}}|\nabla^\perp\log\frac{d\mu}{d\nu}|^2\,d\mu\right)^{1/2}\\ &+ \left(\int_{X_{J,m}} |P^t_{N,K}\nabla\hat\Phi|^2\,d\bar\mu\int_{X_{J,m}}\left|\int_{X_{N,K,y}}\nabla\delta\Psi\,d(\mu_y-\nu_y)\right|^2\bar\mu(dy)\right)^{1/2}\\ &\overset{(32)}{=} \left(\int_{X_{N,m}} |\nabla\Phi|^2\,d\tilde\mu\int_{X_{N,m}}|\nabla^\perp\log\frac{d\mu}{d\nu}|^2\,d\mu\right)^{1/2}\\ &+ \left(\int_{X_{N,m}} |\nabla\Phi|^2\,d\tilde\mu\int_{X_{J,m}}\left|\int_{X_{N,K,y}}\nabla\delta\Psi\,d(\mu_y-\nu_y)\right|^2\bar\mu(dy)\right)^{1/2},\end{aligned}\tag{41}$$

where we have used the definition of $\tilde\mu$ in the last step.

We take a closer look at the inner integral. It holds for any admissible π in the definition of $\mathrm{d}_{X_{N,K,y}}(\mu_y,\nu_y)^2$:

$$\begin{aligned}\left|\int_{X_{N,K,y}}\nabla\delta\Psi\,d(\mu_y-\nu_y)\right|^2 &= \left|\int_{X_{N,K,y}\times X_{N,K,y}}(\nabla\delta\Psi(x)-\nabla\delta\Psi(x'))\,\pi(dxdx')\right|^2\\ &\le \int_{X_{N,K,y}\times X_{N,K,y}}|\nabla\delta\Psi(x)-\nabla\delta\Psi(x')|^2\,\pi(dxdx').\end{aligned}\tag{42}$$

Notice that for any tangent vector $v\in TX_{N,K,y}$,

$$(\nabla\delta\Psi(x)-\nabla\delta\Psi(x'))\cdot v = \sum_{i=1}^N(\frac{d\delta\psi}{dx}(x_i)-\frac{d\delta\psi}{dx}(x_i'))\,v_i,$$

and thus

$$\begin{aligned}|(\nabla\delta\Psi(x)-\nabla\delta\Psi(x'))\cdot v|^2 &\le \sum_{i=1}^N(\frac{d\delta\psi}{dx}(x_i)-\frac{d\delta\psi}{dx}(x_i'))^2\sum_{i=1}^N v_i^2\\ &\le \sup_{\mathbb{R}}|\frac{d^2\delta\psi}{dx^2}|^2\sum_{i=1}^N(x_i-x_i')^2\sum_{i=1}^N v_i^2\\ &= \sup_{\mathbb{R}}|\frac{d^2\delta\psi}{dx^2}|^2\,|x-x'|^2\,|v|^2.\end{aligned}$$

This implies

$$|\nabla\delta\Psi(x) - \nabla\delta\Psi(x')|^2 \;\leq\; \sup_{\mathbb{R}} |\frac{d^2\delta\psi}{dx^2}|^2 \, |x-x'|^2,$$

so that (42), after taking the infimum over all admissible π, turns into

$$\left|\int_{X_{N,K,y}} \nabla\delta\Psi \, d(\mu_y - \nu_y)\right|^2 \;\leq\; \sup_{\mathbb{R}} |\frac{d^2\delta\psi}{dx^2}|^2 \, d_{X_{N,K,y}}(\mu_y,\nu_y)^2. \tag{43}$$

Inserting (43) into (41) yields the lemma.

□

Proof of Lemma 10. According to Lemma 8, ν_y satisfies the LSI with constant $\frac{1}{2}C$, where $C = \exp(K \operatorname{osc}_{\mathbb{R}} \delta\psi)$. According to [OV, Theorem 1] ν_y satisfies the Talagrand inequality with constant $2C$, that is

$$\forall \mu_y \in \mathcal{P}(X_{N,K,y}) \quad d_{X_{N,K,y}}(\mu_y,\nu_y)^2 \;\leq\; 2C \int_{X_{N,K,y}} \log \frac{d\mu_y}{d\nu_y} \, d\mu_y.$$

Using once more the LSI from Lemma 8, we obtain the result.

□

5 Proof of Proposition 1

We prove Proposition 1 by using the representation from Lemma 4:

$$\exp\big(N\,(\bar\psi(m) - \psi_N(m))\big) \;=\; \frac{dg_{N,m}}{d\mathcal{L}^1}(0) \tag{44}$$

and developing the uniform bounds

$$\frac{1}{C} \;\leq\; \frac{dg_{N,m}}{d\mathcal{L}^1}(0) \;\leq\; C, \qquad \left|\frac{d^2}{dm^2}\frac{dg_{N,m}}{d\mathcal{L}^1}(0)\right| \;\leq\; C, \tag{45}$$

where C denotes a generic constant independent of both m and N. Notice that by interpolation, it also follows that

$$\left|\frac{d}{dm}\frac{dg_{N,m}}{d\mathcal{L}^1}(0)\right| \;\leq\; C.$$

Therefore, taking derivatives in (44) and applying the bounds, we deduce the uniform convergence of $\frac{d^2\psi_N}{dm^2}$ as $N \uparrow \infty$.

To begin, recall that $g_{N,m}$ describes the distribution of a sum of independent variables. Hence $\frac{dg_{N,m}}{d\mathcal{L}^1}$ can be written as a convolution. Defining the Fourier transform as

$$\mathcal{F}[f](\xi) := \int_{\mathbb{R}} \exp(i\xi x) f(x) \, \mathcal{L}^1(dx)$$

and recalling that convolution turns into multiplication under the Fourier transform, we can re–express the right–hand side of (44) as

$$\begin{aligned}\frac{dg_{N,m}}{d\mathcal{L}^1}(0) &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\left[\frac{dg_{N,m}}{d\mathcal{L}^1}\right](\xi)\,\mathcal{L}^1(d\xi) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} h^N(m, N^{-1/2}\xi)\,\mathcal{L}^1(d\xi), \qquad (46)\end{aligned}$$

where

$$\begin{aligned}h(m,\xi) \;&:=\; \exp(-i\xi m)\,\mathcal{F}\left[\frac{d\mu_m}{d\mathcal{L}^1}\right](\xi) \\ &\overset{(17)}{=} \int_{\mathbb{R}} \exp\big(-i\xi m + i\xi x - \bar{\psi}^*(\lambda) + \lambda x - \psi(x)\big)\,\mathcal{L}^1(dx). \qquad (47)\end{aligned}$$

Using (46) and the non–negativity of $g_{N,m}$, it follows that (45) is proved once we establish:

$$\frac{1}{C} \;\le\; \Big|\int_{\mathbb{R}} h^N(m, N^{-1/2}\xi)\,\mathcal{L}^1(d\xi)\Big| \;\le\; C, \qquad (48)$$

and

$$\Big|\frac{d^2}{dm^2}\int_{\mathbb{R}} h^N(m, N^{-1/2}\xi)\,\mathcal{L}^1(d\xi)\Big| \;\le\; C. \qquad (49)$$

We will establish (48) and (49) by splitting the integrals into "inner" integrals over $\{N^{-1/2}|\xi| \le \delta\}$ and "outer" integrals over the complement. More precisely, on the one hand we show that there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$ and all $m \in \mathbb{R}$,

$$\begin{aligned}\Big|\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N(m, N^{-1/2}\xi)\mathcal{L}^1(d\xi)\Big| \;&\le\; C, \qquad (50)\\ \mathrm{Re}\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N(m, N^{-1/2}\xi)\mathcal{L}^1(d\xi) \;&\ge\; 1/C, \qquad (51)\\ \Big|\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N(m, N^{-1/2}\xi)\mathcal{L}^1(d\xi)\Big| \;&\le\; C. \qquad (52)\end{aligned}$$

On the other hand, we will argue that for any $\delta > 0$, we have

$$\begin{aligned}\lim_{N\uparrow\infty}\int_{\{N^{-1/2}|\xi|\ge\delta\}} h^N(m, N^{-1/2}\xi)\mathcal{L}^1(d\xi) \;&=\; 0, \qquad (53)\\ \lim_{N\uparrow\infty}\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\ge\delta\}} h^N(m, N^{-1/2}\xi)\mathcal{L}^1(d\xi) \;&=\; 0, \qquad (54)\end{aligned}$$

uniformly in m. The combination of (50)–(54) yields (48) and (49).

First consider the outer integrals. We control h and its derivatives using:

Lemma 11. *For h defined by (47) and any $\delta > 0$, there exists a positive constant C_δ (uniform in m) such that for all $|\xi| > \delta$:*

i)

$$|h(m,\xi)| \;\leq\; \frac{1}{1+|\xi|/C_\delta},$$

ii)

$$\left|\frac{\partial h}{\partial m}(m,\xi)\right| \;\leq\; C_\delta|\xi|,$$

iii)

$$\left|\frac{\partial^2 h}{\partial m^2}(m,\xi)\right| \;\leq\; C_\delta|\xi|^2.$$

Lemma 11 i) implies (53):

$$\begin{aligned}
&\Big|\int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N(m,N^{-1/2}\xi)\,\mathcal{L}^1(d\xi)\Big| \\
&\qquad = N^{1/2}\left|\int_{\{|\hat\xi|\geq\delta\}} h^N(m,\hat\xi)\mathcal{L}^1(d\hat\xi)\right| \\
&\qquad \leq N^{1/2}\left(\frac{1}{1+\delta/C_\delta}\right)^{N-2}\int_{\{|\hat\xi|\geq\delta\}}\left(\frac{1}{1+|\hat\xi|/C_\delta}\right)^2\mathcal{L}(d\hat\xi) \underset{N\uparrow\infty}{\rightarrow} 0.
\end{aligned}$$

For (54) we notice that

$$\begin{aligned}
&\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N(m,N^{-1/2}\xi)\mathcal{L}^1(d\xi) \\
&= \int_{\{N^{-1/2}|\xi|\geq\delta\}} N^2 h^{N-2}(m,N^{-1/2}\xi)\left(\frac{\partial h}{\partial m}(m,N^{-1/2}\xi)\right)^2\mathcal{L}^1(d\xi) \\
&+ \int_{\{N^{-1/2}|\xi|\geq\delta\}} N h^{N-1}(m,N^{-1/2}\xi)\frac{\partial^2 h}{\partial m^2}(m,N^{-1/2}\xi)\,\mathcal{L}^1(d\xi),
\end{aligned}$$

so that by Lemma 11 ii) and iii),

$$\begin{aligned}
&\left|\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N(m,N^{-1/2}\xi)\mathcal{L}^1(d\xi)\right| \\
&\qquad\qquad \leq C_\delta\int_{\{N^{-1/2}|\xi|\geq\delta\}} N^2|h(m,N^{-1/2}\xi)|^{N-2}|\xi|^2\mathcal{L}^1(d\xi) \\
&\qquad\qquad = C_\delta\,N^{5/2}\int_{\{|\hat\xi|\geq\delta\}}|h(m,\hat\xi)|^{N-2}|\hat\xi|^2\mathcal{L}^1(d\hat\xi).
\end{aligned}$$

We appeal once more to Lemma 11 i) to conclude

$$\left|\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N(m,N^{-1/2}\xi)\mathcal{L}^1(d\xi)\right|$$
$$\leq C_\delta\, N^{5/2}\left(\frac{1}{1+\delta/C_\delta}\right)^{N-6}\int_{\{|\hat\xi|\geq\delta\}}\left(\frac{1}{1+|\hat\xi|/C_\delta}\right)^4|\hat\xi|^2\mathcal{L}^1(d\hat\xi)\underset{N\uparrow\infty}{\to}0,$$

establishing (54).

We now turn to the inner integrals. Since μ_m is a probability measure with mean m, cf. (12), we have

$$h(m,0) \;=\; 1,\quad \frac{\partial h}{\partial\xi}(m,0)=0,\quad\text{and} \tag{55}$$

$$-\frac{\partial^2 h}{\partial\xi^2}(m,0) \;=\; \int(x-m)^2\mu_m(dx) \;=\; \mathrm{Var}(\mu_m)>0.$$

According to Lemma 13 ii) in Subsection 5.1, the variance of μ_m is bounded uniformly above and below:

$$1/C \;\leq\; \mathrm{Var}(\mu_m) \;\leq\; C. \tag{56}$$

It follows from the lower bound and Taylor's theorem (see also the proof of Lemma 12 in Subsection 5.1) that there exists $h_2(m,\xi)$ defined on a uniform δ–neighborhood of $\xi=0$ such that

$$h(m,\xi)=\exp(-\xi^2 h_2(m,\xi)), \tag{57}$$

and

$$h_2(m,0) \;=\; \mathrm{Var}(\mu_m). \tag{58}$$

The motivation for introducing h_2 is the formula

$$h(m,N^{-1/2}\xi)^N=\exp(-\xi^2 h_2(m,N^{-1/2}\xi)). \tag{59}$$

The necessary control on h_2 is given by:

Lemma 12. *There exist $\delta>0$ and $C<\infty$ (uniform in m) such that for $|\xi|\leq\delta$ and all $m\in\mathbb{R}$:*

i)

$$\left|\frac{\partial h_2}{\partial\xi}(m,\xi)\right| \;\leq\; C,$$

ii)

$$\left|\frac{\partial h_2}{\partial m}(m,\xi)\right| \;\leq\; C,$$

iii)

$$\left|\frac{\partial^2 h_2}{\partial m^2}(m,\xi)\right| \;\le\; C.$$

Equipped with Lemma 12, we will now establish (50)–(51). In view of (59), we have

$$\begin{aligned}\int_{\{N^{-1/2}|\xi|\le\delta\}} & h(m,N^{-1/2}\xi)^N\,\mathcal{L}^1(d\xi)\\ &= \int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\left(-\xi^2\,h_2(m,N^{-1/2}\xi)\right)\mathcal{L}^1(d\xi).\end{aligned} \tag{60}$$

According to (56) and Lemma 12 i), we have for $|\hat{\xi}|\le\delta$

$$\mathrm{Re}\,h_2(m,\hat{\xi}) \;\ge\; 1/C.$$

Thus, for $N^{-1/2}|\xi|\le\delta$, we have

$$|\exp(-\xi^2 h_2(m,N^{-1/2}\xi))| \;\le\; \exp(-\xi^2/C), \tag{61}$$

so that

$$\left|\int_{\{N^{-1/2}|\xi|\le\delta\}} \exp(-\xi^2 h_2(m,N^{-1/2}\xi))\mathcal{L}^1(d\xi)\right| \;\le\; C.$$

In view of (60), this proves (50).
The proof of (52) is similar. Applying (60), we have

$$\begin{aligned}&\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\le\delta\}} h(m,N^{-1/2}\xi)^N\,\mathcal{L}^1(d\xi)\\ &= \int_{\{N^{-1/2}|\xi|\le\delta\}} -\xi^2\frac{\partial^2 h_2}{\partial m^2}(m,N^{-1/2}\xi)\exp(-\xi^2\,h_2(m,N^{-1/2}\xi))\mathcal{L}^1(d\xi)\\ &\quad + \int_{\{N^{-1/2}|\xi|\le\delta\}} \xi^4\Big(\frac{\partial h_2}{\partial m}\Big)^2(m,N^{-1/2}\xi)\exp(-\xi^2\,h_2(m,N^{-1/2}\xi))\mathcal{L}^1(d\xi).\end{aligned}$$

According to Lemma 12 ii) and iii) and (61), this identity yields the estimate

$$\begin{aligned}&\left|\frac{d^2}{dm^2}\int_{\{N^{-1/2}|\xi|\le\delta\}} h(m,N^{-1/2}\xi)^N\,\mathcal{L}^1(d\xi)\right|\\ &\qquad\le\; C\int_{\{N^{-1/2}|\xi|\le\delta\}} (\xi^2+\xi^4)\exp(-\xi^2/C)\mathcal{L}^1(d\xi)\le C.\end{aligned}$$

Finally, consider (51). It will be convenient to introduce h_3 via

$$h_2(m,\hat{\xi}) \;=\; h_2(m,0)+\hat{\xi}h_3(m,\hat{\xi}), \tag{62}$$

which, according to Taylor and Lemma 12 i), satisfies

$$\sup_{|\hat{\xi}|\le\delta}|h_3(m,\hat{\xi})| \;\le\; \sup_{|\hat{\xi}|\le\delta}\left|\frac{\partial h_2}{\partial \xi}(m,\hat{\xi})\right| \;\le\; C. \tag{63}$$

By the definition of h_3 in (62), we have

$$\begin{aligned} & \exp(-\xi^2 h_2(m, N^{-1/2}\xi)) - \exp(-\xi^2 h_2(m,0)) \\ = \quad & \exp(-\xi^2 h_2(m,0)) \left(\exp(-N^{-1/2}\xi^3 h_3(m, N^{-1/2}\xi)) - 1\right) \\ \stackrel{(58)}{=} \quad & \exp(-\xi^2 \mathrm{Var}(\mu_m)) \left(\exp(-N^{-1/2}\xi^3 h_3(m, N^{-1/2}\xi)) - 1\right). \end{aligned} \tag{64}$$

We use the fact:

$$|\exp(z) - 1| = \left|\sum_{j=1}^{\infty} \frac{z^j}{j!}\right| \leq \sum_{j=1}^{\infty} \frac{|z|^j}{j!} = \exp(|z|) - 1,$$

with

$$z = -N^{-1/2}\,\xi^3\,h_3(m, \hat{\xi})$$

to conclude from (64) that

$$\begin{aligned} & \left|\exp(-\xi^2 h_2(m, N^{-1/2}\xi)) - \exp(-\xi^2 h_2(m,0))\right| \\ & \qquad \leq \exp(-\xi^2 \mathrm{Var}(\mu_m)) \Big(\exp\left(N^{1/2}|\xi|^3 |h_3(m, N^{-1/2}\xi)|\right) - 1\Big). \end{aligned}$$

Together with (56) and (63), this yields for ξ with $N^{-1/2}|\xi| \leq \delta$:

$$\left|\exp(-\xi^2 h_2(m, N^{-1/2}\xi)) - \exp(-\xi^2 h_2(m,0))\right| \leq \exp(-\xi^2/C)(\exp(C\delta\xi^2) - 1).$$

Hence, for δ sufficiently small,

$$\begin{aligned} & \left|\int_{\{N^{-1/2}|\xi|\leq\delta\}} \exp(-\xi^2 h_2(m, N^{-1/2}\xi)) - \exp(-\xi^2 h_2(m,0))\mathcal{L}^1(d\xi)\right| \\ & \qquad \leq \int_{\mathbb{R}} \exp(-\xi^2(1/C - C\delta)) - \exp(-\xi^2/C)\mathcal{L}^1(d\xi) \\ & \qquad = C\left(\frac{1}{\sqrt{1/C - C\delta}} - \frac{1}{\sqrt{1/C}}\right) \\ & \qquad \leq C\delta. \end{aligned} \tag{65}$$

On the other hand, we have by (58) and (56) that

$$\exp(-\xi^2 h_2(m,0)) = \exp(-\xi^2 \mathrm{Var}(\mu_m)) \geq \exp(-\xi^2/C),$$

so that

$$\begin{aligned} \int_{\{N^{-1/2}|\xi|\leq\delta\}} \exp(-\xi^2 h_2(m,0))\mathcal{L}^1(d\xi) & \geq \int_{\{N^{-1/2}|\xi|\leq\delta\}} \exp(-\xi^2/C)\mathcal{L}^1(d\xi) \\ & \geq 1/C - C\exp(-N^{1/2}\delta/C). \end{aligned} \tag{66}$$

The combination of (65) and (66) yields

$$\begin{aligned} & \mathrm{Re} \int_{\{N^{-1/2}|\xi|\leq\delta\}} \exp(-\xi^2 h_2(m, N^{-1/2}\xi))\mathcal{L}^1(d\xi) \\ & \qquad \geq 1/C - C(\exp(-CN^{1/2}\delta) + \delta), \end{aligned}$$

which establishes (51) for δ sufficiently small and N sufficiently large.

□

5.1 Proofs of the lemmas

Before turning to the proofs, we collect a few ingredients which we will use repeatedly. First, recall that by assumption, μ_m is a perturbation of a shifted Gaussian. To be precise, letting

$$\frac{dg_\lambda}{d\mathcal{L}^1}(x) := (2\pi)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\lambda)^2\Big), \tag{67}$$

we may write μ_m as

$$\begin{aligned}\frac{d\mu_m}{d\mathcal{L}^1}(x) &= Z^{-1} \exp\left(-\delta\psi(x) + \lambda x - \frac{1}{2}x^2\right)\\ &= \tilde{Z}^{-1} \exp\left(-\delta\psi(x) - \frac{1}{2}(x-\lambda)^2\right),\end{aligned}$$

and observe that

$$\exp\big(-\operatorname{osc}_{\mathbb{R}}\delta\psi\big)\frac{dg_\lambda}{d\mathcal{L}^1}(x) \le \frac{d\mu_m}{d\mathcal{L}^1}(x) \le \exp\big(\operatorname{osc}_{\mathbb{R}}\delta\psi\big)\frac{dg_\lambda}{d\mathcal{L}^1}(x). \tag{68}$$

A second elementary but important observation is that the mean of a measure μ is optimal in the sense that for all $c \in \mathbb{R}$,

$$\begin{aligned}\int_{\mathbb{R}}(x-c)^2\mu(dx) &= \int_{\mathbb{R}} x^2\mu(dx) - 2c\int_{\mathbb{R}} x\mu(dx) + c^2\\ &\ge \int_{\mathbb{R}} x^2\mu(dx) - \left(\int_{\mathbb{R}} x\mu(dx)\right)^2\\ &= \int_{\mathbb{R}}\left(x - \int_{\mathbb{R}} y\mu(dy)\right)^2\mu(dx).\end{aligned} \tag{69}$$

Finally, we state and prove a lemma about the map between m and λ (cf. Subsection 2.6) which is useful in the proofs of Lemmas 11 and 12. Here and below, we refer to the measure μ_m as μ_λ in order to emphasize the λ–dependence.

Lemma 13. *Consider the change of variables*

$$m = \frac{d\bar{\psi}^*}{d\lambda}(\lambda) \tag{70}$$

and the corresponding measure $\mu_m = \mu_\lambda \in \mathcal{P}(\mathbb{R})$ *with density*

$$\frac{d\mu_\lambda}{d\mathcal{L}^1}(x) = \exp\big(-\bar{\psi}^*(\lambda) + \lambda x - \psi(x)\big)$$

and mean $\int_{\mathbb{R}} x\mu_\lambda(dx) = m$*, cf.* (13), (16) *and* (12)*. Then:*

i) The first two derivatives of m *are related to the moments of* μ_λ *as:*

$$\begin{aligned}\frac{dm}{d\lambda} &= \frac{d^2\bar{\psi}^*}{d\lambda^2} = \int_{\mathbb{R}}(x-m)^2\mu_\lambda(dx),\\ \frac{d^2m}{d\lambda^2} &= \frac{d^3\bar{\psi}^*}{d\lambda^3} = \int_{\mathbb{R}}(x-m)^3\mu_\lambda(dx).\end{aligned}$$

ii) The moments of μ_λ satisfy the uniform bounds:

$$\frac{1}{C} \leq \int_{\mathbb{R}} (x-m)^2 \mu_\lambda(dx) \leq C,$$
$$\Big| \int_{\mathbb{R}} (x-m)^3 \mu_\lambda(dx) \Big| \leq C,$$
$$\int_{\mathbb{R}} (x-m)^4 \mu_\lambda(dx) \leq C.$$

iii) The second derivatives of the inverse map are uniformly bounded:

$$\left| \frac{d^2\lambda}{dm^2} \right| \leq C.$$

iv) The map is uniformly close to the identity: $|\lambda - m| \leq C$.

Proof of Lemma 13.
To show the equalities in i), we first notice that for the variance we have from (12) and (16)

$$\begin{aligned} \frac{dm}{d\lambda} &= \frac{d}{d\lambda} \int_{\mathbb{R}} x \exp(-\bar{\psi}^*(\lambda) + \lambda x - \psi(x)) \mathcal{L}^1(dx) \\ &= \int_{\mathbb{R}} x(x-m)\mu_\lambda(dx) = \int_{\mathbb{R}} x^2 \mu_\lambda(dx) - m^2 \\ &= \int_{\mathbb{R}} (x-m)^2 \mu_\lambda(dx). \end{aligned} \tag{71}$$

Together with (13), this establishes the first equality of i). For the second equality, we take a derivative in (71) and notice that because μ_λ has mean m,

$$\begin{aligned} \frac{d^2 m}{d\lambda^2} &\overset{(70),(16)}{=} \int_{\mathbb{R}} (x-m)^3 \mu_\lambda(dx) - 2\frac{dm}{d\lambda} \int_{\mathbb{R}} (x-m)\mu_\lambda(dx) \\ &= \int_{\mathbb{R}} (x-m)^3 \mu_\lambda(dx). \end{aligned}$$

Next we prove point iv), which follows from

$$\begin{aligned} |\lambda - m|^2 &= \Big| \int_{\mathbb{R}} (\lambda - x)\mu_\lambda(dx) \Big|^2 \leq \int_{\mathbb{R}} (\lambda - x)^2 \mu_\lambda(dx) \\ &\overset{(68)}{\leq} \exp(\operatorname{osc}_{\mathbb{R}} \delta\psi) \int_{\mathbb{R}} (\lambda - x)^2 g_\lambda(dx) \leq \exp(\operatorname{osc}_{\mathbb{R}} \delta\psi). \end{aligned}$$

Turning to the first estimate in ii), we observe that on the one hand,

$$\begin{aligned} \int_{\mathbb{R}} (x-m)^2 \mu_\lambda(dx) &\overset{(68)}{\geq} \exp(-\operatorname{osc}_{\mathbb{R}} \delta\psi) \int_{\mathbb{R}} (x-m)^2 g_\lambda(dx) \\ &\overset{(69)}{\geq} \exp(-\operatorname{osc}_{\mathbb{R}} \delta\psi) \int_{\mathbb{R}} \left(x - \int_{\mathbb{R}} y g_\lambda(dy) \right)^2 g_\lambda(dx) \\ &= \exp(-\operatorname{osc}_{\mathbb{R}} \delta\psi). \end{aligned}$$

On the other hand,

$$\begin{aligned}\int_{\mathbb{R}}(x-m)^2\mu_\lambda(dx) &\overset{(69)}{\leq} \int_{\mathbb{R}}\Big(x-\int_{\mathbb{R}} y g_\lambda(dy)\Big)^2\mu_\lambda(dx)\\ &\overset{(68)}{\leq} \exp(\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}\Big(x-\int_{\mathbb{R}} y g_\lambda(dy)\Big)^2 g_\lambda(dx)\\ &= \exp(\operatorname{osc}_{\mathbb{R}}\delta\psi).\end{aligned}$$

The bound in ii) on the fourth moment follows from:

$$\begin{aligned}\int_{\mathbb{R}}(x-m)^4\mu_\lambda(dx) &\leq C\Big(\int_{\mathbb{R}}(x-\lambda)^4\mu_\lambda(dx)+\int_{\mathbb{R}}(\lambda-m)^4\mu_\lambda(dx)\Big)\\ &\overset{(68)}{\leq} C\Big(\exp(\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}(x-\lambda)^4 g_\lambda(dx)+\int_{\mathbb{R}}(\lambda-m)^4\mu_\lambda(dx)\Big)\\ &\leq C,\end{aligned}$$

by iv) and the definition (67) of g_λ. Hölder's inequality then implies the bound on the third moment.

Finally, iii) follows immediately from i), ii), and:

$$\frac{d^2\lambda}{dm^2} = \frac{d}{d\lambda}\left(\frac{d^2\bar{\psi}^*}{d\lambda^2}\right)^{-1}\frac{d\lambda}{dm} = -\frac{d^3\bar{\psi}^*}{d\lambda^3}\left(\frac{d^2\bar{\psi}^*}{d\lambda^2}\right)^{-3}.$$

□

Proof of Lemma 11. We prove i) by splitting it into two pieces. First we bound h by a constant smaller than one, uniformly in m for ξ bounded away from zero. Then we show the decay for large ξ. By (47) and (11) we have

$$h(m,\xi) = \exp(-i\xi m)\int_{\mathbb{R}}\exp(i\xi x)\mu_\lambda(dx).$$

Thus

$$\begin{aligned}|h(m,\xi)|^2 &= \Big|\int_{\mathbb{R}}\cos(\xi x)\mu_\lambda(dx) + i\int_{\mathbb{R}}\sin(\xi x)\mu_\lambda(dx)\Big|^2\\ &= \Big(\int_{\mathbb{R}}\cos(\xi x)\mu_\lambda(dx)\Big)^2+\Big(\int_{\mathbb{R}}\sin(\xi x)\mu_\lambda(dx)\Big)^2\\ &= 1-\Big(\int_{\mathbb{R}}\cos^2(\xi x)\mu_\lambda(dx)-\Big(\int_{\mathbb{R}}\cos(\xi x)\mu_\lambda(dx)\Big)^2\Big)\\ &\quad -\Big(\int_{\mathbb{R}}\sin^2(\xi x)\mu_\lambda(dx)-\Big(\int_{\mathbb{R}}\sin(\xi x)\mu_\lambda(dx)\Big)^2\Big)\\ &=: 1-\operatorname{Var}_{\mu_\lambda}(\cos(\xi x))-\operatorname{Var}_{\mu_\lambda}(\sin(\xi x)).\end{aligned} \tag{72}$$

Therefore, to bound h by a constant smaller than one, we need to bound the variances away from zero. Recalling the elementary observations (68) and (69), we have:

$$\begin{aligned}
&\mathrm{Var}_{\mu_\lambda}(\cos(\xi x)) \\
&= \int_{\mathbb{R}} \Big(\cos(\xi x) - \int_{\mathbb{R}} \cos(\xi y)\mu_\lambda(dy)\Big)^2 \mu_\lambda(dx) \\
&\geq \exp(-\mathrm{osc}_{\mathbb{R}}\, \delta\psi) \int_{\mathbb{R}} \Big(\cos(\xi x) - \int_{\mathbb{R}} \cos(\xi y)\mu_\lambda(dy)\Big)^2 g_\lambda(dx) \\
&\geq \exp(-\mathrm{osc}_{\mathbb{R}}\delta\psi)\Big(\int_{\mathbb{R}} \cos^2(\xi x) g_\lambda(dx) - \Big(\int_{\mathbb{R}} \cos(\xi x) g_\lambda(dx)\Big)^2\Big).
\end{aligned} \tag{73}$$

Since the Fourier transform of a Gaussian is again Gaussian, the right–hand side of (73) can be computed explicitly. Looking at the second integral, we have

$$\begin{aligned}
&\Big(\int_{\mathbb{R}} \cos(\xi x) g_\lambda(dx)\Big)^2 \\
&= \frac{1}{4}\left((2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \left(e^{i\xi x} + e^{-i\xi x}\right) e^{-\frac{1}{2}(x-\lambda)^2} \mathcal{L}^1(dx)\right)^2 \\
&= \frac{1}{4}\left(e^{i\xi\lambda} (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{i\xi y} e^{-\frac{1}{2}y^2} \mathcal{L}^1(dy) + e^{-i\xi\lambda} (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-i\xi y} e^{-\frac{1}{2}y^2} \mathcal{L}^1(dy)\right)^2 \\
&= \frac{1}{4}\left(e^{i\xi\lambda} e^{-\frac{1}{2}\xi^2} + e^{-i\xi\lambda} e^{-\frac{1}{2}\xi^2}\right)^2 \\
&= \frac{1}{4}\left(e^{2i\xi\lambda} e^{-\xi^2} + e^{-2i\xi\lambda} e^{-\xi^2} + 2e^{-\xi^2}\right) = \frac{1}{2}\big(\cos(2\xi\lambda) + 1\big)\, e^{-\xi^2}.
\end{aligned}$$

The second part of the right–hand side of (73) can be computed similarly. We get:

$$\int_{\mathbb{R}} \cos^2(\xi x) g_\lambda(dx) = \frac{1}{2}\,\big(\cos(2\xi\lambda)e^{-2\xi^2} + 1\big).$$

Therefore,

$$\begin{aligned}
\int_{\mathbb{R}} \cos^2(\xi x) g_\lambda(dx) - \Big(\int_{\mathbb{R}} \cos(\xi x) g_\lambda(dx)\Big)^2 \\
&= \frac{1}{2}\left(1 - e^{-\xi^2}\cos(2\xi\lambda)\right)\left(1 - e^{-\xi^2}\right) \\
&\geq \frac{1}{2}\big(1 - e^{-\xi^2}\big)^2.
\end{aligned}$$

Inserting this into (73) and then (72) (the same inequality holds for $\mathrm{Var}_{\mu_\lambda}(\sin(\xi x))$), we obtain

$$|h(m,\xi)|^2 \leq 1 - \exp\big(-\mathrm{osc}_{\mathbb{R}}\,\delta\psi\big)\,\big(1 - e^{-\xi^2}\big)^2.$$

Hence for any $\delta > 0$ there exists a $C_\delta < \infty$ (uniform in m) such that

$$|h(m,\xi)| \leq 1 - \frac{1}{C_\delta} \qquad \text{for } |\xi| > \delta. \tag{74}$$

To complete the proof of Lemma 11 i), we need to establish decay of h for large $|\xi|$–values. This is done by an integration by parts argument; we have

$$\begin{aligned} h(m,\xi) &= Z^{-1}\int_{\mathbb{R}} \exp(i\xi x)\,\exp(\lambda x-\psi(x))\mathcal{L}^1(dx) \\ &= Z^{-1}\int_{\mathbb{R}} \frac{1}{i\xi}\exp(i\xi x)\left(\lambda - x - \frac{d\delta\psi}{dx}(x)\right)\,\exp(\lambda x - \frac{x^2}{2} - \delta\psi(x))\mathcal{L}^1(dx). \end{aligned}$$

This yields the estimate

$$\begin{aligned} &|h(m,\xi)| \\ &\leq\ |\xi|^{-1}\int_{\mathbb{R}}\left(|\lambda-x| + \left|\frac{d\delta\psi}{dx}(x)\right|\right)\,\mu_\lambda(dx) \\ &\overset{(68)}{\leq}\ |\xi|^{-1}\exp(\mathrm{osc}_{\mathbb{R}}\,\delta\psi)\int_{\mathbb{R}}\left(|\lambda-x| + \left|\frac{d\delta\psi}{dx}(x)\right|\right)\,g_\lambda(dx) \\ &\leq\ |\xi|^{-1}\exp(\mathrm{osc}_{\mathbb{R}}\,\delta\psi)\left((2\pi)^{-\frac{1}{2}}\int_{\mathbb{R}}|y|e^{-\frac{1}{2}y^2}\mathcal{L}^1(dy) + \sup_{\mathbb{R}}\left|\frac{d\delta\psi}{dx}(x)\right|\right). \end{aligned} \tag{75}$$

Since by elementary interpolation

$$\sup_{\mathbb{R}}\left|\frac{d\delta\psi}{dx}(x)\right| \leq C\sup_{\mathbb{R}}|\delta\psi(x)|\ \sup_{\mathbb{R}}\left|\frac{d^2\delta\psi}{dx^2}(x)\right| \overset{(5),(6)}{<} \infty,$$

we infer from (75) that

$$|h(m,\xi)|\ \leq C\,|\xi|^{-1}. \tag{76}$$

The combination of (74) and (76) yields Lemma 11 i).

We turn now to the estimates for ii) and iii). Since $|\xi|\leq\delta$, it suffices to prove

$$\begin{aligned} \left|\frac{\partial h}{\partial m}\right| &\leq\ C(1+|\xi|), \\ \left|\frac{\partial^2 h}{\partial m^2}\right| &\leq\ C(1+|\xi|^2). \end{aligned}$$

We appeal to the change of variables (69):

$$\begin{aligned} \frac{\partial h}{\partial m} &= \frac{\partial h}{\partial \lambda}\frac{d\lambda}{dm} \\ \text{and}\quad \frac{\partial^2 h}{\partial m^2} &= \frac{\partial^2 h}{\partial\lambda^2}\left(\frac{d\lambda}{dm}\right)^2 + \frac{\partial h}{\partial\lambda}\,\frac{d^2\lambda}{dm^2}. \end{aligned} \tag{77}$$

According to Lemma 13 i), ii), and iii), it thus suffices to prove

$$\begin{aligned} \left|\frac{\partial h}{\partial \lambda}\right| &\leq\ C(1+|\xi|), \\ \left|\frac{\partial^2 h}{\partial \lambda^2}\right| &\leq\ C(1+|\xi|^2). \end{aligned} \tag{78}$$

The starting point is formula (47):

$$
\begin{aligned}
h &= \int_{\mathbb{R}} \exp(i\xi(x-m))\mu_\lambda(dx) \\
&= \int_{\mathbb{R}} \exp\left(i\xi x - i\xi m - \bar{\psi}^*(\lambda) + \lambda x - \psi(x)\right) \mathcal{L}^1(dx).
\end{aligned} \tag{79}
$$

Using (70), we infer the identities

$$
\frac{\partial h}{\partial \lambda} = \int_{\mathbb{R}} \left(-i\xi \frac{dm}{d\lambda} + x - m\right) \exp(i\xi(x-m))\,\mu_\lambda(dx), \tag{80}
$$

$$
\begin{aligned}
\frac{\partial^2 h}{\partial \lambda^2} &= \int_{\mathbb{R}} \left(-i\xi \frac{d^2 m}{d\lambda^2} - \frac{dm}{d\lambda}\right) \exp(i\xi(x-m))\,\mu_\lambda(dx) \\
&\quad + \int_{\mathbb{R}} \left(-i\xi \frac{dm}{d\lambda} + x - m\right)^2 \exp(i\xi(x-m))\,\mu_\lambda(dx).
\end{aligned} \tag{81}
$$

By Jensen's inequality and (12), these yield the inequalities

$$
\begin{aligned}
\left|\frac{\partial h}{\partial \lambda}\right|^2 &\le \int_{\mathbb{R}} \left| -i\xi \frac{dm}{d\lambda} + x - m\right|^2 \mu_\lambda(dx) \\
&= \xi^2 \left|\frac{dm}{d\lambda}\right|^2 + \int_{\mathbb{R}} (x-m)^2 \mu_\lambda(dx),
\end{aligned}
$$

and

$$
\begin{aligned}
\left|\frac{\partial^2 h}{\partial \lambda^2}\right| &\le |\xi| \left|\frac{d^2 m}{d\lambda^2}\right| + \left|\frac{dm}{d\lambda}\right| + \int_{\mathbb{R}} \left| -i\xi \frac{dm}{d\lambda} + x - m\right|^2 \mu_\lambda(dx) \\
&= |\xi| \left|\frac{d^2 m}{d\lambda^2}\right| + \left|\frac{dm}{d\lambda}\right| + \xi^2 \left|\frac{dm}{d\lambda}\right|^2 + \int_{\mathbb{R}} (x-m)^2 \mu_\lambda(dx).
\end{aligned}
$$

Hence (78) follows from Lemma 13 i) and ii).

□

Proof of Lemma 12.

Since $h(m,0) = 1$ and $\frac{\partial h}{\partial \xi}(m,0) = 0$, cf. (55), we may introduce by Taylor the complex–valued function $h_1(m,\xi)$ by

$$
h(m,\xi) = 1 - \xi^2\, h_1(m,\xi) = \exp(-\xi^2\, h_2(m,\xi)), \tag{82}
$$

so that

$$
h_2(m,\xi) = \begin{cases} -\xi^{-2} \log\left(1 - \xi^2\, h_1(m,\xi)\right) & \xi \neq 0 \\ h_1(m,0) & \xi = 0. \end{cases} \tag{83}
$$

We claim that Lemma 12 is a consequence of the following bounds on h_1:

$$|h_1(m,\xi)| \le C, \qquad \left|\frac{\partial h_1}{\partial \xi}\right| \le C, \tag{84}$$

$$\left|\frac{\partial h_1}{\partial m}\right| \le C, \qquad \left|\frac{\partial^2 h_1}{\partial m^2}\right| \le C. \tag{85}$$

Indeed, Lemma 12 i) follows from (84) after rewriting (83) in the form:

$$h_2(m,\xi) = h_1(m,\xi)\, f(\xi^2\, h_1(m,\xi)),$$

for the function $f(z) = -z^{-1}\log(1-z)$ which is smooth in a neighborhood of zero. Points ii) and iii) follow from (85) via the chain rule applied to (83):

$$\frac{\partial h_2}{\partial m} = \frac{1}{1-\xi^2 h_1}\frac{\partial h_1}{\partial m}$$

and

$$\frac{\partial^2 h_2}{\partial m^2} = \frac{1}{1-\xi^2 h_1}\frac{\partial^2 h_1}{\partial m^2} + \frac{\xi^2}{(1-\xi^2 h_1)^2}\left(\frac{\partial h_1}{\partial m}\right)^2,$$

along with the fact that $|\xi| \le \delta$. As in the proof of Lemma 11, it will be convenient to consider derivatives with respect to λ instead of m. By (77) and Lemma 13 i)–iii), we can establish (85) by showing

$$\left|\frac{\partial h_1}{\partial \lambda}\right| \le C(1+|\xi|) \quad \text{and} \quad \left|\frac{\partial^2 h_1}{\partial \lambda^2}\right| \le C(1+|\xi|^2). \tag{86}$$

In view of definition (82) which can be reformulated as

$$h_1(m,\xi) = \frac{1}{\xi^2}\int_0^\xi (\xi'-\xi)\frac{\partial^2 h}{\partial \xi^2}(m,\xi')\,d\xi',$$

(84) and (86) are consequences of

$$\left|\frac{\partial^2 h}{\partial \xi^2}\right| \le C, \qquad \left|\frac{\partial^3 h}{\partial \xi^3}\right| \le C, \tag{87}$$

$$\left|\frac{\partial^3 h}{\partial \xi^2 \partial \lambda}\right| \le C, \qquad \left|\frac{\partial^4 h}{\partial \xi^2 \partial^2 \lambda}\right| \le C. \tag{88}$$

The estimates (87) are easily established. We infer from (79)

$$\frac{\partial^k h}{\partial \xi^k}(m,\xi) = \int_{\mathbb{R}} \big(i(x-m)\big)^k \exp(i\xi(x-m))\mu_\lambda(dx).$$

Thus, (87)follows from Lemma 13 ii).
For (88) we turn to (80) and (81), which we write as

$$\frac{\partial h}{\partial \lambda} = \int_{\mathbb{R}} a_1(\lambda,\xi)\,\mu_\lambda(dx) \quad \text{and} \quad \frac{\partial^2 h}{\partial \lambda^2} = \int_{\mathbb{R}} a_2(\lambda,\xi)\,\mu_\lambda(dx),$$

where we set for abbreviation:

$$a_1(\lambda,\xi) = \left(-i\xi\frac{dm}{d\lambda} + x - m\right)\exp(i\xi(x-m))$$
$$a_2(\lambda,\xi) = \left(\big(-i\xi\frac{d^2m}{d\lambda^2} - \frac{dm}{d\lambda}\big) + \big(-i\xi\frac{dm}{d\lambda} + x - m\big)^2\right)\exp(i\xi(x-m)).$$

Since for $|\xi| \le \delta$ we have

$$\left|\frac{\partial^2 a_1}{\partial\xi^2}\right| \le C\left(\left|\frac{dm}{d\lambda}\right| + 1\right)\left(|x-m|^3 + 1\right)$$

and

$$\left|\frac{\partial^2 a_2}{\partial\xi^2}\right| \le C\left(\left|\frac{d^2m}{d\lambda^2}\right| + \left|\frac{dm}{d\lambda}\right|^2 + \left|\frac{dm}{d\lambda}\right| + 1\right)\left(|x-m|^4 + 1\right),$$

(88) follows from Lemma 13 i) and ii).

□

Acknowledgements

N. Grunewald was supported by the BMBF Förderbereich Mathematik. F. Otto was partially supported by the DFG through SFB 611. M.G. Reznikoff was supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

References

[AGS] L. Ambrosio, N. Gigli, and G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures,*Lectures in Mathematics ETH Zürich*, Birkhäuser Verlag, (2005).

[BE] D. Bakry and M. Emery, Diffusions Hypercontractives, *Sem. Probab. XIX, Lecture Notes in Math*, **1123**, Springer–Verlag, (1985), pp 177–206.

[BB] G. Blower and F. Bolley, Concentration of Measure on Product Spaces with Applications to Markov Processes, submitted for publication.

[BH] Th. Bodineau and B. Hellfer, On Log Sobolev Inequalities for Unbounded Spin Systems, *J. Funct. Anal.*, **166**, (1999), pp. 168–178.

[Cha] D. Chafaï, Glauber versus Kawasaki for Spectral Gap and Logarithmic Sobolev Inequalities of Some Unbounded Conservative Spin Systems, *Markov Processes Relat. Fields*, **9**, (2001), pp. 341–362.

[Cap] P. Caputo, Uniform Poincaré Inequalities for Unbounded Conservative Spin Systems: The Non–Interacting Case, *Stochastic Process. Appl.*, **106**, (2003), no. 2, pp. 223–244.

[E] R. S. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer–Verlag, (1985).

[F] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 2, Wiley, (1971).

[GZ] A. Guionnet and B. Zegarlinski, Lectures on Logarithmic Sobolev Inequalities, unpublished lecture notes.

[GPV] M.Z. Guo, G.C. Papanicolau, and S.R.S.Varadhan, Nonlinear Diffusion Limit for a System with Nearest Neighbor Interactions, *Commun. Math. Phys.*, **118**, (1988), pp. 31–59.

[H] B. Helffer, Remarks on Decay of Correlations and Witten Laplacians — Brascamp–Lieb Inequalities and Semi–classical Analysis, *Ann. Inst. H. Poincaré*, **35**, (1999), pp. 483–508.

[HS] R. Holley and D.Stroock, Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Statist. Phys.*, **46**, (1987), pp. 1159–1194.

[KL] C. Kipnis and C. Landim, Scaling Limits of Interacting Particle Systems, Springer–Verlag, (1999).

[LPY] C. Landim, G. Panizo, and H.T. Yau, Spectral Gap and Logarithmic Sobolev Inequality for Unbounded Conservative Spin Systems, *Ann. Inst. H. Poincaré Probab. Statist*, **38**, (2002), no. 5, pp. 739–777.

[L] M. Ledoux, Logarithmic Sobolev Inequalities for Unbounded Spin Systems Revisted, *Sem. Probab. XXXV*, Lecture Notes in Math, **1755**, Springer–Verlag, (2001), pp. 167–194.

[LY] S.L. Lu and H.T. Yau, Spectral Gap and Logarithmic Sobolev Inequality for Kawasaki and Glauber Dynamics, *Commun. Math. Phys.*, **156**, (1993), pp. 399–433.

[O] F. Otto, The Geometry of Dissipative Evolution Equations: The Porous Medium Equation, *Commun. Partial Differential Equations*, **26**, (2001), pp. 101–174.

[OV] F. Otto and C. Villani, Generalization of an Inequality By Talagrand and Links with the Logarithmic Sobolev Inequality, *J. Funct. Analysis*, **173**, (2000), pp. 361–400.

[R] G. Royer, Une Initiation aux Inégalités de Sobolev Logarithmiques, Cours Spécialisés, Soc. Math. de France, (1999).

[SZ] D. Stroock and B. Zegarlinski, The Equivalence of the Logarithmic Sobolev Inequality and the Dobrushin–Shlosman Mixing Condition, *Comm. Math. Phys.*, **144**, (1992), 303–323.

[Y] H.T. Yau, Logarithmic Sobolev Inequality for Lattice Gases with Mixing Conditions, *Commun. Math. Phys.*, **181**, (1996), pp. 367–408.

[Yo] N. Yoshida, The Log–Sobolev Inequality for Weakly Coupled Lattice Fields, *Probab. Theor. Relat. Fields*, **115**, (1999), pp. 1–40.

[Yo2] N. Yoshida, The Equivalence of the Log–Sobolev Inequality and a Mixing Condition for Unbounded Spin Systems on the Lattice, *Ann. Inst. H. Poincaré Probab. Statist.*, **37**, (2001), pp. 223–243.

[Z] B. Zegarlinski, The Strong Decay to Equilibrium for the Stochastic Dynamics of Unbounded Spin Systems on a Lattice, *Comm. Math. Phys.*, **175**, (1996), pp. 401–432.